# HIGHLIGHTS OF 2011

## Table of Contents

### FEATURES

**Editorial—2011 Highlights Issue**

The Blossoming of Biology Education Research  
Erin L. Dolan ................................................................. 1–2

**Letter to the Editor**

Implications for Undergraduate Education of Two Interdisciplinary Biological Sciences: Biochemistry and Biophysics  
Jonghoon Kang and Crystal B. Purnell ................................................................. 3–4

**Approaches to Biology Teaching and Learning**

Moving Theory into Practice: A Reflection on Teaching a Large, Introductory Biology Course for Majors  
Kimberly D. Tanner ................................................................. 5–14

**WWW.Life Sciences Education**

Where Do New Medicines Come From?  
Dennis Liu ................................................................. 15–21

**WWW.Life Sciences Education**

Life Sciences Career Exploration  
Louisa A. Stark ................................................................. 22–25

**From the National Science Foundation**

Undergraduate Research Experiences in Biology: Alternatives to the Apprenticeship Model  
Cynthia A. Wei and Terry Woodin ................................................................. 26–34

**Current Insights**

Recent Research in Science Teaching and Learning  
Deborah Allen ................................................................. 35–37

**Book Review**

A Life in Science  
Jay Brewster ................................................................. 38–39

**Meeting Report**

Society for the Advancement of Biology Education Research (SABER)  
Erika G. Offerdahl, Teri Balser, Clarissa Dirks, Kathryn Miller, Jennifer L. Momsen, Lisa Montplaisir, Marcy Osgood, Karen Sirum, Mary Pat Wenderoth, Brian White, William B. Wood, Michelle Withers, and Robin Wright ................................................................. 40–42

*Table of Contents Continued on Back Cover*
Bringing The World’s Best Biology To You...for FREE!

iBioSeminars
www.ibioseminars.org
Full-length seminars from leading scientists

iBioMagazine
www.ibiomagazine.org
5-15 minute videos on people, discoveries, and opinions

Educational Resources for Teachers and Students:

iBio 101
Videos for high school/undergrad

Short Clip Library
1-6 minute video clips

Teaching Tools
Educational resources for teachers

Bio Techniques
Lectures on methodology

75 Talks and growing...
View by web streaming or download

Sponsored by:
A searchable collection of images, videos, and animations of cells and cellular functions available to students, educators, and the interested public.

New features with optional free accounts:
- Save favorite images for later use
- Automatically find new images that match your areas of interest

Coming Soon!
- Annotate images
- Invite colleagues to connect and share images
- Create multiple folders for saving/sharing images
- Subscribe to receive emails when new matching images are available

The ASCB welcomes image submissions and feedback and encourages you to visit www.cellimagelibrary.org for more information. Or email David Orloff (dorloff@ascb.org) Manager, Image Library.

Editorial—2011 Highlights Issue

The Blossoming of Biology Education Research

Erin L. Dolan

Department of Biochemistry and Molecular Biology, University of Georgia, Athens, GA 30602

Welcome to the 2011 Highlights issue of CBE—Life Sciences Education (CBE-LSE)! This special issue is printed annually and circulated widely with multiple aims, including attracting a broader readership, prompting new authors to consider CBE-LSE as a venue for sharing their work, and providing exemplar articles for those who are new to the scholarship of teaching and learning. This year’s Highlights is a compilation of articles from the Spring, Summer, and Fall issues; articles from Winter 2011 will be considered for inclusion in the 2012 Highlights.

As we assembled this issue, I was prompted to reflect on the collective experiences of the journal, which has achieved several notable milestones in recent years. The journal’s name was changed from Cell Biology Education to CBE—Life Sciences Education to more accurately reflect its wider scope. The editorial board was expanded to represent more life science disciplines and a greater diversity of educational environments and expertise. CBE-LSE has published work by authors from 12 countries and, in the United States, 43 states and the District of Columbia. Readership has grown such that almost 8000 people are registered to receive electronic Table of Contents alerts for each new issue.

The discipline of biology education research has also grown in recent years. I am proud that CBE-LSE has been an integral part of this trend. The National Research Council (NRC) and the Board on Science Education commissioned a systematic review of the literature, revealing that the number of undergraduate biology education research publications increased substantially in the past decade (Dirks, 2011). Of the 195 articles identified, 50 were published in CBE-LSE (26% of the total), three times more than any other journal. CBE-LSE’s impact is probably an important reason the Howard Hughes Medical Institute awarded another 3-yr grant to the American Society for Cell Biology (ASCB) to partially support publication of the journal.

The accomplishments of the journal are the result of time, energy, and resources invested by many stakeholders, including authors, reviewers, ASCB staff, and editorial board members. I especially thank Brad Kincaid and Gary Reiness, who will end their terms on the editorial board this year. I also thank editorial board members Bruce Alberts (University of California, San Francisco), Diane Ebert-May (Michigan State University), John Jungck (Beloit College), Jay Labov (NRC), and Vivian Siegel (Broad Institute), all of whom have agreed to serve on the board for another term. I am pleased to welcome editorial board members who will start their terms in 2012: Linnea Fletcher (Austin Community College), Graham Hatfull (University of Pittsburgh), José Herrera (Truman State University, currently at the National Science Foundation), James Hewlett (Finger Lakes Community College), Nancy Pelaez (Purdue University), Maria Ruiz-Primo (University of Colorado, Denver), Hannah Sevian (University of Massachusetts, Boston), and José Vázquez (New York University). These individuals will increase the diversity of institutional and methodological perspectives on the editorial board.

Several professional development efforts are also supporting the growth of biology education research. For example, CBE-LSE is a partner in the Biology Scholars Program of the American Society for Microbiology. Through a series of residencies, Biology Scholars prepares biologists to identify problems of importance in biology education research, design appropriate studies, and craft clear and compelling manuscripts. In addition, the broader biology education research community is uniting through the recently established Society for the Advancement of Biology Education Research (SABER). SABER supports networking among biology education researchers through a member listserv, a wiki, and in-person meetings, the next of which is tentatively scheduled for July 12–15, 2012. CBE-LSE continues to employ a constructive review process as a form of professional development. Rather than simply rejecting manuscripts, editors and reviewers at CBE-LSE help authors refine their manuscripts or suggest alternative venues in which authors can publish their work.

Not surprisingly, the biology education research community is experiencing some “growing pains.” We are learning
that our biology jargon has different meaning in the context of biology education research. We are also learning that methods with which we are facile in the lab are not so easily translated to the classroom. For example, in the lab, we are able to control most aspects of our investigations. We use controls as standards for comparison to make inferences about the causal relationships between variables. In most educational settings, such as college classrooms, K–12 schools, and science centers, we are not able to control the myriad factors that may affect learner outcomes, including student backgrounds, motivations, interests, and prior knowledge. When we want to identify trends that indicate causal relationships, we are often limited to other research designs. Those of us who conduct quasi-experimental studies identify a treatment group that experiences an intervention and a comparison group that experiences a different intervention, but not a “no intervention” group (i.e., a control group). To make inferences from the resulting data, we need to think carefully about other variables that may be affecting the outcomes we are studying. I encourage readers to explore Shavelson and Towne (2002) and Slater and colleagues (2011) for more in-depth treatments of this issue.

Another source of “growing pains” in biology education research is the elevated expectation of reviewers regarding the nature of evidence. Although CBE-LSE and other journals have published self-reports of knowledge and skills gains, biology education research as a discipline is moving beyond this methodology. Several studies have shown that people typically overestimate their performance on intellectual tasks (e.g., Mabe and West, 1982; Kruger and Dunning, 1999; Dunning et al., 2003). Now reviewers want more direct measures of student learning. Student reports of knowledge and skills gains can be included as a complement to more direct measures of student learning, such as systematic collection and analysis of student work. Assignment or exam scores are also not sufficient, unless they are accompanied by examples of student work or rubrics that indicate how scores were assigned (Allen and Tanner, 2006).

Finally, I would like to offer some advice for authors on the difference between an essay and an article. Many manuscripts submitted to CBE-LSE describe the implementation and evaluation of an educational innovation. As outlined in the Instructions for Authors, these manuscripts fit the article category. Essays more often resemble a scholarly review or position piece, which can be framed by personal experience or include data from one’s own work (e.g., Hue et al., 2010; Klymkowsky, 2010; Fuhrmann et al., 2011). Essays can also illustrate the translation of research into practice (e.g., Marbach-Ad et al., 2010). I encourage future authors to read these and other recent examples to determine which category is the best fit for their manuscripts.

As I begin my second year as Editor-In-Chief of CBE-LSE, I hope readers will continue to consider this journal their home for biology education research. Furthermore, I hope more biologists will utilize CBE-LSE as a mechanism for getting involved in biology education research: a source for potential collaborators and for professional development in a discipline that affects all of our students.

REFERENCES


Letter to the Editor

Implications for Undergraduate Education of Two Interdisciplinary Biological Sciences: Biochemistry and Biophysics

Jonghoon Kang and Crystal B. Purnell

Department of Biology, Valdosta State University, Valdosta, GA 31698

Dear Editor:

As CBE—Life Sciences Education readers know, modern biology has become more quantitative and mechanistic rather than qualitative and phenomenological. Indeed, the American Society for Cell Biology (2010) notes that “biology learning encompasses diverse fields, including math, chemistry, physics, engineering, and computer science, as well as the interdisciplinary intersections of biology with these fields.” Because of this, interdisciplinary biology courses are increasingly critical for preparing undergraduates for further education and careers in the life sciences. The problem is that many institutions do not reflect this in their undergraduate course offerings. In this letter, we show that the growth of two biological disciplines, namely biochemistry and biophysics, can be described in terms of their representation in the research literature. On this basis, we make predictions about the future trends of research in these disciplines and the implications of these trends for undergraduate education.

The most widely taught interdisciplinary biological science is probably biochemistry. Many undergraduate biology and chemistry departments recognize its importance and offer it as a core course in their curricula. Biophysics is another interdisciplinary biological science. Although it is not as popular as biochemistry, the growing importance of the subject should be recognized by teaching institutions and reflected in their course offerings. Unfortunately, it seems that the number of institutions that have noticed the demand for it is limited.

For example, out of 34 undergraduate institutions in the University of Georgia System (2010), only 5 offer an undergraduate biophysics course or related courses, such as physical biochemistry. In contrast, 8 of the top 10 universities listed in U.S. News and World Report (2010) teach biophysics or physical biochemistry.

Here, we offer a brief analysis of the primary life sciences literature to support our argument for widespread teaching of biophysics at the undergraduate level. Specifically, we used the number of papers indexed by PubMed (National Center for Biotechnology Information, 2010) as a measure of subject importance. We searched PubMed with search terms “biochemistry OR biochemical” for biochemistry and “biophysics OR biophysical” for biophysics with a limit of the search to title/abstract for each year from 1950 to 2009. We then normalized the number of papers for each subject to the total number of papers in each year. We recognize that there are limitations to this approach. For example, the DNA structure paper by Watson and Crick (1953) is not counted as either a biophysics or a biochemistry paper because it does not contain the search terms in its title or abstract. Yet we believe that this metric reasonably represents the activity of the subjects relative to one another.

Figure 1 shows the normalized number of papers of biochemistry and biophysics, from which we draw three conclusions. First, biochemistry has been continuously growing, with a large increase between 1974 and 1975. This growth may explain why biochemistry is now widely taught. Second, the normalized number of biochemistry publications approaches an asymptotic value suggesting saturation (Figure 1A), although we cannot rule out the potential for major transitions in biochemistry research to prompt new growth in the field in the future. Third, although the discipline of biophysics is younger than that of biochemistry, it has been growing exponentially (Figure 1B). This feature can be more clearly seen in Figure 1C, where the number of biophysics papers was normalized to that of the biochemistry papers. Figure 1C indicates that the relative importance of biophysics was ~12% in year 2009, but if the trend of the last 25 years is maintained for the next 10 years, it will be ~20% in year 2020. Current
biology seems to be facing the same “cultural transition” that chemistry faced ~100 years ago when it started including physics-based approaches in education and research. This incorporation of physics in chemistry resulted in physical chemistry, one of the core areas in chemistry today (Servos, 1996).

This analysis reveals the major upswing in biochemistry publications that occurred between 1974 and 1975. We offer a hypothesis to explain this phenomenon: the introduction of a great textbook. We propose that Biochemistry by Lehninger (1970), which has been widely recognized since its initial publication (Sable, 1971), is the original modern biochemistry textbook. Perhaps groundbreaking instructional resources such as this can trigger a quantum leap in research in the discipline. Unlike biochemistry, biophysics suffers from the lack of consensus on the standard topics and skills that should comprise an undergraduate course. This is an issue that biologists and biophysicists need to solve together.

In conclusion, if training future biologists is one of the goals of undergraduate biology education, then our analysis indicates that we need to make biophysics a standard component of the curriculum, similar to biochemistry.

ACKNOWLEDGMENTS

This work was partly supported by Valdosta State University (Reassigned Time fund to J. Kang). The content in this letter is not necessarily the view of Valdosta State University.

REFERENCES


Feature

Approaches to Biology Teaching and Learning

Moving Theory into Practice: A Reflection on Teaching a Large, Introductory Biology Course for Majors

Kimberly D. Tanner

Department of Biology, San Francisco State University, San Francisco, CA 94132

PREFACE

While this feature in *CBE—Life Sciences Education* most often translates some aspect of the research literature—from cognitive science, psychology, science education, and other fields—into practical strategies for biology instructors, this installment is purposefully different in kind. Below, I offer a more personal reflection on teaching, motivated by several considerations. First, many of these features have focused on a single pedagogical approach or issue. However, the disaggregation of the many aspects of teaching and learning can belie the complexity of the task of making things happen in a classroom. Practical teaching strategies inspired by research findings may appear only tangentially related to one another, and many strategies are dependent on intangibles related to individual characteristics of an instructor and an institution. In addition, while evaluating research evidence is key in developing effective approaches to teaching, I have found that many colleagues have been inspired to try something new in their teaching from hearing stories of what could be, of the possibilities, regardless of the research behind the stories. Here I am aiming to use my own recent first-time experience of teaching a large introductory biology course for majors to weave together several practical approaches that are grounded in the research literature.

Since the reflection that follows is a personal story, I will emphasize my teaching experiences without unpacking the direct connections to research literatures. That said, all of the teaching strategies I utilized are strongly influenced by and grounded in numerous papers, books, and research studies. Much of the relevant research literature has been previously presented in installments of this feature (e.g., Tanner and Allen, 2002, 2004, 2007; Tanner et al., 2003; Allen and Tanner, 2005; Chamany et al., 2008). More specifically, I have learned immensely from those science educators who pioneered active learning and similar teaching techniques (e.g., Felder and Brent, 1996; Silberman, 1996; Wiggins and McTighe, 1998; Bransford et al., 1999; Shipman and Duch, 2001; Fink, 2003), in particular those who have done so in the biological sciences (e.g., Ebert-May et al., 1997; Eisen, 1998; Donham et al., 2001; Handelsman et al., 2004; Wood and Knight, 2004; Smith et al., 2009). The research literatures on situated learning, cognitive apprenticeship, and effective tutoring have played a key role in my perspective on the role of the instructor (Flavell, 1979; Posner et al., 1982; Collins et al., 1987; Brown et al., 1989; Chi et al., 1994; Bandura, 1997; Bransford et al., 1999; Shimamura, 2000). And my approach to cultivating instructor-student relationships has been profoundly altered by research on issues of equity, diversity, and identity in the sciences (Tobias, 1990; Rosenthal, 1992; Ladson-Billings, 1995; Steele and Aronson, 1995; Seymour and Hewitt, 1997; Steele, 1999; Brown, 2004; Chamany, 2006; Cohen et al., 2006; Johnson, 2007). These references will not be discussed in this article, but should provide the interested reader with starting points for considering the research that influenced my teaching and this reflection on my experiences.

PROLOGUE

I recently taught the largest biology course, by far, that I have ever taught—an introductory biology course for majors that enrolled ~300 students—at my urban, public, 4-yr university. No doubt many readers have faced similar challenges. As a neurobiologist who has enjoyed more formal training in pedagogical methods than most scientists, I was eager to put to work the many innovative teaching tools I have learned over the years and to apply them in this particular context. As I have considered reflections from my students and reflected on the experience myself, I have been struck by several insights that I would like to share.

I had many worries and concerns in preparing to teach this course for the first time. Would students revolt when
I asked them to talk to the people next to them? Would the teaching strategies that so successfully engaged my nonmajor students be soundly rejected by majors in the same department? Would I, as a colleague jokingly suggested, really be eaten alive by the experience? Make no mistake, years of experience and training did not prevent my confidence from temporarily waning, given reports about student resistance to innovative pedagogical approaches (e.g., Anderson, 2002, 2007; Silverthorn, 2006).

To presage the conclusion of this story, I did indeed teach the course. I not only survived, I think I thrived. And, no, I did not come anywhere close to being eaten alive! As a means of processing the experience metacognitively, I wrote a final reflection, something I require my students to do in all my courses. This final reflection is usually 1200–1500 words in length and is in response to the following prompt: “What have you learned in this class that will continue to influence you for years to come? How have you learned these things?” This final reflection inevitably gives me insights into what my students have valued in the course, often things that I did not realize were critical for them.

So, what did I learn in teaching an ~300-student introductory biology course for majors for the first time? I had five realizations that I believe will continue to influence me for years to come.

Realization #1: Even in a classroom of ~300, students can feel a personal connection with their instructor, and I can learn their names (mostly).

Realization #2: A ~300-student lecture classroom can be a welcoming and engaging place to be at 8:00 am.

Realization #3: It is important to be on the same team as my students.

Realization #4: Students value learning how to think like a biologist.

Realization #5: I can minimize student resistance by being explicit about the reasons behind my teaching choices.

While I learned a great deal more from this teaching experience than I can record here, I am sharing those insights I think are most widely applicable to all biology instruction in a variety of settings. In addition, since much of what I personally learned came from reading my ~300 students’ own final reflections, I am occasionally including their voices here as a complement to my own. The quotes from students presented below were not rare examples, nor are they intended as formal, systematically analyzed research evidence of any sort. The student voices offered are shared because they were a primary source of insight in my own metacognitive reflections on my experiences teaching this course.

MY FINAL COURSE REFLECTION

Realization #1: Even in a Classroom of ~300, Students Can Feel a Personal Connection with Their Instructor and I Can Learn Their Names (Mostly)

I have always considered knowing students personally to be a key part of my teaching. With ~300 students, I was told that was simply going to be impossible. However, just a few months earlier, I had asked a colleague what college science course was her favorite. Without hesitation she had replied, “Oh, that’s easy, General Chemistry with Dr. Laird.” Why? “Oh, because he really wanted each one of us to learn and tried to learn all 600 of our names in the process. He did not learn them all, but what mattered to me was that he tried, that he cared enough to try.”

Learning a student’s name, or even just attempting to learn it, appears to go a long way in building a functional teaching–learning relationship between an instructor and students. While simple in concept, it appears to make students feel a personal connection, even in the presence of, in my case, ~300 students. From my students’ final reflections came comments like those below:

“When I signed up for the class I was nervous signing up for such a large lecture, but now that the class is coming to an end I must confess that the class was never any less personal than a class of 30 students and not once did I feel lost in the mass of students.”

“This has been one of the most personal classes I have ever taken at the college level. Even though there were more than 270 students in the class, the professors made each individual feel like they were somebody and strived to make sure every student was doing well in the class.”

As another example, I received many emails over the course of the semester that suggested students felt I was aware of their individual behavior during class...

“Hi Kimberly, I wasn’t sure if you noticed I kind of dozed off for a little while in class today. I want to apologize and let you know that it’s not you. It’s me because I didn’t sleep due to homework. I feel bad because you try to be positive and enthusiastic. I will try to stay awake next time. Sorry, Elena”

While it may seem daunting to learn the names of students in such a large class, there were a few simple strategies, recorded below, that aided me greatly.

Use Name Cards in Lecture. In most of my classes of fewer than 50 students, I have used a folded 8.5 × 11 piece of cardstock marked with the first name of each of my students to make an individual name card for each student. These name cards have been indispensable tools in my courses for learning names, constructing groups, and encouraging students to get to know one another. On the first day of this large course, every student received a bright green name card and was asked to write his or her name in large letters on the front and the back. I expressed the desire not only to learn their names, and explained the importance of learning together as a community, even a big one. It was not just important for me to know their names; it was important for their colleagues (other students) as well. See Figure 1.

Further, these name cards are also assessment and memory devices. Prior to folding them, I print a grid on the back that has an entry space for each class meeting during the semester. I often pre-enter into this grid important dates, like exams and holidays. The rest of the spaces are left blank. In classes of smaller size, where I have the luxury of more time, I ask a question at the end of each class that students can respond to in the space for that day. On the first day of this large course, I explained the dual purpose of these name cards. I expressed that as a neurobiologist I recognized the name cards were an excellent tool for them as students to record...
something—be it an aha moment, a frustrating confusion, or any other anecdote—that would help them to remember the class session experience for that day and distinguish it from the other 44 class sessions we would be spending together. While recording these entries was not required, it was strongly encouraged. At many points during the course, students asked whether I wanted to check their name cards. They would share particularly funny or meaningful entries with me. And at the final exam, a surprising number brought in their card to show me how useful it had been for them in tracking their experiences through the semester.

**Build a Personal Information Index Card Directory on the First Day of Class.** While the name cards were immediately useful in class, even more helpful was the collection of a “Personal Information Index Card” from each student on the first day of class. Students were asked to provide the following:

- Your given name, and the name you want to be called
- Your major
- Your email
- The city, state, and country where you were born
- Something unique about you that is not true of anyone else in this room

I referred to this stack of ~300 index cards regularly while I was teaching, and I still have the cards and refer to them on occasion. It was immensely useful in learning students’ names by associating names with majors and unique characteristics. See Figure 2.

**Take Their Pictures with Their Name Cards.** The most critical thing that I did with the name cards was to take each student’s picture—with his or her explicit permission—with the name card in front of the individual in the picture. I did this in the laboratory sections during the first 2 wk of the course. As I took the pictures, I asked what they had written on their index card the first day of class that was unique about them. This was immensely helpful! Oh, you were the one who: was attacked by a monkey, writes jingles, is a professional photographer, died for 4 min, eats only macaroni and cheese, etc. The mechanics of obtaining the photographs was probably even more important than taking them. It meant that within the first 2 wk of class I had had a personal conversation, albeit ~1–2 min long, with each student in my course. See Figure 3.

**Assign a “More about You” Survey as Homework on the First Day of Class.** While the name cards and personal information index cards were immediately useful, they were quite limited in the amount and type of information collected about a student. So, the first homework assignment given on the first day of class was an online “More About You” survey that consisted of 30 questions of four types. Many of these questions were closed-end, multiple-choice questions, so that the online survey software provided a synthesized output of the profile of students in the course. These questions were mostly factual: major, class standing, contact information, transfer status, gender identification, languages spoken, and cultural/ethnic
identification. Strikingly, I found out from this survey that about one-third of our students in this “majors” course were not biology majors, but rather kinesiology, biochemistry, and other majors. Other questions probed students’ current career goals. There were also conceptual questions that probed for common misconceptions in biology and gave me insight into students’ baseline ideas on a few core ideas that threaded through the course. Finally, there were questions that asked students what had been most effective in supporting their learning in previous biology courses, what had been least effective, and what else should be known about them as a learner. These last questions became most useful later in the course. For common misconceptions in biology and giving me insight into students’ baseline ideas on a few core ideas that threaded through the course. Finally, there were questions that asked students what had been most effective in supporting their learning in previous biology courses, what had been least effective, and what else should be known about them as a learner. These last questions became most useful later in the course. For common misconceptions in biology and giving me insight into students’ baseline ideas on a few core ideas that threaded through the course. Finally, there were questions that asked students what had been most effective in supporting their learning in previous biology courses, what had been least effective, and what else should be known about them as a learner. These last questions became most useful later in the course.

**Make Coming to Office Hours a Part of the Course.** As I have done for smaller classes, I included in the syllabus the following: extra credit is available for meeting for at least 10 min during an office hour session to discuss your progress and feedback for class (15 points). By the end of the semester, more than one-half of the students in the course had taken advantage of these office hours appointments for extra credit and many were repeat visitors, even though credit was only awarded for the first visit. While I do not have a rigorous comparison, I predict that I would not have had these hundreds of one-on-one conversations with students about their career goals, their school struggles, and, importantly, about biology, if this office-hour incentive had not been structured into the course.

**Say “Hello” on Campus and Do Not Be Afraid to Ask Their Names.** There were many times when I simply could not remember students’ names. The final strategy I offer is honesty. Many times I had to share with a student that I could not remember his or her name. Never did I feel that students were insulted, nor did I feel that they were disappointed in me. Remembering someone’s name is a placeholder for knowing them and wanting to interact with them. I got a lot of credit just for trying, much as one may when attempting to speak the language in another country. I see those students around campus constantly. I call many but not all of them by name. And they call me by my name, too.

**Realization #2: A 300-Student Lecture Classroom Can Be a Welcoming and Engaging Place to Be at 8:00 am**

While the class size could have been bigger, the class time could not have been earlier or shorter—it was scheduled at 8:00 am on Mondays, Wednesdays, and Fridays for 50 min. Despite this, a variety of simple teaching techniques appeared to cultivate a positive atmosphere and build a sense of community, making the classroom a welcoming, comfortable place to be, as reported by the students:

> “I wasn’t sure what to expect when I walked into Biology 230 for the first time, so I think it’s fitting to start my reflection with that experience and how it affected me throughout the semester. I was extremely surprised at the size of the massive lecture hall with music playing when I first walked in. I quickly grabbed a spot in the middle of the movie style seating and noticed a strange contraption on the stage at the front of the class. A large rectangular box on stilts with a question mark on the front and a funnel protruding from the bottom was placed in the middle of the stage on top of a table. I had no idea what the ‘mystery box’ could have to do with a biology class, but was intrigued by how it worked and what was going on inside…”

> “I’ve always been a very solitary person, especially in my academic career. This class helped me realize that I don’t always have to do things on my own and in many cases it is detrimental to do so.”

> “In lecture I also really enjoyed taking time to discuss things with my other classmates. Anytime there was a clicker question and we got to discuss it, I felt that had two purposes. The first was to get us to talk to people we may not have known very well or at all, and we got to throw ideas off each other to come to a conclusion about the question given to us. Another reason it was nice to discuss questions among my peers was because it was an 8 am class, and talking to other people helped wake me up in the morning…”

**Make the First Day of Class about Science and Not the Syllabus.** On the first day of class, I think the most important thing that I did was NOT read the syllabus. The first day of any course, like the beginning of any new experience or relationship, is full of expectation and excitement. In our first 50 min together, we did science together and got to know one another. As described above, collection of the personal information index cards, making their student name cards, and meeting some of the students sitting around them were critical activities. But that was not science. So, on the first day, we did a problem-solving challenge called the Mystery Box, in which water is poured into the top of a box and may or may not come out of the bottom of the box, changed or unchanged, depending on the setup of the box. The students were immediately engaged in making predictions, recording observations, proposing and comparing models, and revising these models based on new observations. While the Mystery Box is a wonderful tool, there are dozens of ways to engage students in the habits of mind of scientists on the first day of class.

**Come to Class Early and Make Yourself Available.** This is not always possible in the busy lives of instructors, but
simply being in the classroom for some period of time prior to the start of class appeared to be important in making me accessible and uniniminating to students. While the 8:00 am start time of this course was not optimal in many ways, it was a luxury in that our classroom was open and available. I used this time to mingle with the students, not to busy myself with my computer, shuffle papers, or otherwise appear behind the podium. In the beginning of the semester, I needed to be the initiator of conversations, to actively talk with students, approach them, ask them how it was going, and be friendly. Later in the semester, I generally did not need to approach students, because they were approaching me. I did not talk with every student, but used this opportunity to meet and talk with students I did not usually see in office hours or who tended to sit in the back row. I consciously made sure I did not always talk to the same students. We talked about the course, but I also asked them how the semester was going, what they did over the weekend, and things I would talk about with any other professional colleague.

**Look Thrilled to Be There.** Like most faculty, I was at many points overwhelmed during the semester. In addition, this class was at 8:00 am, Mondays, Wednesdays, and Fridays. Ugh. But when I was in class, I viewed it as my job to look thrilled to be there. And this did seem to have an important effect on my students. If I was not excited to be there, how could I possibly expect them to be? There are a variety of ways to express enthusiasm for being there, and even more ways not to. It takes conscious effort to remember that you as the instructor set the tone for the classroom.

**Share about Yourself.** At many points in the semester, I strategically and purposefully shared my own experiences with students. On the first day of class, I shared that I was a first-generation college-going student, and that there were no other scientists in my family. For many of the students at my university, this is an important point of connection. After their first exam grades were returned, I reminded them that I was a PhD-level neurobiologist and then promptly shared that I had, in fact, failed my first neurobiology exam as an undergraduate. At multiple points during the semester, I shared that I was a working parent, like many of them. On numerous occasions, I emphasized the vastness of the field of biology and how it was really no longer possible for individual biologists to be expert in all the subfields of the discipline. Numerous students have recounted that these stories somehow “took the edge off” for them. I became one of them instead of an “other” without similar experiences or an empathetic stance, and my success looked like something they could achieve. A danger in this approach is one may appear either unknowable or unprofessional. Another danger is cultural incompetence, and making assumptions that the stories you share are relevant, when in fact they are not, or they inadvertently reinforce stereotypes of science and scientists. So, sharing of yourself can be a teaching tool, but it must be judiciously used, usually for specific purposes at specific junctures during the course.

**Commit to Making Time for Students to Talk during Class.** There was no class session during my teaching in this course in which students did not have an opportunity to exercise their voice at least once, and usually multiple times, during class. The simple “think-pair-share strategy” is as possible with 1000 as it is with 10. Like any other novel strategy that students may be unaccustomed to, it takes time to entrain such classroom behaviors and interactions. I most often used teaching strategies to get students talking at the beginning of a class period, usually to discuss homework that they had been doing in preparation for class. In addition, a quick 2-min discussion around clicker questions became an expectation, such that not talking about a clicker question would be a surprise and a disappointment to some students. Not all the students in my ~300-person community were thrilled to talk with their neighbors, but in examining students’ final reflections, I found these students were a small minority; the majority of students valued a chance to hear the perspectives of their peers and share their own views. Why the lack of resistance? Four things may have made a difference. First, I openly acknowledged that having students talk in class with others would put some students out of their comfort zone, while simultaneously putting other students in their comfort zone. I expressed that my job as an instructor was to have no individual chronically out of his or her comfort zone. Second, I kept the time for talking short, but periodic and recurring. Third, I actively worked with those students I saw not talking with others to get them engaged in a group. After several class sessions, I had to do this much less often, because they knew I was coming their way if I did not see them engaged in talking with peers. Fourth, the things I chose to have them talk about in class appear to have been interesting to them and complex enough that multiple perspectives were possible and intriguing. The choice of what to have students discuss appears to be much more challenging than the mechanics of getting them to talk.

**Explicitly Encourage Student Collaboration and Study Groups.** On multiple occasions during those 10–15 min before class, I introduced students to others who had something in common, shared interests I had discovered through their online Biologist Journal entries. I encouraged students to find study buddies and form study groups, especially in the context of the laboratory sections. As this cohort of students has gone on to the second semester of this introductory biology series, I have heard from numerous students that it was a relief to walk into that next course knowing so many classmates and having some established study partnerships and groups that would continue to serve them in their biology studies.

**Play Music during the Time before Class Starts.** I had the luxury of an 8:00 am class with no prior class to prevent my occupation of the classroom. Since I needed to wake up and so did my students, I made it a policy to play music for the 20 min prior to the start of class. In fact, I invited and began to get requests on a regular basis, which was wonderful, as we had music in a variety of languages and styles over the course of the semester. This strategy is certainly not possible for everyone, but if feeling welcome, comfortable, and at ease promotes learning, strategies such as these are not inconsequential.

**Realization #3: It Is Important to Be on the Same Team as My Students**

One of the most striking things I have experienced as an undergraduate biology educator is the assumption that instructors and students play opposing and sometimes adversarial
roles. While I had experience in cultivating a relationship that put me on the same side as my students in classrooms of fewer than 50 students, I had sincere concerns about whether I could accomplish this with ~300 people sitting down in a room and me generally standing up. Even the body language of the situation put us in roles that I wanted to change. Reflections from students suggested that, even in a large class, we could cultivate a teaching and learning partnership where we were on the same side.

“The professors in Biology 230 were amazing because they seemed to really care that the students understood the material.”

“The fun I had learning this year in this biology course was really unexpected… Before enrolling in the course I heard horror stories about how the class was really difficult and that there was a lot of material to cover for each exam and how it was very difficult to pass the class with a high grade. This is all true but I felt like this year things changed, the teaching was so effective that it literally made things a lot simpler and easier to learn.”

“[This class] made me realize that I haven’t been learning but competing in school, which is such a tragedy. All I have been thinking about is grades, and not really what I could get out of it in the end…This class has taught me pursue those things, and try not to make everything about grades.”

Strategies that appeared to cultivate a partnership between students and instructors were neither specific activities nor anything particularly dramatic, but rather habits of language and interactions that were purposefully collegial.

**Treat Students as You Would Your Professional Colleagues.** In my own scientific training experiences as an undergraduate, graduate, and postdoctoral student, I have personally valued being treated as a scientific colleague even by those who had far more experience and wisdom. In response, I have generally taken the stance that students are my colleagues, and refer to and treat them as such. The use of the term “colleague” seemed to aid a great deal in forging a partnership between my students and myself. Even simple acts that indicate respect, such as replying to an email in a timely fashion, and interactions that were purposefully collegial.

**Use Language That Puts You All on the Same Team.** The use of “you” is problematic. It has an accusatory tone and is generally unhelpful. Phrasing my professional interactions with students using “I” or “we” instead of “you” language served me very well in a ~300-person classroom. Early on, the phrase “Team Bio 230” emerged as a way to refer to all of us collectively. In addition, when students’ performance on an assignment was unexpected, I did not use language of “you all” or “you” or “they,” rather it was about “we.” While this may seem trivial, our language often belies the true nature of our beliefs. If we really believe that we are partners with our students in learning, our language should support that.

**Solicit Students’ Feedback Regularly on How the Course Is or Is Not Supporting Their Learning.** Through weekly online journal assignments, students were involved in writing regularly about their experiences in the course. While most of these Biologist Journal entries were focused on grappling with conceptual confusions or unpacking case studies, some were used to solicit student feedback on their learning in the course. They were not asked what they “liked” or “enjoyed,” but rather what “most supported” or “least supported” their learning of biology. The regular invitation for student voice and opinion on the effectiveness of the teaching and learning process is an important way to construct a situation where the instructor and the students are on the same team. Interestingly, the aspect of the course that students consistently said least supported their learning was reading the textbook, which they found disjointed, jargon-laden, and confusing.

**Strive to Think Like a Student.** I believe it is easy to make assumptions about why students are not successful in our courses, rather than attempting to understand their experience. One of the key habits of mind that guided teaching and learning in this course was to “think like a biologist.” At one point, a student approached me and said, “I really like the fact that you seem to think like a student.” In structuring our work together, I did take into account when students would be experiencing intensive work due to the laboratory portion of the course and tried to then minimize the demands occurring from the lecture portion of the course, which I controlled. I did assign work over weekends and breaks, but I tried to assign it early enough so that individual students who had family obligations or just wanted a real break could complete the work in advance. In addition, I assigned page ranges in the textbook for reading, as opposed to entire chapters. I often find that there is far more information in a textbook than either I want students to learn or that they are likely to be able to learn. Being thoughtful and judicious about indicating the pages that were really, really important was an appreciated aid to students. Finally, every day in class, I posted an agenda for what we were going to do over the course of the next 50 min and the guiding question that we were trying to understand (e.g., “How can mutations in DNA contribute to the development of cancer?”). This was perhaps in response to my own experiences as an undergraduate student, where I was frequently at a loss to express what had transpired in my class sessions or what the point had been. And, importantly, I was explicit with students about my attempts to “think like a student” in structuring our work together.

**Make the Conversation about Learning and Not about Grades.** Early in the course I spoke very directly about the purpose of the course being about learning biology. I expressed to students that with ~300 students, there would be a lot of accounting, and I always wanted them to alert me to any recording errors in keeping track of the grades they had earned; however, I expected our one-on-one conversations to be focused on issues of learning and issues in biology. Grades were something they earned, not something I assigned. Grades were determined not solely from exams, but also from weekly Biologist Journal entries, homework assignments, in-class clicker participation, in-class index card participation, the laboratory portion of the course, and even their final course reflection. These strategies reduced discussion of grades to only a few students, and with them only rarely.

**Care about and Believe in Your Students.** This strategy seems like stating the obvious. That said, I have encountered many faculty who express something to the effect of “some students will do fine no matter what, some are hopeless, and there are
On every day of class, from the first day of class, there were six posters with colorful lettering that were hung on music stands on the stage of the classroom. These six posters read as follows: Use Evidence, Ask Questions, Be Skeptical, Cultivate Wonder, Identify Confusions, and Think Like a Biologist. These scientific habits of mind, which constitute part of the learning goals for all of my biology courses, appeared to play an important role in helping students aim for learning, rather than just memorizing information during the course. Of course, there were things to memorize, and many students seemed to do this well. Most, though, appeared at a loss for what else to do. The U-ABC-IT posters (as they came to be known) were used in lectures and made constant appearances in students’ writing in their Biolologist Journals, aiding all of us in keeping a focus on the habits of mind that we were attempting to cultivate in this introductory biology course. See Figure 4.

Realization #4: Students Value Learning How to Think Like a Biologist

I have asked colleagues who teach upper-division biology courses what they really want students to leave introductory biology being able to do. The answers are quite diverse, but no colleagues have ever provided me with lists of detailed mechanisms or even conceptual ideas in biology. Most, at the core of their answer, express a desire for students to be able to think critically and skeptically and know how to learn. This goal of teaching students how to learn and think like biologists was central to the course and valued by students.

“Biology 230 has also taught me to think. Quite a bold statement, I agree. But I simply realized after finishing this course (completely finishing after I submit this essay), that it’s not enough to read, memorize, or even comprehend. I think that in order to really grasp a lesson one must ask questions, identify confusions, use evidence, and be skeptical. In order for one to cultivate wonder, one must look at something from every angle and want to learn it and not just think about it in a normal sense of things.”

“No, another important concept or concepts, I will always remember is U-ABC-IT, which is something that will be always useful in whatever class I take whether its lower or upper division classes. Use Evidence, Ask Questions, Be Skeptical, Cultivate Wonder, Identify Confusions, and Think Like a Biologist are the main concepts one should use always in life and maintain their mind open to new ideas and view things skeptically. I believe that if we use all these concepts, we can become better scientists and analytic persons.”

Be Explicit about the Scientific Habits of Mind That You Want Students to Cultivate. On every day of class, from the first day of class, there were six posters with colorful lettering that were hung on music stands on the stage of the classroom. These six posters read as follows: Use Evidence, Ask Questions, Be Skeptical, Cultivate Wonder, Identify Confusions, and Think Like a Biologist. These scientific habits of mind, which constitute part of the learning goals for all of my biology courses, appeared to play an important role in helping students aim for learning, rather than just memorizing information during the course. Of course, there were things to memorize, and many students seemed to do this well. Most, though, appeared at a loss for what else to do. The U-ABC-IT posters (as they came to be known) were used in lectures and made constant appearances in students’ writing in their Biolologist Journals, aiding all of us in keeping a focus on the habits of mind that we were attempting to cultivate in this introductory biology course. See Figure 4.

Realization #5: I Can Minimize Student Resistance by Being Explicit about the Reasons behind My Teaching Choices

While much has been written about the potential perils of integrating active learning into undergraduate science courses, I have trouble identifying much of any resistance in this case. I would speculate that the primary reason was that I was quite explicit with students about what we were doing in class and why. In fact, the constant stream of small assignments that engaged students in active learning and self-assessment throughout the semester appeared to be welcomed. In their reflective journals, students reported feeling that a number of the teaching choices I had made, many involving teaching strategies initially unfamiliar to them, did effectively support their learning:

“What I really liked about the lecture class was the homework that was given to us. I felt that the homework really contributed to my studying because we had to utilize all of the concepts we learned in order to do them.”
“At the start of the semester I honestly thought that the journal entries would be such a time consuming, tedious job, but it contributed to my learning a great deal!! Without the journal entries, I don’t think that I would have understood the material as well as I have. I love learning more fascinating facts about science, and the journal entries along with the articles we were required to read, provided me with materials to explore science further.”

“When I found a way to connect what I was learning in the classroom to something in my life, it made it easier to remember. When we were learning about cellular respiration and photosynthesis I remember putting my hand up next to my mouth and breathing out. I knew we breathed out carbon dioxide, but actually feeling the water come out too helped me to remember which equation was which. This I will always remember.”

“Concept mapping really helped me understand the digestive system... I believe that this technique of concept mapping I will always use in future classes when I don’t understand certain things or have a problem relating terms, ideas, or concepts.”

From my experiences, the following strategies are concrete ways to engage students in their own learning and be explicit with them about how this will benefit them not just in the current course, but also in their careers and lives.

Give Them More to Do Than Just Read the Textbook. For many students, the lecture (not the laboratory) portion of science courses consists of lectures, assigned textbook readings, and exams. However, the tenets of active learning suggest that we engage students in a variety of small assignments to help them grapple with the material, direct them toward changing their minds about common misconceptions, and provide explicit opportunities to connect biology learning to the real world, their personal lives, and their own opinions. The most highly praised (by students) pedagogical tool that I used in this course was reflective writing, followed closely by case studies and concept mapping. For some students, reflective writing in online Biologist Journals and case studies were beloved, while for others concept mapping was a transformative discovery. While the mode of active learning preferred by students varied, active learning in general was welcomed.

Allude to How Students Can Apply Learning Strategies to Other Courses. Throughout the course, there was an emphasis on the applicability to other courses of the learning tools being used in our course. Early and often, I stated that the purpose of this introductory biology course was to prepare students to learn in upper-division biology courses, where the content would be more detailed and the pedagogical approach likely more traditionally lecture-based. Students were encouraged to take the reins for their own learning in this course, and in others. Students seemed shocked by the statement that college would be over before they knew it, and they would still need to know how to go out and learn on their own.

Connect Learning Experiences in Your Course to Other Courses They Will Take. A common complaint among biology faculty is that students enter upper-division courses without a solid understanding of the basics, for example, how genes encode traits (central dogma). So, when we worked on this conceptual area, I introduced pictures of upper-division faculty and the courses that they teach in which central dogma would play a key role. And I challenged my students to learn the concepts not just for the present course, not just for our exams, but also for upper-division courses, in particular, and for their professional lives, more generally. Motivating learning in this way appeared to raise the stakes and minimize the importance of learning for the short term, for exams, and for grades.

Make Learning in Your Course about Learning for the Real World. Biology is the study of living things. Unfortunately, even with a great deal of change in textbooks and curricular materials over the years, biology courses can still feel abstract to students and unrelated to things they care about. I was explicit about how we would be doing a lot of real world–like work to prepare them for grappling with biology problems postcollege. Case studies were a wonderful strategy that made the relevance of the biology we were learning explicit. Students were challenged to defend or refute a connection between the candy Altoids and mad cow disease early in the course to guide our evaluation of “living” and our exploration of the different classes of biological macromolecules. The case of marathon runner Cynthia Lucero (who died of hyponatremia) and the exploration of the cholera outbreak in Haiti were both striking entry points for students in understanding the importance of osmosis. Finally, the integration of real people from a variety of different backgrounds and their involvement in the discoveries that we had studied were noted by students on the final exam: Henrietta Lacks, Rosalind Franklin, Stanley Prusiner, Maurice Wilkins, Lynn Margulis, Erwin Chargaff, Francis Crick, George Washington Carver, Phineas Gage, Elizabeth Blackburn, Louise Brown, and Henry Molaison.

Highlight the Diversity of Learners and the Role of Being out of One’s Comfort Zone. Finally, I have thought a lot about why my students were more accepting of active learning strategies than predicted. I believe communication was critical. I was explicit with my students that aspects of my teaching might be highly effective for some of them while simultaneously being benign or perhaps even frustrating to others. I was explicit with them that they would at times be in their comfort zone and at other times out of their comfort zone. My commitment to them was to provide points of entry for all of them and to make sure that no learner was always out of his or her comfort zone. I compared teaching to having a tool belt: the hammer (lecture) is a great tool, but not for all tasks. Sometimes you would be better served by using a wrench (case study) or a screwdriver (concept map). This explicit discussion of the diversity of learners in the room and the requisite need for a diversity of teaching styles may have been the preventative strategy to blunt the possibility of revolt. Even students who most enjoyed lecture and would have thrived in a lecture–text–exam cycle classroom would reflect in their Biologist Journals that they saw the merit of the different teaching strategies for some of their colleagues.

In Summary
So, what did I learn in teaching a ~300-student class for the first time? In general, I learned that many assertions I had previously heard about the values and behaviors of
students in a large introductory biology class for majors did not resonate with my own experiences. The same pedagogical approaches may work differently in the hands of different instructors, and sometimes intangibles may be at play, aspects of teaching that are stylistic, affective, and centered on the student–instructor relationship. These “intangibles” should make everyone skeptical of teaching discussions that make claims about “what works” in any general sense. Teaching is a social endeavor about personal relationships. What I learned in teaching the largest class that I had ever attempted was that these personal relationships did not just disappear. They are there if you notice them, cultivate them, and honor them; in fact, they are a major asset in accomplishing our teaching goals. Finally, metacognition is not only important for students learning about biology, but also for instructors’ learning about biology teaching. While it is neither possible nor desirable for every course we teach to become a full-scale research project, we can cultivate a scholarly, scientific, and metacognitive approach to teaching by purposefully reflecting on our experiences and student evidence collected along the way.

POSTSCRIPT

The reflection above is singularly my own and in no way represents the perspective and opinion of others affiliated with the teaching of this course. I am immensely grateful for the collaborative efforts of my co-instructor Dr. Zheng-Hui He. While I led instruction during the first and last quarters of the course, Zheng-Hui led during the second and third quarters with his own approach and style. We met weekly to collaborate and strategize, attended one another’s lecture sessions, and attempted to make connections across our sessions. The graduate teaching assistants and lecturers who were the pedagogical leaders in the laboratory sections graciously used name cards in the laboratories to aid me in learning names, but they were not asked to adjust their teaching styles to match mine in any way. Two part-time graduate assistants—Stephen Ingalls and Issam Jadrane—shared in the participation-based grading of online Biologist Journals, index cards, and concept maps, as well as the rubric-based grading of open-ended exam questions. They attended and audiotaped every lecture session, and graciously assisted when classroom materials setup was more complex. Everyone affiliated with the course functioned as wonderful thought partners through hallway conversations, email chats, and other informal interactions.

The inevitable question will arise about how much time was involved in my approach to this course. I did not track this, but note that the approaches I took that seemed to be most effective did not require extra time, but a different use of time before class, during class, in preparing class sessions, and in structuring office hours. For example, I did not read every Biologist Journal from every student every week. Generally, reading about 10%—about 30 journal entries—of ~300–400 words each would give me a good sense of the range of ideas present among students. I worked on learning the names of students while waiting in line for my lunch or in any other place where I was waiting and could use my phone to access their pictures. I tried to make my appointments with students ~15–20 min each, so that I could see 8–10 students a week during my regularly scheduled office hours. So, while assistance in grading written work may be critical, many of the strategies presented above can serve as small starting steps toward changing the nature of an introductory biology course with as little or as much investment of time as desired by an individual instructor.

ACKNOWLEDGMENTS

Sarah (Sally) Elgin made significant contributions to the tone and flow of this feature as my CBE-LSE editor, and Erin Dolan contributed references to student resistance to innovative pedagogical approaches for inclusion.

REFERENCES


Wood W, Knight J (2004). Teaching large biology classes: active-engagement alternatives to lecturing and evidence that they work. MBoC 15, 5338a.
Ask your students where they think new medicines come from. For that matter, how about old medicines? Aspirin, perhaps the most famous medicine, has a very interesting history that entails folk medicine, chemistry, clinical research, and molecular biology. Bayer has a website called Wonderdrug (www.wonderdrug.com/pain/asp_history.htm), where you can see a timeline of aspirin’s history, which reaches back to before ancient Greece. On the subject of painkillers, I was musing over the yin and yang of basic and applied research the other day while in the endodontist’s chair for a root canal. My “endo” was telling me about all the different “medicines” he was using. To keep myself occupied, I tried ticking off what was primarily the fruit of basic versus applied research. The procaine injections were a result of mostly applied research, but using basic chemistry to improve cocaine. Sodium hypochlorite as well as the drill and tiny canal files were definitely the products of applied research. The microscope has its origins in basic research but has matured into applied technology. EDTA and osteoclast-osteoblast modulating factors are the fruits of basic research. But none of my determinations were really tidy or definitive.

The research enterprise is confusing to most people, even for advanced students. How do results get transferred to medical advances? Converting basic research into new treatments has many names such as translational, clinical, applied, and disease-oriented research, to name just a few. Clearly, we need all these kinds of research, and they depend on one another. In this Web-review feature, my primary interest is to consider three research stories that highlight the importance of basic research to improving medicine because basic research is fundamental to health advances and yet often underappreciated. After all, it’s not that long ago that the National Institutes of Health (NIH) and National Science Foundation (NSF) were regular recipients of Senator Proxmire’s Golden Fleece awards for wasting taxpayer money. While emphasizing the critical importance of basic research, these stories reveal how important and interdependent the various categories of research are.

Analgesics like aspirin are an important group of painkillers primarily for treating relatively minor aches and pains. There is also a great need for treating pain-associated chronic debilitating illness as well as the intense pain that can follow certain traumatic injury or some surgical procedures. The best-known drugs for chronic pain management are the opioids and other narcotics, which also have the well-known downside of being addictive. Have a look at the University of Utah Learn Genetics website for an excellent feature on addiction (http://learn.genetics.utah.edu/content/addiction). There is a long-standing need for drugs that manage chronic pain more effectively, with fewer side effects, and without being addictive. The opening scene of my first story features a consummate predator stalking her prey (Figure 1). But this predator is a snail, of the genus *Conus*. The cone snails are a large and diverse group of marine mollusks. They prey on worms, other snails, and even fish, but like all snails, their locomotion is slow. Unlike more familiar predators, they don’t have strong jaws and sharp teeth. The cone snails have evolved several fascinating hunting strategies to capture their prey.
Baldomero “Toto” Olivera has done pioneering research on cone snail venoms, motivated by the triple purpose of learning new things about how the nervous system works, developing new research tools, and discovering new medicines. Toto talks about his research in his 2009 HHMI Holiday Lectures (www.hhmi.org/biointeractive/biodiversity/lectures.html) and has developed a website aimed at providing teachers and students with information about cone snails (www.theconesnail.com). Each cone snail species mixes a unique cocktail of up to 200 peptide toxins. Given that there are hundreds of cone snail species, there are estimated to be tens of thousands of peptides targeting a variety of cellular targets such as ion channels, transmitter vesicles, and receptor complexes. In addition to Toto’s Holiday Lectures, aimed at a high school audience, you can view a more technical version of the story delivered by Toto on Ron Vale’s iBioSeminars website (www.ibioseminars.org/lectures/chemicalbiologybiophysics/baldomero-olivera.html) (Figure 2).

Although Conus toxin research has uncovered previously unknown types of ion channels in the name of basic research, I’m telling this story because new medicines have come from this research. Toto began purifying peptide toxins from cone snail venom in the early 1970s. Research psychiatrist Michael McIntosh, then an undergraduate working in Toto’s lab at the University of Utah in the early 1980s, had found that one peptide—ω-conotoxin—caused an interesting “shaker” phenotype when injected into mice (view a short interview with Dr. McIntosh at www.hhmi.org/biointeractive/biodiversity/McIntosh_bio.html). ω-Conotoxin turned out to be a calcium channel blocker and in the mid-2000s was approved for the management of chronic, intractable pain. Currently a number of other peptide toxins derived from cone snails are in development to treat Parkinson’s disease, epilepsy, heart disease, and pain. The animation
New Medicines

found at www.hhmi.org/biointeractive/biodiversity/2009_prialt_blocks_motor.html shows the physiological action of \( \omega \)-conotoxin. To learn more about the synthesized medical form, trademarked as Prialt, search the European Medicines Agency website (www.ema.europa.eu) or use a direct link to the Prialt pages (http://tinyurl.com/4pouwgh). Prialt’s current patent holder, the French company Ilan, has produced a website aimed at doctors and patients (www.prialt.com/patients/product_info/about_prialt).

The nearly three-decade time frame from basic discovery to a clinical drug approved for use with patients might surprise students. There is of course lots of information about the drug approval process on the U.S. Food and Drug Administration (FDA) website, but I found one of the best concise graphics, including time frames, on the website of the pharmaceutical industry publication NGP (www.ngpharma.com/news/Glenmarks-new-drug-fails-in-trials). The time frame ranges from 12 to 25 years to go from the discovery process to FDA approval, and they present time estimates for each phase of the process. You might challenge students to come up with ways to streamline this process while maintaining patient safety.

For conotoxin researchers, it’s been a long journey from the reefs of the Indo-Pacific to new pharmaceutical treatments. Students are likely aware that there is a huge public health problem concerning antibiotic resistance, a perennial search for new antibiotics, and hopes for entirely new classes of antimicrobial agents. The next story has not reached the point of a new drug for patients yet but holds great promise. By sheer coincidence this story also features a predatory mollusk, the diminutive bobtail squid. Bobtail squid bury themselves during the day and come out at night to prey on crustaceans and fish. You can see videos of how they bury themselves on the vimeo website (vimeo.com/15490567) and on the Biointerative Biodiversity pages (www.hhmi.org/biointeractive/biodiversity/2009_bobtail_squid.html). These night hunters have a challenge when the moon is out. Moonlight causes them to cast a perceptible shadow as they cruise along the sandy shallows in search of prey. Their adaptation to this situation is to deploy a form of countershading by emitting luminance that they can tune to the amount of moonlight. They counter being silhouetted by the moon to avoid detection by the eyes of upward-glancing prey. The problem is that squid don’t have the genes to directly produce light. Instead, they have evolved a symbiosis with Vibrio bacteria that can produce light. The Vibrio–squid symbiosis is a classic example of curiosity-driven basic research aimed at understanding life on our planet. Learn more about the symbiosis from two researchers at the University of Wisconsin in a feature at www.hhmi.org/biointeractive/biodiversity/Bioluminescence01.html (Figure 3). Margaret McFall-Ngai and Ned Ruby have devoted their careers to understanding the squid–Vibrio symbiosis. Students at Davidson College have put together a website on bioluminescence that includes information on the symbiosis as well (www.bio.davidson.edu/people/midorcas/animalphysiology/websites/2005/plekon/
in her iBioSeminars (www.ibioseminars.org/lectures/...sufficient population. Bonnie presents her science at length
factors and coordinate attacks on their hosts when they have

to coordinate light production, but to turn on virulence
Most bacterial species use quorum-sensing signaling, not

simply cycles of population growth. When bacteria multiply
to a certain density, the community begins to produce light.
The mechanism for this coordinated activity has come to
have the catchy name “quorum sensing.” Bonnie Bassler,
now at Princeton University, had heard about this fascinat-
ing symbiosis and as a microbiologist became particularly
interested in understanding the genetics and biochemistry of
light generation as well as the molecular signals for turning
the lights on and off. Ultimately Bonnie turned to a free-living
luminescent species of Vibriobacteria that are easier to culture.
You can view an animation of the signaling phenomena in
action at www.hhmi.org/biointeractive/biodiversity/2009
_QS_molecular_cascade.html (Figure 4) as well as a molec-
ular animation of its genetic control at www.hhmi.org/bio-

As a budding scientist launching a career, Bonnie had chal-

lenges for those who dismissed bacterial light production as

an interesting phenomenon not in the mainstream of biomed-
cal research. The PBS program NOVA scienceNOW has an ex-
cellent video profile of Bonnie on its website (www.pbs.org/
wgbh/nova/body/bonnie-bassler.html). Bonnie talks about
the challenges, her good fortune, and her desire for her very
basic research to lead to new drugs. In this regard, Bonnie has
had good instincts and good fortune. It appears that quorum
sensing is a universal phenomenon among bacteria of every
species studied, including many important pathogenic
species from Vibriocholera to Salmonella and Escherichia coli.
Most bacterial species use quorum-sensing signaling, not
to coordinate light production, but to turn on virulence
factors and coordinate attacks on their hosts when they have

sufficient population. Bonnie presents her science at length
in her iBioSeminars (www.ibioseminars.org/lectures/

chemicalbiologybiophysics/besslar.html), in her Holiday
Lectures (www.hhmi.org/biointeractive/biodiversity/
lectures.html), and in an energetic 18-min TED talk
that should appeal particularly to students (www.ted .com/talks/bonnie_bassler_on_how_bacteria_communicate
.html).

A main thrust of the quorum-sensing drug development
research is to identify molecules that can disrupt quorum-
sensing systems of pathogens and therefore perhaps disrupt
the production of toxins. The Bassler lab and collaborators
are screening large chemical libraries and have identified
some interesting compounds that disrupt quorum sensing and
when administered to mice have had some success in
fighting off bacterial infections. This research is now in the
highly unpredictable preclinical phase of medicine develop-
ment.

The first story described basic research to clinical ap-
plications while the second story is still in the preclinical
phase of drug discovery. My third story relates to a
blockbuster drug, one that’s making a lot of money and
saving a lot of lives. Cardiovascular disease is a leading
cause of illness and death globally. Perhaps it’s no surprise
then that the biggest blockbuster drugs of all time are the
statins, developed to fight atherosclerosis, a leading factor
in heart disease. The PharmaLive website has a report on
the world’s best-selling medicines (www.pharmalive.com/
special_reports/sample.cfm?reportID=314). Over 40 million
people take statins worldwide, generating many billions of
dollars in revenue for the pharmaceutical industry annually.
Despite some leveling off of individual brands, the class con-
tinues to dwarf the sale of more famous drugs, like Viagra.
Understanding how statins work and developing new statins
that are more effective with fewer negative side effects are
highly dependent on basic research. However, our initial
interest in controlling cholesterol levels and developing
statins originated in the field of clinical medicine.

Statins act to lower the level of serum cholesterol. Choles-
terols are lipids, and many students find lipid chemistry
difficult, or at least hard to remember, perhaps because hy-
drophobic molecules get short shrift in the science curricula
on a watery planet. Of course their hydrophobicity is what
makes them so important and interesting. Without lipids


Figure 4. Bacteria use a signaling sys-

Figure 4. Bacteria use a signaling sys-


index.htm). The Microbial Life Educational Resources pages
also have information on the squid–Vibriosymbiosis (http://
serc.carleton.edu/microbelife/topics/marinesymbiosis/
quidi-vibrio/collection.html), as do the Why Files web-

pages (http://whyfiles.org/2010/sustaining-symbiosis-new-
class).
there would be no cellular and subcellular compartments, no way to sequester all that watery chemistry, not to mention the interesting signaling functions of lipids. I think cholesterol is a good entry point into lipid chemistry because students can be motivated by the health connections. Wiley has a good interactive tutorial on cholesterol that you can find at www.wiley.com/college/boyer/0470003790/animations/cholesterol/cholesterol.swf (Figure 5). The graphics are highly schematic, but it's a good overview of cholesterol with some emphasis on aspects that relate to heart disease and statin action. The occasional pop-up quizzes help keep students' attention and discourage just clicking through. It's good that this feature puts cholesterol in the context of the steroid family of molecules, but it's unfortunate that they don't put them in the broader context of fat and lipid molecules. William Reusch in the Department of Chemistry at Michigan State University has put together a nice online primer on organic chemicals, including lipids (www2.chemistry.msu.edu/faculty/reusch/VirtTxtJml/biomol.htm). It's not interactive or graphics driven, but it has excellent information, clearly and concisely presented. Satoshi Amagai has developed a pair of very nice features for the BioInteractive website on the molecular structure of fats (www.hhmi.org/biointeractive/obesity/obesity_molecular/01.html) and how the body uses fat (www.hhmi.org/biointeractive/obesity/obesity_processing_fat/01.html). The "How the Body Uses Fat" feature in particular helps clarify that the much discussed "good" (HDL) and "bad" (LDL) cholesterols are not cholesterol; they are large lipoprotein particles that transport cholesterol through the blood. Our understanding of the chemistry of cholesterol, its metabolism, and its physiological regulation is a triumph of basic research. But importantly, the earliest concerns about the potential health hazards of cholesterol came from physicians. To understand the relationship between clinical and basic research, it's necessary to consider some history.

By the 1950s and 1960s, research had associated atherosclerosis with heart disease and established that artery-clogging plaques were composed largely of cholesterol. It was also known that HMG-CoA reductase was the rate-limiting enzyme on the path to making cholesterol. These facts led to a search for an inhibitor of HMG-CoA reductase that could be used to reduce cholesterol in the body and perhaps slow or stop plaque formation. For a review of the plaque formation process, have a look at the animation produced by pharmaceutical company AstraZeneca and available on YouTube (www.youtube.com/watch?v=fLonh7ZesKs). The graphics are much better than the narration. By the early 1970s, drug company employee Akira Endo was screening bacterial and fungal cultures to find inhibitors of HMG-CoA reductase. He soon found a candidate, the first member of the class now known as statins. In 2003 John Simons published an article in Fortune Magazine (http://money.cnn.com/magazines/fortune/fortune_archive/2003/01/20/335643) covering the fascinating commercial aspects of these drugs.

In the late 1960s, physicians Michael Brown and Joseph Goldstein saw their first patients suffering from a severe
form of inherited hypercholesterolemia called familial hypercholesterolemia (FH). These individuals can have serum cholesterol levels 10 times that of an average person, and they develop thick deposits of cholesterol called xanthomas that can be seen on various parts of the body (Figure 6). Brown and Goldstein’s approach was to use this genetic disease as a way to understand the more general phenomena of high cholesterol affecting the general population. They were also interested in the basic research problem of how insoluble cholesterol could be delivered to cells—“the delivery problem.” I recommend visiting their Nobel Prize webpages and in particular reading the transcript of their Nobel lecture (http://nobelprize.org/nobel_prizes/medicine/laureates/1985/goldstein-lecture.html). Their lab website presents a good short history of this work as well (www4.utsouthwestern.edu/moleculargenetics/pages/gold/past.html). Although I think Brown and Goldstein would probably call themselves disease-oriented researchers, it’s significant that they sought training in basic research and have followed a basic research approach to understanding disease. As a result, their work established important concepts in cell biology, a case of medically oriented research contributing to basic research advances. Their cholesterol work has elucidated receptor-mediated endocytosis, recycling of membrane receptors, and feedback regulation of receptors. These principles are outlined in a review article on their website (www4.utsouthwestern.edu/moleculargenetics/pdf/msb_cur_res/2009%20ATVB%20Brown%20431.pdf).

Brown and Goldstein discovered the answer to the delivery problem: Cells had receptors on their surface that bound cholesterol-rich LDL particles. Bound particles were subsequently internalized and processed (coated pits and vesicles) to make the cholesterol available to the cell to make new membrane. Once separated from the LDL, the receptor could be recycled to the cell surface. The simple animation found on the W.H. Freeman website (http://bcs.whfreeman.com/thelifewire/content/chp05/0502003.html) illustrates endocytosis and recycling of LDL receptors, but not feedback regulation. Cells deprived of cholesterol increase the number of receptors on their surface and decrease the number of receptors when cholesterol is plentiful. Goldstein and Brown’s hook for finding and cloning the LDL receptor was that their FH patients had a paucity of LDL receptors. Understanding these various genetic and biochemical equilibria in the context of cellular function led to a hypothesis that was experimentally validated. Lowering the cholesterol content in liver cells could up-regulate LDL receptors, providing more receptors for taking LDL out of the bloodstream, thus lowering serum cholesterol levels and inhibiting plaque formation. Brown and Goldstein’s work has been essential to understanding the action of statins and forms the rational basis for ongoing statin development. They received the 1985 Nobel Prize in Physiology or Medicine, and the lab continues to work on cholesterol regulation.

There are so many more stories that could be used to hone students’ appreciation for the multifarious dimensions of the research enterprise, and I think these stories help students realize how many different sorts of research and research careers there are. Cancer is one story that really needs an article all its own. There is continuing controversy about the war on cancer that dates to Nixon’s 1971 National Cancer Act. Because of the devastating and far-reaching nature of cancer, the government and various nongovernmental organizations are under continuous pressure to be waging a war. The National Cancer Institute has published a useful cancer timeline (http://dtp.nci.nih.gov/timeline/noflash/index.htm) emphasizing therapeutics. There are many versions of why the war on cancer failed: Cancer is so many separate diseases; the research emphasis was wrong; environmental factors can cause cancer. A much-discussed article published in Fortune Magazine in 2004 (http://money.cnn.com/magazines/fortune/fortune_archive/2004/03/22/365076/index.htm) blamed the failure on faulty animal models, in particular mice, while a 2007 Washington Post article suggested that the $100 billion spent on cancer drug development is wasted because cancer is predominantly an environmental disease (www.washingtonpost.com/wp-dyn/content/article/2007/11/02/AR2007110201648.html). A 2008 Newsweek article
All of these popular press articles present a piece of the story that has some truth but entirely misses the nature of basic research. I think this also underlines why waging war on a disease is not a good metaphor for what research does. Dueling metaphors aside, it probably was premature to focus on eliminating cancer in the 1970s before we understood enough about the biology of cancer and before we had enough information from basic research. Due to a large body of research done on model systems starting in the 1970s, we now understand that what all cancers have in common is faulty cell cycle regulation. Because of the Nobel Prize–winning basic research of Paul Nurse, Leland Hartwell, Tim Hunt, and many others on cell cycle regulation (http://nobelprize.org/nobel_prizes/medicine/laureates/2001/nurse-lecture.html), we understand the biology of cancer and finally have a chance to develop rational therapies. As a 2010 article published in the journal PLoS ONE documents (www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0009584), deaths due to cancer are declining in the United States. Although these gains are most likely due primarily to reduced smoking, basic research has helped and is likely to accelerate gains as therapies for specific cancers become available.

I’ll close with a short list of websites of related interest:

- FASEB has a Breakthroughs in Science section (www.faseb.org/News-and-Publications/Breakthroughs-in-Science.aspx) with some good examples, including hypertension research. I think the common media usage of “breakthrough” is a problem, but I understand why we fall back on that language.

- The NIH National Institute of General Medical Sciences has a good website describing the use of model organisms in research (www.nigms.nih.gov/Publications/modelorg_factsheet.htm).

- The University of Wisconsin has a website that focuses on Caenorhabditis elegans in teaching but also includes information on other model systems and the general approach of using model organisms (www.wormclassroom.org/teaching-model-organisms).

- A thoughtful EMBO report is available (www.nature.com/embor/journal/v9/n8/full/embor2008142.html) that discusses animal models and their future as human stem cell science and computer modeling matures.

- In a Science Perspective (www.sciencemag.org/content/307/5717/1885.full), Stanley Fields and Mark Johnston have authored a succinct and specific statement of how model organisms will continue to be important for biomedical research in the coming decades.


- BioCentury published a very good report on the politics of the 2011 NIH budget (www.biocentury.com/promotions/budgetfight/us-budget-fight-over-basic-translational-research-spending-by-nih-a1.htm) and the political battle between basic and translational research priorities.

- The Science Coalition, primarily representing U.S. research universities, has published a thorough report on how basic research pays an investment dividend in jobs and new companies (www.sciencecoalition.org/successstories). The report points out that 55% of basic research takes place at universities and that 60% of the funding is federal.

- The Journal of Clinical Investigation publishes a review series (www.the-jci.org/publiTron.php?list=review_series) that presents bundles of articles on particular diseases or organs. Reproductive Biology is an interesting example; it includes articles on reproduction in placental mammals focusing on getting bench-to-bedside insights.

- The NSF has been conducting an annual poll of citizen attitudes toward science and engineering for many years, and the results are quite consistent. Americans trust scientists and think science is important, but they don’t understand science very well. The 2010 report is available online (www.nsf.gov/statistics/seind10/c7/c7s3.htm).

- Virginia Commonwealth University conducts an annual survey focused on attitudes toward the life sciences, especially controversial issues like cloning and stem cells. The most recent report is available online (www.vcu.edu/lifesci/centers/cen_lse_surveys.html).


Email me at dliu@hhmi.org to tell me your favorite bench-to-bedside story.

ACKNOWLEDGMENTS

Thanks to Malcolm Campbell for helpful editorial suggestions.
Life Sciences Career Exploration
Louisa A. Stark

Is there a career that combines the environment, genetics, and engineering? What is a theriogenologist, and what education and skills does one need for this career? What careers can one enter with an associate’s or a bachelor’s degree? What careers have the best outlook in terms of projected new positions, and what salary ranges can one expect? This article provides an overview of free, online resources for career exploration, with a focus on life sciences careers. This review is intended to be of use to secondary and postsecondary students seeking information about careers in this field of science. Hopefully, teachers and faculty members advising students who are planning their career goals and majors will find this a helpful resource to share with students.

All of the online resources in this review include the following information for each career in their database: a description of the types of work involved, the level of education needed, the median income, the outlook for employment, and links to related careers and more information, such as professional society websites. Additional types of information and resources are noted in the reviews.

GENOMIC CAREERS
One of the newest resources is Genomic Careers (www.genome.gov/GenomicCareers/index.cfm; Figure 1), developed by the National Human Genome Research Institute at the National Institutes of Health (NIH), with input and feedback from student users. A short video (1:24 min) on the home page provides an introduction to genomics careers and the resources on the site. One of the most engaging ways to explore these careers is via the 52 videos. Most of the videos are interviews (0:49–14:30 min) with individuals representing various careers, conducted by three college-age interviewers. Five videos (2:45–20:47 min) provide a peek into working environments in industry, academia, and the government via tours. Videos also provide an introduction to DNA testing (7:56 min) and BLAST software (2:25 min). The longer videos offer an interactive option: a list of questions the user can select to view sections of the movies that provide answers. The videos have a closed-captioning option, and downloadable transcripts are available. The Career Profiles section of the site allows users to sort careers based on five areas of interest, nine types of careers, and median income. The results provide a one-sentence description of each career, education level needed, and median income, with a link to more detailed information.

Genomic Careers has two unique features. First, the Genomics Challenge allows users to assess their knowledge of genomics careers. Second, users can register for the Career Tracker, which sets up a page for them to keep track of their career exploration. Once users have signed up, they can rate each video and career profile they visit on a scale of one to five stars. Based on their ratings, the site provides a summary of their interest level in each of the nine career categories, and lets them know the categories in which they have shown the highest interest. The MyFavorites and MyStatistics features help users track the careers they have rated. The Career Tracker home page also includes a link to a page for educators and advisors that provides an overview of the site’s features. Visitors to this well-designed site will have the opportunity to hear from people in a wide range of genomics-related jobs, allowing them to gain insight into careers in this field.

HEALTH CAREERS
Students interested specifically in health-related careers will want to visit the ExploreHEALTHCareers.org website (http://explorehealthcareers.org/en/home), hosted by the American Dental Education Association. Users can narrow their search of 119 careers by selecting a minimum salary, maximum number of years of education required, and career field. The academic requirements section for each career
Figure 1. Genomic Careers website from the NIH National Human Genome Research Institute.

Figure 2. LifeWorks website from the NIH Office of Science Education.
provides recommendations for courses and extracurricular activities in high school and college courses and degrees, in addition to postbaccalaureate degrees and certifications. All of this information can be downloaded as a PDF. Interviews with six professionals and five students provide insights into several of the careers.

ExploreHEALTHCareers.org includes two unique features in the Resources section of each career profile. One feature is a link to enrichment programs related to the career, such as summer programs for high school students and internship programs for college students. The second feature is a link to funding opportunities related to the career, such as scholarships, fellowships, and loan forgiveness programs. Both features include a description of each program, eligibility criteria, a link to the program website, and contact information.

Other interesting features of ExploreHEALTHCareers are a blog with responses to site visitors’ questions, an FAQ with answers to questions such as “Do health workers ever get laid off?,” and a section on issues in healthcare, with links to articles on health and healthcare disparities, workforce diversity, cultural competence, humanism in healthcare, and health policy. The News & Articles section on the home page provides articles on useful topics, such as how to apply for financial aid. Visitors to ExploreHEALTHCareers have the opportunity to learn about a broad range of health-related careers and to access information about enrichment opportunities and financing their studies.

LIFE SCIENCES CAREERS

The LifeWorks website (http://science-education.nih.gov/LifeWorks; Figure 2), developed by the Office of Science Education at the NIH, provides information on 128 life sciences careers. Visitors can use the Career Finder feature to search for careers that best align with particular job categories, as well as with their interests and skills. After users mark selections on each page, a status box lets them know how many careers match their selections and provides a link to the search results. Definitions for each of the six interest areas and 49 skills assist students with their selections. One can also browse careers based on the level of education required, interest areas, and median salary, and alphabetically. Unfortunately, the Career Finder navigation is a bit frustrating. The only way I found to get back to the list of careers Career Finder had identified for me was by using the back arrow, which meant I had to click back through the sections for each career I perused.
In addition to the standard types of information on each career, LifeWorks includes recommended high school courses for particular career paths, and a list of postsecondary degree programs that prepare one for a specific career. Some career profiles include links to interviews with individuals in the career. I particularly liked that the interviews include not just why individuals chose their careers, their typical day, and what they like best and least about their work, but also a section about what each person likes to do outside of work. Photographs illustrate each part of the interviews, and some interviews include short movies (1:41–2:06 min) with a closed-captioning option. A printer-friendly version of each career profile is available.

For secondary school students, the site provides extensive information on preparing for college, with information on what students need to be doing each year in grades 8–12. The latter two years are broken down into month-by-month lists. Visitors to LifeWorks will be introduced to a wide variety of life sciences careers, with a peek into the daily lives of people already in these careers.

Personalized advice and information about life sciences careers is available through the LifeWorks E-mentoring program (http://science.education.nih.gov/LifeWorks/Ementoring). Students who register for this free service can browse the list of prescreened mentors and request the individual who might be a best fit for them. All email communications between mentors and students pass through a private and secure NIH website.

STATE-SPECIFIC CAREER INFORMATION

CareerOneStop (www.careeronestop.org; Figure 3) is sponsored by the U.S. Department of Labor. CareerOneStop covers a very broad range of occupations, including ones that do not require a high school diploma or postsecondary education. Visitors can browse the database of more than 600 careers in several ways, including by occupation type, employment trends, jobs most in demand, and “green” careers. A unique feature of CareerOneStop is that the career profiles are individualized for each state. Each overview includes a brief description and a video, which can only be viewed using Windows Media Player or Real Player. The videos are narrated, closed-captioned, and ~1:30 min; many are also available in Spanish. Users select how much additional information they want to view about each career. Among the options are tables that allow the user to compare wages and employment trends between a selected state and the nation overall. In addition to career-specific information, the site includes sections on conducting a job search, writing resumes and cover letters, and preparing for an interview.

ADDITIONAL INFORMATION SOURCES

Many professional societies have career sections on their websites. A browser search for the name of the career (for example, “genetic counselor”) or the scientific field and the word “society” (for example, “neuroscience society”) will provide links to these societies.

ACKNOWLEDGMENTS

I thank A. Malcolm Campbell for critical comments and helpful suggestions on this article.
Feature
From the National Science Foundation

Undergraduate Research Experiences in Biology: Alternatives to the Apprenticeship Model

Cynthia A. Wei and Terry Woodin

National Science Foundation, Arlington, VA 22230

This is the first in a series of articles exploring some of the approaches advocated in the American Association for the Advancement of Science’s (AAAS) Vision and Change in Undergraduate Biology Education (AAAS, 2011a), an effort within the biology community to address the needs of undergraduate education in the life sciences (Woodin et al., 2009, 2010) in response to the dramatic and rapid transformations in biology in recent decades (National Research Council, 2009). The Vision and Change report describes a number of ways to meet the needs of the 21st-century undergraduate. Here, we address one of the changes advocated in that report—the call to “introduce the scientific process to students early, and integrate it into all undergraduate biology courses.” We review a representative sampling of recent innovations integrating scientific research experiences within the biology curriculum. Most (but not all) of the examples given are drawn from the recent literature and from projects presented at a recent meeting of principal investigators from the National Science Foundation’s Course, Curriculum and Laboratory Improvement/Transforming Undergraduate Education in Science, Technology, Engineering and Mathematics (STEM) program (AAAS, 2011b). We hope that this sampling will provide new insights and ideas that will encourage more faculty members to consider ways to involve their undergraduate students in research. In addition to outlining a variety of approaches being used, this article briefly addresses, first, the way in which different biological subdisciplines (e.g., ecology, molecular biology, genomics) and different types of institutions are incorporating this approach into their curriculum and, second, the outcomes that are beginning to emerge and the tools being developed to document and evaluate these outcomes.

The opportunity to conduct independent research as an undergraduate has often been cited as the compelling experience that launches a scientific career (National Research Council, 2003; A. Roe, as cited in Lopatto, 2010). The benefits of such experiences have been chronicled both anecdotally (Cejda and Hensel, 2009) and in studies that span institutional types and disciplinary approaches (Russell et al., 2007; Lopatto, 2010). The apprenticeship model, in which students conduct independent research projects in an individual faculty member’s laboratory, is a well-established approach to providing independent research experiences. As the demand for undergraduate research experiences increases, the strain on institutions and faculty trying to meet this demand becomes more evident. Whereas this approach is critical for providing students with an inside view of how science proceeds and for socializing them into the scientific community, it requires a great deal of financial and faculty resources. Thus, its reach is very limited. The apprenticeship model is especially difficult for institutions where research is not a large part of their institutional mission, and many students, particularly those from populations currently underrepresented in the STEM professional community, may not seek out these opportunities. The need for alternative ways of bringing the benefits of undergraduate research experiences to students and engaging them in the scholarly community is becoming increasingly evident.

INTEGRATING SCIENTIFIC RESEARCH EXPERIENCES THROUGHOUT THE BIOLOGY CURRICULUM

As faculty and departments recognize that the benefits of engaging in the scholarship of science apply to a broad spectrum of students, they have been finding creative ways of providing these experiences to a broader population. The following sections attempt to describe some of the many approaches being used to accomplish this goal. The projects highlighted here are a small sampling (see Table 1) that collectively represent a spectrum of approaches; these approaches range
Table 1. Undergraduate research experiences: examples of alternative approaches

<table>
<thead>
<tr>
<th>Project, institution, principal investigator</th>
<th>Project description</th>
<th>Outcomes</th>
<th>Comments, subject area, website</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Introducing students to research; demystifying science and scientists through close investigations of individual research labs</strong></td>
<td><strong>Project, institution, principal investigator:</strong> Consider, Read, Elucidate the hypotheses, Analyze and interpret the data, and Think of the next Experiment (CREATE), City College of the City University of New York, Sally Hoskins</td>
<td><strong>Focuses on the research of an individual scientist to give students a sense of how scientific progress is made and how research unfolds over time.</strong> Using novel and adapted pedagogical tools, students read and analyze a series of four research papers from a single laboratory, design their own follow-up experiments, vet these in a “grant panel” activity, interact with the laboratory involved by posing a set of questions in the form of an email survey sent to each author. Thoughtful responses from authors illuminate “the research life.”</td>
<td><strong>Student Assessment of Learning Gains (SALG) surveys indicate students develop greater interest in science and greater confidence in their ability to understand science. Critical Thinking Tests (adapted from Field-tested Learning Assessment Guide) and class observations indicate improved student scientific problem-solving and critical thinking abilities. Postcourse interviews indicated gains in understanding of “who does science and why,” decreased misconceptions about the research life, increased confidence in ability to become scientists, and increased enthusiasm for research careers. Outside evaluations of CREATE implementations on multiple campuses found student gains in critical thinking abilities, understanding the nature of science, and attitudes toward science and scientists.</strong></td>
</tr>
<tr>
<td><strong>Connecting Researchers, Educators, and Students (CREST), Center for BioMolecular Modeling, Milwaukee School of Engineering, Tim Herman</strong></td>
<td><strong>Undergraduate student teams interact with a research lab investigating an intriguing protein and learn about the research and the protein; develop a physical model of the protein; and work closely with an undergraduate educator to develop instructional materials that incorporate the physical model and the research as a means of highlighting protein function and structure.</strong></td>
<td><strong>Proposed outcomes include student understanding of complex protein structure and function and the process and culture of science.</strong></td>
<td><strong>Physical models are available for loan. Subject area: protein structure and function <a href="http://cbm.msoe.edu/stupro/crest/index.html">http://cbm.msoe.edu/stupro/crest/index.html</a></strong></td>
</tr>
</tbody>
</table>

**Experiencing the process of science: integrating scientific research into the student laboratory**

**Approaches that are based on and contribute to faculty research**

| Authentic Research Experience in Microbiology (AREM), Brooklyn College of the City University of New York, Theodore Muth | In AREM courses, students isolate a strain of the plant pathogen Agrobacterium tumefaciens from an environmental sample they have collected; examine its infectivity using Arabidopsis root segments; and use bioinformatics tools to relate genomics to host-pathogen interactions. | Preliminary results suggest that content knowledge, as measured by standard exams, is equal for those in AREM sections and those in traditional labs (AAAS, 2011b). However, as measured by the California Critical Thinking Skills Test, students in AREM courses showed gains in some aspects of critical thinking compared with those in traditional labs. | Example of a course-based research experience at a diverse, urban college with a dominant commuter population. In a recently introduced urban metagenomics approach designed to study urban bacterial community dynamics, students analyze 16S RNA sequences amplified from samples they collect from local sites. Subject area: microbiology approach can be adapted to a variety of disciplines or research subjects. |  

(Continued)
### Table 1. Continued

<table>
<thead>
<tr>
<th>Project, institution, principal investigator</th>
<th>Project description</th>
<th>Outcomes</th>
<th>Comments, subject area, website</th>
</tr>
</thead>
<tbody>
<tr>
<td>Project Laboratory in Genomics and Genetics, Brandeis University, Waltham, MA, Susan Lovett</td>
<td>Engages students in research through a three-stage approach: 1) read primary literature and complete training exercises to learn key concepts and techniques, 2) collaboratively develop and conduct group projects, and 3) individually develop and conduct independent research projects. Students isolate random <em>Escherichia coli</em> transposon mutations affecting rates of genetic variation; analyze these mutants to discover functions essential to genetic stability; integrate their findings with information found in public domain genomic information resources; and write a research paper reporting their results.</td>
<td>On the SURE survey, students self-report increased understanding of the process of scientific research, increased interest in science, and enhanced reading and writing skills. Using a principal investigator–generated survey to find out how the course helped students, the students report that the writing component enhanced their understanding of course content and general writing skills (AAAS, 2011b).</td>
<td>Principal investigator reports that the percentage of underrepresented minorities in the course is higher than the college’s or major’s percentage. This approach seems to have a beneficial effect on their feeling of being able to be a scientist. Subject area: genomics approach can be adapted to a variety of disciplines or research subjects. For example, a Neurobiology Project Lab has been developed for 2011.</td>
</tr>
</tbody>
</table>

### Approaches that provide opportunities for students to participate in original collaborative research and/or contribute to a broader research effort

| Subject area: genomics | A two-semester course supported by the Howard Hughes Medical Institute. In semester one, students isolate and characterize bacteriophage from local soils and name the newly identified life form; extract and purify its DNA; and send one phage DNA sample to the U.S. Department of Energy Joint Genome Institute for sequencing. In the second semester, students use bioinformatics tools to annotate the sequenced genome. | Pooled survey data from the first-year’s cohort indicate that students taking SEA labs, compared with their peers in introductory laboratory courses, were more likely to complete their course (only 2–5% dropped out vs. a school-wide average of 14%) and scored better on exams in introductory biology courses by an average of 6 of 100 points. Both of these findings were true regardless of school size or whether the students were honors, at-risk, biology majors, or undeclared majors (www.hhmi.org/news/SEA20091217.html). | SEA provides instrumentation, reagents, and protocols and support for faculty through training, workshops, and a learning network. The undergraduate research has resulted in a peer-reviewed paper with 192 authors, most of them undergraduates (Pope et al., 2011). By 2012 the course will be offered at 60 schools in 29 states and Puerto Rico (www.hhmi.org/grants/sea/institutions.html). Subject area: genomics bacteriophage www.hhmi.org/grants/sea |
| Subject area: genomics | A collaborative venture providing access to databases, Web-accessible tools, curriculum materials, and other resources to support: undergraduates in upgrading draft-quality *Drosophila* genomic sequences to high-quality and/or thoughtful and detailed annotation of these sequences, generating curated gene models; pooling of data and ideas from student courses in many institutions to result in significant improvement in existing databases; and joint publication in the scientific and science education literature. | Based on a postcourse survey adapted from Lopatto’s SURE and CURE surveys, students reported professional and learning gains similar to students in apprenticeship experiences (Lopatto et al., 2008). | This project has resulted in peer-reviewed papers in both scientific (Leung et al., 2010) and educational (Shafer et al., 2010) journals. Subject area: genomics http://gep.wustl.edu |

(Continued)
<table>
<thead>
<tr>
<th>Project, institution, principal investigator</th>
<th>Project description</th>
<th>Outcomes</th>
<th>Comments, subject area, website</th>
</tr>
</thead>
<tbody>
<tr>
<td>Community College Genomics Research Initiative, Bellevue Community College (ComGen), in partnership with the U.S. Department of Agriculture’s Agricultural Research Service group at Washington State University, Pullman, WA, Gita Bangera</td>
<td>The student laboratory centers on sequencing of DNA from the biocontrol bacterium <em>Pseudomonas fluorescens</em>, which attacks the plant pathogen that causes take-all, a fungal disease of wheat and barley.</td>
<td>On modified CURE surveys, students report gains in scientific skills such as laboratory techniques, understanding how knowledge is constructed, and how to read and understand primary literature; and personal development, including tolerance for obstacles faced in the research process and understanding how scientists solve problems.</td>
<td>Website includes materials so others may emulate this approach. Subject area: genomics <a href="http://scidiv.bellevuecollege.edu/comgen/CCabout.html">http://scidiv.bellevuecollege.edu/comgen/CCabout.html</a></td>
</tr>
<tr>
<td>Partnership for Research and Education in Plants for Undergraduates (PREP-U), Virginia Polytechnic Institute and State University, Blacksburg, VA, in partnership with University of California, Davis, and Richard Bland College, Petersburg, VA, Erin Dolan</td>
<td>This project challenges students to develop an interdisciplinary approach to biology by examining interactions between <em>Arabidopsis</em> (both wild type and mutant) and herbivores at the genetic, biochemical, organismal, and population levels. Students design their own experiments to investigate whether changes in <em>Arabidopsis</em> genes affect their interactions with herbivores; share their findings via video chat or email, and add them to the PREP Online Lab Notebook, a site accessible to scientists and other students (<a href="http://www.prep.biochem.vt.edu">www.prep.biochem.vt.edu</a>).</td>
<td>In progress</td>
<td>The module offers experimental and analytical techniques adaptable for both beginning and advanced laboratories and to the addition of other components such as molecular genetics and bioinformatics. Subject area: general biology; plant biology; behavioral ecology; genetics; bioinformatics. <a href="http://www.prepu.biochem.vt.edu">www.prepu.biochem.vt.edu</a></td>
</tr>
<tr>
<td>Bringing Field Research into the Classroom, Rocky Mountain Biological Laboratory, Ian Billick</td>
<td>Provides improved access to online databases and curricular materials (this part of the project is in development). Enables students to use related long-term climate, weather, and biodiversity data to address ongoing problems of environmental and ecological interest; and track the stepwise progression of a particular research project done by others from planning to publication.</td>
<td>In progress</td>
<td>Introduces students to the work of individual scientists. Subject area: ecology, evolutionary biology, environmental science</td>
</tr>
<tr>
<td>Online Research in Biology (ORB), Cornell Lab of Ornithology, Nancy Trautmann</td>
<td>Facilitates student research using ecology and animal behavior data from online databases such as the Cornell Lab’s collection of citizen-science–generated databases (i.e., eBird, Great Backyard Bird Count) or the Macaulay Library’s collection of animal sounds and videos. Currently creating and piloting Web-based curriculum resources that make use of visualization and analysis tools such as Raven sound software and Science Pipes scientific workflow software.</td>
<td>In progress</td>
<td>Subject area: ecology, conservation, animal communication, and animal behavior <a href="http://www.birds.cornell.edu/orb">www.birds.cornell.edu/orb</a></td>
</tr>
</tbody>
</table>

Projects that provide faculty support and address barriers to adoption and implementation

| Community College Undergraduate Research Initiative (CCURI), Finger Lakes Community College, James Hewlett | Designs, implements, and evaluates a model for integrating undergraduate research into community college science curriculums. Model involves the use of inquiry-based materials and activities in freshman courses, which are then expanded into an undergraduate research experience at the sophomore level. | In progress | Sophomore course with research experience will be credit bearing and transferable to 4-yr institutions. Addresses barriers to undergraduate research in community colleges, such as lack of resources and limited access to research collaborations and networks. |

(Continued)
Table 1. Continued

<table>
<thead>
<tr>
<th>Project, institution, principal investigator</th>
<th>Project description</th>
<th>Outcomes</th>
<th>Comments, subject area, website</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bring Your Own Science in society: connecting research experiences to real-world issues</strong></td>
<td></td>
<td></td>
<td>Approach can be adapted to a variety of disciplines or research subjects. <a href="http://www.ccuri.org">www.ccuri.org</a></td>
</tr>
<tr>
<td><strong>Cassava, University of Puerto Rico, Mayaguez, Dimuth Siritunga</strong></td>
<td>Incorporates research to assess the genetic diversity of cassava populations in Puerto Rico into Genetics and Cell Physiology courses. Students in the Genetics course locate, transport, extract, and quantify DNA and then use simple sequence repeat DNA markers to assess genetic diversity from cassava leaves collected from around their homes. Students in the Cell Physiology course characterize the structure of root cells of cassava varieties and perform analysis on root contents.</td>
<td>Surveys and content tests created by the principal investigator, administered premodule and postmodule, show student gains in both content learning and confidence in a variety of scientific skills, such as “constructing a testable hypothesis” and “designing an experiment to test a hypothesis.” Student data are contributing to an understanding of the genetic diversity of Puerto Rican cassava and cassava conservation (Montero Rojas et al., 2011). Novel cassava accessions found by the students have been added to the Puerto Rican cassava germplasm maintained in vitro in the principal investigator’s lab and in the field at the Isabela Agricultural Research Station.</td>
<td>Reaches a student population that is 99% from underrepresented groups and 59% regarded as economically disadvantaged. Project is being expanded to investigate the sweet potato. Impacts approximately 800 students per year. Subject area: genetics, cell physiology</td>
</tr>
<tr>
<td><strong>Application-Based Service Learning (ABSL), Duquesne University, Pittsburgh, PA, Nancy Trun</strong></td>
<td>Incorporates service learning and uses community-based problems to engage students in scientific research. Course has two components: First, students learn about a specific community-based problem in the service portion of the course. Then, in laboratory classes they conduct research to help understand and solve the problem.</td>
<td>Pre- and posttests, surveys, and content exams show preliminary results including a 40% increase in retention of knowledge 5 mo postcourse when compared with a lecture-only class (55% retention without ABSL vs. 95% with ABSL). Student evaluations indicate strong positive attitudes about the course (100% would recommend the class) (AAAS, 2011b).</td>
<td>Approach can be adapted to a variety of disciplines or research subjects. <a href="http://serc.carleton.edu/sencer/application-based_service/index.html">http://serc.carleton.edu/sencer/application-based_service/index.html</a></td>
</tr>
<tr>
<td><strong>Community-Based Participatory Research (CBPR), Crow Environmental Health Steering Committee, Little Big Horn College, Mari Eggens, Montana State University, Bozeman, Anne Camper; University of New England, Biddeford, ME, Tim Ford</strong></td>
<td>CBPR projects are incorporated into the life sciences curriculum and undergraduate research and internship programs. Students do standard water quality tests, as appropriate, on water from local rivers and wells (pH, temperature, conductivity, coliforms/ E. coli, etc.), and deliver samples to another lab for inorganics and metals testing. This summer they will be adding the use of polymerase chain reaction to detect Cryptosporidium. Findings are related to local health risks, and results are disseminated in the community.</td>
<td>When this project began in 2006, few tribal members had undergraduate degrees in biological or environmental sciences and none had graduate degrees in the field. Now persistence rates to degrees for research interns are near 100%. Twelve Little Big Horn College students are earning 4-yr degrees, and two have earned master’s degrees in the disciplines. The students and the community have developed an understanding of risk assessment and testing methodologies and an appreciation of local water issues.</td>
<td>Research involves students collaborating with a local steering committee of community members; researchers from Montana State University, the University of New England, and the University of Wyoming, Laramie, WY; federal agencies such as the U.S. Geological Survey, the Environmental Protection agency, the Fish and Wildlife Service, and the Indian Health Service; and the nonprofit Hopa Mountain, Bozeman, MT, and other organizations. Subject area: natural resources, environmental science, environmental health <a href="http://www.epa.gov/osp/tribes/NatForum10/ntsfl0_3t_Ford.pdf">www.epa.gov/osp/tribes/NatForum10/ntsfl0_3t_Ford.pdf</a>, <a href="http://www.lbhc.edu/waterquality">www.lbhc.edu/waterquality</a></td>
</tr>
</tbody>
</table>
from literature-based investigations that are amenable to lecture classrooms to research activities within the lab, and include research projects that bring students out of the classroom and into their communities to address local problems. Although these approaches may vary in the resources they require and the contexts for which they are best suited, they all share the ability to actively engage and inspire students.

Introducing Students to Research; Demystifying Science and Scientists through Close Investigations of Individual Research Labs

Most students who seek apprenticeship-type research experiences are already self-identified as science majors (Lopatto, 2010), and many students, particularly underrepresented minorities, may not consider participating in apprenticeship-type undergraduate research experiences because they are uncomfortable or unfamiliar with the idea of research. One approach that can reach broader populations of students is to engage students in “thinking like a scientist” and to introduce them to the work of individual scientists in the classroom. This can be accomplished, for example, through close investigations of scientific literature in the classroom, as is done in the Consider, Read, Elucidate the hypotheses, Analyze and interpret the data, and Think of the next Experiment (CREATE) approach (Table 1), in which students study a series of research papers from a single research laboratory and interact with scientists in the focal lab through email surveys, or through meaningful collaborations with scientists, as achieved by students developing physical models of specific proteins of interest in the Connecting Researchers, Educators, and Students (CREST) program at the Milwaukee School of Engineering, Milwaukee, WI (Table 1).

Experiencing the Process of Science: Integrating Scientific Research into the Student Laboratory

There are several approaches to accomplishing the key recommendation of Vision and Change for faculty to integrate research activities into classrooms and student laboratories that can be applied to a broad spectrum of institutional types and subdisciplines within biology. One approach that is often used by faculty with active research programs is to develop projects that integrate teaching with research efforts. For example, the Authentic Research Experience in Microbiology (AREM) project (Table 1) engages students in research projects that contribute to the research objectives of an individual faculty member; in one course, students isolated strains of the plant pathogen Agrobacterium tumefaciens, the subject of the instructor’s research, and explored the biology of this bacterium through a variety of research-based modules. The Project Laboratory in Genetics and Genomics course (Table 1) is another example in which students engage in original research projects that are inspired by local faculty research interests.

Another approach that is gaining momentum is to provide opportunities in the classroom for students to participate in original collaborative research and contribute to a broader research effort. These approaches make research more accessible and affordable for faculty members who do not have active research programs (or are engaged in research that does not translate well to a group project) and can be used in many types of institutions, including community colleges. Genomics represents one such fundamental area of modern biology that has the potential to engage students in cutting-edge research. Genomics research is especially amenable to collaborative research (Shaffer et al., 2010), in part because teaching and research materials are both easily accessible and readily distributed and because raw data can be available at low cost (often free, from nationally funded databases). Furthermore, genomics can be readily adapted to create projects of varying complexity. A growing number of efforts are using genomics to engage students in research in the classroom, including the National Genomics Research Initiative developed by Howard Hughes Medical Institute’s Science Education Alliance (SEA; Table 1), which engages freshmen in a national experiment on bacteriophage genomics; the Genomics Education Partnership at Washington University, St. Louis (Table 1), which currently focuses on a problem in Drosophila evolution, working primarily with upper-level undergraduates (Lopatto et al., 2008; Shaffer et al., 2010); and the Community College Genomics Research Initiative at Bellevue Community College, Bellevue, WA (Table 1), which focuses exclusively on bringing genomics research into community college biology classrooms. This collaborative approach is not limited to genomics. For example, the Partnership for Research and Education in Plants for Undergraduates (PREP-U) project (Table 1) engages students in studying interactions of Arabidopsis with small herbivores at the genetic, biochemical, organismal, and population levels; students share data online with scientists at a distance via video chat, email, or the PREP Online Lab Notebook (www.prep.biochem.vt.edu).

In other fields where research is typically conducted in natural settings, integrating research opportunities into courses is often challenging. However, newly developing online approaches offer economical and accessible ways to bring field research into classrooms and laboratories. For example, at the Rocky Mountain Biological Laboratory (RMBL), Crested Butte, CO (Table 1), cyber resources are being developed that will enable students to access the extensive collection of datasets at RMBL, including those mapping regional climate, weather, and biodiversity trends, as well as supporting materials for inquiry-based activities. The Cornell Lab of Ornithology, Ithaca, NY, is also developing online resources for cyber-enabled research investigations. The project, Online Research in Biology (Table 1), provides access to the Cornell Lab’s world-class databases about birds, other animals, and their habitats, including citizen-science-assembled bird population data and the Macaulay Library’s archives of animal sound recordings and videos. Cyber-based approaches may lack some of the appeal of fieldwork, but they offer students unique opportunities to engage with data gathered across time and from multiple, often remote sites while avoiding the logistical challenges to gathering such data. As online collections of scientific data continue to grow, so do the accessibility and appeal of this approach.

Science in Society: Connecting Research Experiences to Real-World Issues

One of the core competencies outlined by Vision and Change for undergraduate biology students is the ability to understand the relationship between science and society. Research
experiences that are connected to real-world issues and problems help to develop this competency. For example, in the Bring Your Own Cassava project (Table 1), students engage in research that is literally in their backyards and contribute to the understanding of an important local issue—the conservation of a common food in Puerto Rico, the cassava plant (Montero Rojas et al., 2011). Other projects explicitly integrate student research experiences with societal issues; in the Application-Based Service Learning project (Table 1), students first take a service-learning course to understand the problem that they will then research as part of a subsequent biology course. Similarly, the Community-Based Participatory Research projects1 (Table 1; Cummins et al., 2010) immerse students in the scientific and social dimensions of a local issue (e.g., water quality) by collaborating with community stakeholders throughout the process of conducting their research on the problem. This approach can be a very powerful way to engage students in biology and to help them appreciate why an understanding of biology matters to their lives.

CROSS-CUTTING ANALYSES

Tools to Aid Resource-limited Institutions

A number of resources have been instituted to help institutions whose main mission is not research to offer their students the opportunity to participate in a course-based research experience.2 These are probably most fully developed in the field of genomics and include rich resources offered by the SEA phage project, such as reagents and instrumentation, instructional materials, technical support, shared databases, and opportunities for networking; the databases, instructional material, and opportunities for data sharing offered by the Genomics Educational Partnership fly chromosome project; the databases and technical training offered by the Interpret a Microbial Genome Annotation Collaboration Toolkit project (www.jgi.doe.gov/education/annotation_tools.html); and the background materials and microarray support offered by the Genome Consortium for Active Teaching (GCAT) project at Davidson College, Davidson, NC (www.bio.davidson.edu/GCAT).3 In addition, as described earlier, online resources are opening venues for classroom-based research investigations of field-based phenomena by using the growing collections of online biodiversity databases. Computer-based simulations and analytical challenges are also available from a variety of sources to help develop research skills and thinking; BioQuest (www.bioquest.org/BQLibrary) is an excellent example of this type of resource.

A growing number of efforts are also providing support to community colleges to help their faculty incorporate student research into the curriculum (Cejda and Hensel, 2009). The Community College Undergraduate Research Initiative (CCURI) at Finger Lakes Community College, Canandaigua, NY, is developing a model to facilitate this goal. The model takes into consideration several of the barriers faced by community colleges, including the issue of student transfer to 4-yr colleges; in the CCURI model, research experiences provide course credit that is transferable to 4-yr institutions. Other projects address this issue by forming partnerships between 2- and 4-yr institutions to enrich the research experience, as is the case for the Community-Based Participatory Research project. Such partnerships can also facilitate continuity for transfer students, as illustrated by the California State University (CSU), Chico, and Butte Community College, Oroville, CA, partnership, where a two-tiered set of labs in cell and molecular biology provides a smoother transition for students from Butte who transfer to CSU Chico. These students, having taken the first-tier lab at Butte, can enter the second-tier lab with the same background in the course topic as their CSU Chico classmates.

Suitability across Institutions and Biological Subfields

This small sampling of both mature and newly developing projects that engage undergraduates in original research illustrates approaches that are feasible at a wide range of institutions, from community colleges to research-focused universities and from small private institutions to large public ones, including minority-serving institutions. Examples given in Table 1 range across a variety of biological topics and disciplines, from the cellular and molecular levels (genetics, genomics, microbiology) to the organismal and population levels (ecology, animal behavior). Given the diversity of approaches that faculties have developed and the successes they are documenting, we might surmise that the only limits to finding ways to engage students with biological research are the limits of our creativity; it is no longer possible to say this approach cannot be done on a particular campus or in a particular discipline. Rather, the question that faculty understandably wrestle with becomes whether a particular approach is appropriate for its unique situation and group of students.

Outcomes and Observations

A number of tools have evolved for studying the outcomes of engaging students with the scholarship of the discipline. Several studies have documented the benefits of research experiences of the apprenticeship type (Russell et al., 2007; Junge et al., 2010; Laursen et al., 2010; Lopatto, 2010). These have included, at a minimum, student surveys about attitudes toward science and perceptions of the value of the research experiences. One set of commonly used assessment tools is the Summer Undergraduate Research Experiences (SURE) survey and its cousin, the Classroom Undergraduate Research Experiences (CURE) survey, developed by David Lopatto (Grinnell College, Grinnell, IA). These surveys include pre-course and postcourse questions that ask about science attitudes and, in the postcourse survey, estimates of learning gains and perceptions of benefits from the course. Many of the projects described here report outcomes of SURE/CURE.

1There are five independent undergraduate science research projects at Little Big Horn College, Crow Agency, MT. We highlight one of these projects—the Community-Based Participatory Research project on water quality led by Mari Eggers.

2All of these programs provide websites with instructional and problem-solving activities. Typically, faculty members who wish to join the collaborative effort apply to attend a multi-day workshop to learn the experimental approaches, software tools, and such.

3GCAT now provides support in synthetic biology as well (www.bio.davidson.edu/projects/gcat/Synthetic/synthetic.html).
surveys or adapted versions of these tools; in all of these cases, students reported several benefits from their classroom-based research experiences, and in at least some cases they showed gains similar to those reported by students who had spent a summer in a research apprenticeship (Lopatto et al., 2008). Some of these benefits include enhanced understanding of the scientific process, increased interest in science, and increased confidence in various scientific skills. Many projects have also devised content- and competency-based exams to determine whether engaging in classroom-based research experiences enhances students’ abilities to learn and retain content knowledge and to develop understanding of the emerging core concepts and competencies for biology (AAAS, 2011a). Whereas assembling consistent patterns is difficult, given the diversity of approaches, some interesting observations have emerged. First, evidence is emerging that these approaches are introducing more underrepresented minorities to scientific research in the classroom. For example, in Project Laboratory courses, participation by underrepresented minorities was unexpectedly high, 40% (27 of 67 students) across multiple courses (James Morris, personal communication) compared with approximately 15% in the overall student population (www.brandeis.edu/institutionalresearch/2009pdfs/CDS2009_2010.pdf). Second, in many of the examples highlighted here, a common ingredient of success is identifying a suitable research problem that uses a set of common tools (which can be taught to the students as a group) but can be subdivided to provide students with individual projects. Often these parts are reassembled to derive more informative conclusions. Well-designed projects also provide extensive opportunities for peer interaction and mutual support. These elements likely contribute to a student’s sense of responsibility, of ownership of his or her piece of the project, and of the importance of his or her contribution to a broader picture, both of which may be critical for developing an appreciation for scientific research.

CODA

This article has pointed out a few of the promising emerging efforts to integrate research experiences into academic-year classes. This is by no means a comprehensive review. As noted in Vision and Change, no one approach is the best practice for all situations, all faculties, all institutions, or all students. The examples presented here represent approaches that have the potential to be successful practices in diverse settings. More comprehensive reviews of undergraduate research approaches can be found in two interesting publications, both available online. Each of these publications is wide ranging in the disciplines, approaches, and institutions considered. They each include many biology-related examples. One publication, Developing Undergraduate Research and Inquiry (Healey and Jenkins, 2009), has a somewhat global outlook (mostly Australia, Canada, New Zealand, the United Kingdom, and the United States); the other, a publication from the Council on Undergraduate Research, Developing and Sustaining a Research-Supportive Curriculum (Karukstis and Elgren, 2007), confines itself to work in the United States. Both publications include an extensive discussion of the philosophical and pedagogical basis for the approach presented. For additional examples of class- and laboratory-based undergraduate research in biology the reader is directed to articles in this journal and others concerned with biology education. For example, the January/February issue of the journal, Biochemistry and Molecular Biology Education (Volume 39, Issue 1) highlights innovative laboratory exercises for undergraduates; four of the five examples feature research-based approaches, one of them concentrating on an introductory biology course that includes biology majors and students majoring in other fields (Bell, 2011).

ACKNOWLEDGMENTS

We thank Sarah Elgin for her insightful and quick editing. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

REFERENCES


AAAS (2011b). Transforming Undergraduate Education in STEM: Making and Measuring Impacts, Washington, DC.


C. A. Wei and T. Woodin


Shaffer CD et al. (2010). The Genomics Education Partnership: successful integration of research into laboratory classes at a diverse group of undergraduate institutions. CBE Life Sci Educ 9, 55-69.


This feature is designed to point CBE—Life Sciences Education readers to current articles of interest in life sciences education as well as more general and noteworthy publications in education research. URLs are provided for the abstracts or full text of articles. For articles listed as “Abstract available,” the full text may be accessible at the indicated URL for readers whose institutions subscribe to the corresponding journal.

This themed issue focuses on recent studies about the use of external representations (ERs; visualization formats such as three-dimensional models, graphs, diagrams, pictures, and simulations) in science education and the role they play in the formation of internal representations (mental models and visual imagery). The authors discuss the implications of their findings with respect to understanding students’ difficulties with learning science and for the improved design of science curricula.


Modern college-level biology textbooks include diagrams that depict both cladogenic (branching events) and noncladogenic or anagenic (ordered linear progressions) perspectives on evolutionary history. (Speciation via cladogenesis occurs when a population is split into two or more populations that are subjected to different selection pressures; anagenesis is an accumulation of changes over time that leads to transformation of one species or taxon into another.) Although numerous studies have addressed how cladogenic representations are interpreted by students, this study explores the relatively uncharted terrain of the conceptualizations about evolutionary relationships evoked by noncladogenic diagrams. In a first substudy, 50 university students with a wide range of biology backgrounds viewed a pair of noncladogenic diagrams (one of which depicted “hominid” taxa) and then wrote responses to a set of questions about the evolutionary relationships depicted in them. For analysis, three researchers (working independently) assigned codes to the response data; the codes were used to represent variables within major categories such as the nature of evolutionary relationships, time, characters possessed by the taxa, and evolutionary mechanisms. In a follow-up study with 62 students, interpretations of specific evolutionary terms that appeared in the first set of analyzed responses were probed using questions about an unfamiliar context (a scenario about evolutionary relationships between relatively obscure taxa), with the assumption that prior knowledge would thus be less likely to influence the responses. The article discusses in detail the specific interpretations elicited by the diagrams and scenarios used in the two studies and how well these interpretations aligned with scientifically accepted views on the process of speciation (e.g., cladogenesis as the predominant process). The authors conclude that noncladogenic diagrams, although commonly included in college-level textbooks, can lead to misinterpretation of important aspects of macroevolutionary processes and therefore are educationally inappropriate (for this reason as well as others).

   [Abstract available: www.informaworld.com/smpp/content~db=all~content=a788637377~frm=abslink]

Although the literature contains numerous reports on the facets of student learning that ERs can mediate, and on how students interpret them (in intended as well as unintended ways), the authors contend that there is little empirical research on students’ processing of the many types of ERs...
used in the study of biochemistry. They address this deficit in part by the work reported in this article, which recounts the process of development and validation of a model of the factors that influence students’ ability to conceptualize and reason about the types of diagrams commonly used in biochemistry textbooks. The article describes the five-stage iterative process by which the authors used a “model of modeling” framework adapted from Justi and Gilbert (2002) to develop the model and define its component factors. The authors describe operational definitions for the resulting seven-factor model (three major factors termed “conceptual,” “reasoning,” and “representation model,” and the four possible interactions between these major factors) that emerged from the process and represent the model as a Venn diagram. They validated the model by analysis of data obtained from three 60-90-min audio- and videotaped interviews with nine third-year university students who had just completed a biochemistry module on immunology. Each interview was focused on the subject’s interpretations of one of three different diagrams representing antigen–antibody interactions. The semistructured interview process used a series of questions to probe the subjects’ conceptual knowledge prior to viewing the diagram (Phase 1), their reasoning processes and conceptual understandings during interpretation of the diagram presented to them (Phase 2), and their evaluation and critique of the ER and its utility for their learning (Phase 3). The questions were spontaneously modified and supplemented in the course of the interview to effectively probe each subject’s unique responses to the original, scripted line of questioning. Analyses of the empirical data from the interviews were used to test the robustness of the operational definitions of the model. The article describes selected examples of the data and how they were used in the process of verifying that each factor of the model made an important contribution to students’ ability to interpret the diagrams. The authors conclude by discussing the implications of the model for future research on and educators’ understanding of how students use and learn from biochemistry-related ERs. The authors also emphasize the need for future studies to shed light on the transferability of their model to other areas of science and elaborate on the particular ways in which the model could inform instructional choices.


This study addresses the utility of using student-manipulated physical models and modeling activities in large-enrollment, college-level courses. The study context was seven sections of an earth science course for nonmajors (average section size about 150 students). The study was conducted during coverage of the Sun–Earth relationship and the causes of the seasons, topics that many college-level students find difficult to explain. All sections incorporated use of peer instruction and personal response systems (“clickers”). In five sections, students working in groups manipulated and discussed tennis ball/flashlight models of the Earth and Sun in a series of three activities. The activities were bookended by a premodel concept test and a lecture (about the tilt of Earth’s axis and the nature of its incoming solar radiation) on one end and a postmodel concept test at the other. Students were encouraged to discuss postmodel concept test questions within their groups if the initial responses for the entire section showed an inaccuracy rate of 25% or greater, and then to tackle the question again. In the two comparison sections, students listened to a lecture on the causes of the seasons, completed the concept tests, and engaged in peer instruction related to problematic questions, but did not use the physical models. Incorporation of the models in five of the sections thus augmented use of an interactive pedagogy (informative testing with peer instruction) previously shown to be effective in numerous other studies. Classroom observations and interview data about students’ perspectives on model use supplemented the quantitative data (concept test responses) collected via the response systems. The authors found that the sections using the models showed significantly greater improvement in concept test scores and that the students used and valued the use of the models as an aid to...
understanding the underlying concepts and as a visual reference for recalling them, evidence in favor of overcoming the logistical barriers for “hands-on” model use in large-enrollment classrooms.


This study investigated the effect of participation in a sequence of modeling activities on middle school students’ understanding of the particulate nature of matter. Previous studies have shown that although various types of physical models can help ameliorate the difficulties that the abstract nature of atoms and molecules can present for learning fundamental chemistry concepts, the use of animations alone might not be an effective aid for overcoming these difficulties. This study was therefore designed to tease out the comparative benefits of middle school students’ use of animations in three contexts: as part of a combination of activities in which students constructed, interpreted, and evaluated (critiqued one another’s) animations (Treatment 1); as part of construction and interpretation activities only (Treatment 2); or as part of viewing and interpreting teacher-generated animations. Three seventh-grade teachers from different schools, along with 271 of their students (from eight classes), participated in the study. Students in a given class were randomly assigned to one of the three treatment groups for participation in three chemistry lessons that incorporated use of Chemation as a tool to support their learning. Chemation is a program for visualization, construction, interpretation, and exchange of models and animations of atoms and molecules that runs on handheld computers; all participating teachers and students were experienced users of Chemation. The three treatment groups completed the same pre- and postinstructional chemistry achievement tests, which assessed lesson-related content knowledge and content knowledge in combination with animation constructing, interpreting, and evaluating ability. Students also wrote down their interpretations of the animations they viewed or constructed and their connections to the macroscopic phenomena relevant to the lessons. The researchers used two-factor analysis of covariance to analyze the effects of different treatments on the test scores, and they developed and used two separate coding schemes to assess the quality of the student-generated animations and of the students’ interpretation of animations. They found that the quality of the student-generated animations was higher with Treatment 1 (construction, interpretation, and peer evaluation activities) versus Treatment 2 (no peer evaluation activities). Students’ interpretations of the animations were also significantly better in Treatment 1 than in either Treatment 2 or 3 (interpretation of teacher-generated animations); the lack of significant difference between Treatments 2 and 3 indicated that interpreting ability was as good in students who only viewed animations as it was in students who constructed them. The analysis of the test scores revealed that in this context, student design of animations coupled with peer evaluation (Treatment 1) had a significantly positive impact on student achievement in chemistry (as compared with viewing teacher-generated animations, Treatment 3). However, the authors caution that the results did not favor the use of a “design-only” approach (Treatment 2); in the absence of student engagement in peer evaluation of the animations, the value to students’ learning of chemistry appeared to be on a par with that of viewing teacher-generated animations.

The following are recent CBE-LSE articles on use of ERs:


I invite readers to suggest current themes or articles of interest in life science education, as well as influential papers published in the more distant past or in the broader field of education research, to be featured in Current Insights. Please send any suggestions to Deborah Allen (deallen@udel.edu).

REFERENCE

Feature
Book Review

A Life in Science


Reviewed by Jay Brewster, Natural Science Division, Pepperdine University, Malibu, CA 90263

The process of “becoming” is inherently unique for each young mind that chooses a career of scientific exploration and experimentation. Yet within each scientist’s story are themes of energizing curiosity, fulfillment emerging from discovery, and an instinctual desire for the freedom to explore and investigate. Laura L. Mays Hoopes, the Halstead-Bent Professor of Biology at Pomona College in Claremont, CA, has authored a book highlighting her remarkable path in science over the past 50 yr. In this fascinating memoir, you will find the story of a woman in science, drawn with wonder into the biosciences, specifically the study of DNA and aging. Interwoven through her research career is the career of a teacher, a mentor who has welcomed hundreds of undergraduate students into her research laboratory, and who has taught thousands of students in college courses. This is a woman who distinctively chose a scientific career path in the early 1960s, a time when few women entered PhD programs in research science. This is also a story of her pursuit of balance between career and family, and offers painfully honest details of her struggle for that balance. When I picked up this book, I expected a history lesson; what I found was an inspirational experience, walking with this extraordinary woman through her life as a scientist.

Hoopes earned her PhD in biology from Yale University, served postdoctoral fellowships at Scripps Clinic and Research Foundation and the University of Colorado School of Medicine, and was subsequently recruited to a faculty position at Occidental College in Los Angeles. She has enjoyed a wide-ranging career in research, teaching, administration (Vice President, Dean at Occidental College), and has been a national leader in the development of the Council for Undergraduate Research and the Genome Consortium for Active Teaching. This honest and compelling memoir offers insight into both the historical and modern challenges facing anyone choosing to pursue a career in the sciences, but is particularly effective at revealing the struggle of young women...
within this career track. This book will be valuable reading for any student beginning a science-focused career in research or academia.

Hoopes highlights her beginnings as an undergraduate research student at the Marine Biological Laboratory at Woods Hole, MA. She applied to the Woods Hole summer course in marine ecology during her freshman year at Goucher College. She knew little about the program beyond the colorful brochure, but applied and was accepted. She enjoyed the hands-on nature of the coursework, and ended up working with a scientist from the Woods Hole Oceanographic Institute (WHOI) on an embedded 3-wk research project. She decided to extend her stay at Woods Hole, working as a student research assistant through the end of the summer (and returning the following summer). She was hooked, and began to imagine a career in the research sciences. A devastating disappointment during that first summer at Woods Hole involved an opportunity for students completing the marine ecology course. Students were invited to apply to join the famous WHOI research vessel, Atlantis, during a 1-wk cruise at sea. Hoopes recalls being told that she could not go on this cruise with the other students because the sailors viewed women as “bad luck” on the open seas. Her disappointment over this strange disposition (especially for a research vessel) stands in profound contrast to the encouragement she received from her research mentor during the research project she engaged in at Woods Hole.

In 1963, a letter of interest written to Princeton regarding her potential application to graduate school was met with: “We have not sent the catalog and graduate application which you requested. Unless there is a peculiar need for our facilities, we do not consider admission of women to the graduate program here.” She experienced a similar response when visiting the California Institute of Technology to investigate opportunities for graduate school. With the challenges and frustrations of graduate work, it is hard to imagine the lasting discouragement that must have come from such a letter. Hoopes examines issues related to sexism only through the description of events in her own career; there is little dissection of these events beyond an expression of thankfulness for the changes she has witnessed and hope for further progress.

Hoopes delivers her story as a mentor might discuss her career with a young undergraduate research student, and perhaps this book mirrors the many conversations she has had with her students during her years as a college professor.

She is a national leader in the area of undergraduate research, pedagogical styles in the science classroom, and the development of funding mechanisms to enhance the training of young scientists. I am grateful for her leadership in these areas. I enjoyed learning more about the life of her research laboratory, the style of guidance she provides her students, and her advice for those students. What an extraordinary experience it would have been to join her lab as an undergraduate research student.

At times, this book is hard to read, as Hoopes describes the struggles of being a busy college professor, a mother, and a wife. She tried to “‘have it all,’” refusing to choose career over a healthy personal life, and writes openly about the challenges associated with such a decision. The constant tight wire of time shared among home life, teaching, and research is not an uncommon story for women building an academic career. No simple solutions to these challenges are presented, only the compelling and honest description of the choices she made, and the consequences of such choices. Her story of balance brought up memories of two of my favorite books: Time, Love, Memory by Jonathan Wiener, and Natural Obsessions by Natalie Angiers. These books explore the world of laboratory research and describe researchers searching for some kind of balance between home and the laboratory.

The final chapter is entitled “Who am I really?,” and offers a very personal assessment of her career path and the people who have truly influenced her life, along with an analysis of the critical choices she has made along the way. She considers other women who completed their graduate training with her, and wishes that more of them had become independent researchers or professors. She also examines her own path, defending her choices to teach more than is usual for a typical research scientist; to train undergraduates, who require tremendous amounts of time and energy as they learn the basics of research; and to protect the precious time needed for children and family. Her story is compelling, offering painful details regarding personal frustrations, insecurities, and failure, yet offering a story that is familiar to anyone who has chosen science as a lifelong pursuit. Who is Laura L. Mays Hoopes? She is a model to guide generations of scientists to come. I highly recommend this book to you; the words within will encourage and inspire.
**Meeting Report**

**Society for the Advancement of Biology Education Research (SABER)**

Erika G. Offerdahl,*†‡ Teri Balser,*‡§ Clarissa Dirks,‡|| Kathryn Miller,‡¶ Jennifer L. Momsen,†# Lisa Montplaisir,†# Marcy Osgood,†@ Karen Sirum,†** Mary Pat Wenderoth,††‡‡ Brian White,††‡‡ William B. Wood,††§§ Michelle Withers,|||| and Robin Wright¶¶

*Department of Chemistry & Biochemistry, North Dakota State University, Fargo, ND 58108; §Institute for Cross-College Biology Education, University of Wisconsin–Madison, Madison, WI 53706; |||Scientific Inquiry, The Evergreen State College, Olympia, WA 98505; ¶Department of Biology, Washington University in St. Louis, St. Louis, MO 63130; ‡Department of Biological Sciences, North Dakota State University, Fargo, ND 58108; ‡‡Department of Biochemistry and Molecular Biology, University of New Mexico, Albuquerque, NM 87131; **Department of Biological Sciences, Bowling Green State University, Bowling Green, OH 43403; ††Department of Biology, University of Washington, Seattle, WA 98195; †‡Department of Biology, University of Massachusetts, Boston, Boston, MA 02125; §§Department of Molecular, Cellular, and Developmental Biology, University of Colorado, Boulder, CO 80309; |||Department of Biology, West Virginia University, Morgantown, WV 26505; ‡¶Department of Genetic, Cell Biology and Development, University of Minnesota, St. Paul, MN 55455

**INTRODUCTION**

Biology education research (BER) can be a major contributor to the herculean task of modernizing and transforming biology education. However, as researchers in a relatively young and still small field, BER practitioners now find themselves fragmented across 64 biology-related societies and lacking agreement on a core research agenda, a convenient professional network, and venues of dissemination. To begin addressing these needs, the newly formed Society for the Advancement of Biology Education Research (SABER) held an inaugural meeting in September 2010 at the University of Minnesota–Twin Cities. The 29 participants (Figure 1) included faculty, postdoctoral fellows, and graduate students engaged in empirical research and professional development, as well as journal editors, textbook writers, and a textbook editor. The diversity of this group contributed to a thoughtful, reflective, and productive meeting, whose major goals were to 1) define BER, 2) identify challenges to its practice, 3) formulate overarching research questions, and 4) outline the role of SABER in supporting the BER community. The consensus views of the participants on each of these goals are described in the paragraphs below.

The outcomes of this meeting are timely in the context of the recently released summary report *Vision and Change in Undergraduate Education, A Call to Action*, which charged the biology community with "creating, using, assessing, and disseminating effective practices in teaching and learning" (American Association for the Advancement of Science, 2010). By providing the infrastructure needed to create a vibrant network of practitioners, SABER is well positioned to support biology education researchers in generating empirical evidence that can effect meaningful changes in undergraduate biology education.

**WHAT IS BER?**

Twenty years ago, in his widely cited work, *Scholarship Reconsidered* (1990), Ernest Boyer argued that scholarly teaching should receive equal emphasis with disciplinary research at American universities. At about the same time, education researchers in science disciplines began to apply empirical research methods to the assessment of teaching and learning, particularly in physics (e.g., Hestenes et al., 1992). Some practitioners used assessment results to measure and enhance the effectiveness of their teaching (sometimes called action
research), and this tradition came to be known as the scholarship of teaching and learning (SoTL). Others used assessment tools to push beyond immediate practical application in their own classrooms and link research on how students learn a specific discipline to results from education and cognitive sciences on how people learn in general (reviewed in National Research Council, 1999), giving rise to the field of discipline-based education research (DBER).

The relation of SoTL to DBER has been the subject of considerable debate (e.g., Kreber, 2002; Boshier, 2009). SABER participants agreed that BER is separate from, though not exclusive of, SoTL or scientific teaching (Handelsman et al., 2007). The group adopted the definition that BER is hypothesis-driven research seeking to create new knowledge about the teaching and learning of biology and to disseminate that knowledge to the broader scientific community.

CHALLENGES TO THE PRACTICE OF BER

Compared to other areas of DBER, such as chemistry or physics, BER is an adolescent field, and its practitioners face a number of challenges both at the career and day-to-day levels. For example, many of those currently engaged in BER have a PhD in a life sciences discipline but have developed expertise in BER through informal routes and continue to publish biology research in addition to BER. Further, although there are increasing numbers of tenure-track BER faculty members in biology departments, the evaluation criteria for promotion and tenure decisions are often less well defined than for traditional biology faculty. This situation stems, in part, from the fact that the research methods in BER are often distinct from those of other biology colleagues, and there is seldom more than one tenure-track BER faculty member in any given biology department. Moreover, the multiple entry points into the field are not well articulated, and pathways for training and preparing BER scholars have yet to be established. Participants in the SABER meeting anticipated that these career challenges will lessen as BER matures into an accepted sub-discipline of biology.

Biology education researchers also encounter the daily challenges of keeping abreast of the most current research design and analytical techniques in a field that requires a particularly broad familiarity with the literature in both education and biology. BER is disseminated in a wide range of journals, from discipline-specific venues such as *Bioscience* and *Genetics* to educational journals like *Journal of Research in Science Teaching*. Furthermore, many educational journals in which BER results might be disseminated, such as *CBE—Life Sciences Education*, have yet to acquire the conventional impact factors that would allow traditional biologists to evaluate the impact of BER work.

Perhaps the greatest challenge in the practice of BER is the isolation in which many BER practitioners work. Unlike Physics Education Research or Chemistry Education Research, BER does not have a unifying society, entity, or venue to enhance collaboration among practitioners and support the growth of BER. Not surprisingly, SABER participants were invigorated by the opportunity to interact with other biology education researchers whom they had not met at other meetings. In fact, this was a primary benefit noted by the SABER founding members in a meeting follow-up survey. When asked about their sense of connection with colleagues in BER as a result of the meeting, all 21 respondents said it had either increased (16%) or significantly increased (84%).

OVERARCHING RESEARCH QUESTIONS IN BER

The progressive research agenda envisaged by SABER will focus on the systematic investigation of questions that are unique to teaching and learning biology, while drawing on foundational and ongoing research from the cognitive

---

**Figure 1.** SABER Founding Members: Teri Balser, Clarissa Dirks, Mary Pat Wenderoth, Janet Branchaw, Rob Brooker, Peggy Brickman, Malcolm Campbell, Mark Connelly, Erin Dolan, Scott Freeman, Mark Hens, Jenny Knight, Kathryn Miller, Jennifer Monsen, Lisa Montplaisir, Erika Offerdahl, Marcy Osgood, Nancy Pelaez, Becky Ruden, Jonathon Schramm, Michele Shuster, Karen Sirum, Amber Smith, Michelle Smith, Brian White, Devin Wixon, William Wood, Robin Wright. (Photo Credit: Becky Ruden, founding member)
sciences and other areas of DBER. Meeting participants reached consensus on four overarching lines of inquiry:

1. How does learning in biology compare to learning in the other STEM disciplines? What aspects of learning biology are unique to the discipline?
2. How does scientific teaching impact 1) students’ long-term conceptual development, 2) ability to think like a scientist, and 3) career path choices in biology?
3. What practices, activities, and assessments promote the acquisition and transfer of those science process skills (or competencies) that make a biologist?
4. What are the most effective pathways for institutionalizing evidence-based teaching?

ROLE OF SABER

Participants envisioned the new society as supporting a research community dedicated to improving the teaching and learning of biology, particularly at the undergraduate level. One outcome of the meeting was to articulate the role of SABER in the following mission statement:

SABER is a scientific community whose members develop theory and generate evidence with the goal of improving biology education. SABER fosters Biology Education Research (BER) and its dissemination by defining the standards for BER practice, supporting the BER community through training and faculty development programs, and fostering collaborations among BER investigators.

To grow the membership of SABER, the inaugural participants agreed to invite an initial group of about 100 biologists and educators with interests in BER to join the society as charter members. Additional members will be recruited through the SABER website (http://saber-biolyeducationresearch.wikispaces.com), which is now accessible although still under development, as described further below.

One priority of SABER will be to sponsor an annual national meeting that brings together biology education researchers to share their research findings. The first such conference will take place in the summer of 2011 at the University of Minnesota–Twin Cities, from Friday July 29 through Sunday July 31. SABER invites all who are interested in BER, including undergraduate and graduate students, postdoctoral researchers, and administrators, to attend. Chairs and deans will have the opportunity to take part in a preconference administrators’ lunch and tour the spectacular new technology-equipped collaborative-learning biology classrooms in the Science Teaching and Student Services building on Friday afternoon of the meeting. The meeting will include plenary sessions, research talks, poster presentations, networking opportunities, and professional development workshops as well as the society’s business meeting and a working session of the advisory board.

A call for presentation proposals will be made in January 2011, with abstracts due March 15. Authors of abstracts selected for presentation will be notified in April 2011. Details will be available on the SABER website.

The SABER concept that germinated in Minneapolis this fall must now be transformed into a vibrant and functional society. SABER will provide support and a community of practice for biology education researchers across all disciplines of biology. The society is intended to bring together BER practitioners from other biology societies as well as those from traditional education societies and societies that focus on the SoTL. In coming months, the SABER website will expand to include links to relevant meeting announcements, graduate programs, job postings, and funding opportunities. In addition, SABER plans to offer several members-only services, including access to an expanding and annotated directory of professionals (for students seeking mentors and committee members, and for faculty seeking collaborators or tenure package reviewers), a current annotated bibliography of BER-related literature, and professional development opportunities. SABER will cultivate high research standards, thereby positioning it to collaborate with other professional societies to support faculty and graduate students who are either active in BER or seeking to transition to BER, as well as journals seeking to expand or redefine submission guidelines to include BER-related articles. SABER will periodically survey its members to ensure a dynamic society, in tune with the needs of its membership.

The inaugural meeting of SABER harnessed the energy and enthusiasm of an emerging field and created the inspiration needed to guide its future growth. As the BER community coalesces, we envision SABER not only as an advocate for the BER community but also as a direct response to national calls for the transformation of undergraduate biology education. The meeting participants hope that SABER will provide a framework to unite biology education researchers and shape BER in the 21st century.

ACKNOWLEDGMENTS

Special thanks to Robin Wright for hosting the event at the University of Minnesota–Twin Cities. This work was funded by NSF Research Coordination Network for Undergraduate Biology Education incubator grant 0955572.

REFERENCES


42 CBE—Life Sciences Education
Today’s doctoral programs continue to prepare students for a traditional academic career path despite the inadequate supply of research-focused faculty positions. We advocate for a broader doctoral curriculum that prepares trainees for a wide range of science-related career paths. In support of this argument, we describe data from our survey of doctoral students in the basic biomedical sciences at University of California, San Francisco (UCSF). Midway through graduate training, UCSF students are already considering a broad range of career options, with one-third intending to pursue a non–research career path. To better support this branching career pipeline, we recommend that national standards for training and mentoring include emphasis on career planning and professional skills development to ensure the success of PhD-level scientists as they contribute to a broadly defined global scientific enterprise.

INTRODUCTION

Forty years ago, the career trajectory of PhD-level basic biomedical scientists could be described as a linear pipeline. Trainees moved from doctoral to postdoctoral training, and, ultimately, to tenure-track faculty positions. As the number of trainees has outpaced the availability of academic positions, an increasing number of PhD-trained scientists have pursued paths outside of academia. These scientists are often described as “leaking” from the pipeline. Unfortunately, this metaphor perpetuates the negative perception that scientists who “leak” are outside the norm and represent failures within the system. In fact, today’s PhD students and postdoctoral scholars commonly follow diverse career paths. Not only are PhD-trained scientists pursuing research careers beyond academe, but increasing numbers are leaving research altogether.

This is not new information—several national reports (National Research Council [NRC], 1998, 2005, 2011) and prominent articles (Golde and Dore, 2001; Teitelbaum, 2008; Benderly, 2010) have highlighted shifting career patterns of life scientists and challenges faced by those who do pursue the academic path. Consequently, promising new initiatives have been established in the area of trainee career development. However, most of these initiatives have focused on assisting young biomedical investigators as they transition to independent academic positions. A gaping hole remains; as a scientific community we have ignored the many trainees who will pursue nontraditional positions.

Our lack of action in this arena is shocking, because “nontraditional” career paths are not “alternative.” Since 2001, fewer than 20% of PhDs in the biological sciences have been moving into tenure-track academic positions within 5–6 yr of receiving a PhD. In fact, the most recent data (2006) show only 14% of these PhDs in tenure-track positions. Forty-three percent were employed full-time in nonacademic settings.
(Stephan, 2012). With increasing numbers of domestic and international PhDs being trained (Cyranoski et al., 2011), and research funding to employ and support them becoming tighter, the proportion of PhDs pursuing nontraditional career paths is likely to continue to increase. Students in these career paths need a broad set of skills to succeed in such positions (Smith et al., 2002; Melin and Janson, 2006; Rudd et al., 2008). Yet, graduate program curricula and federal training standards changed little in response.

To address this issue, those of us leading graduate education need to increase our awareness of when, and why, career choices are made. Career outcomes data for doctoral alumni are informative, but to design and implement doctoral-level career development curricula most effectively, we need insight into the career planning of current trainees, including the career paths they are considering, the factors influencing their career decisions, and the timing of when their career decisions are made. Both universities and funding agencies need this information to ensure the training being offered appropriately prepares these talented scientists for their future careers.

Two prior studies have looked at doctoral student career preferences and how these career preferences change over time (Golde and Dore, 2001; Goulden et al., 2009; Mason et al., 2009). In a 1999 national survey of doctoral students in 11 arts and sciences fields, Golde and Dore found most students entered graduate school strongly considering a faculty career, but students reported a change in interest for this career path during their training: 35% of students reported becoming less interested in this career path and 21% reported becoming more interested (Golde and Dore, 2001). In 2006, Mason and colleagues surveyed graduate students in all disciplines across several University of California campuses, and learned that while 45% of men and 39% of women initially planned to become a professor with research emphasis, this proportion had decreased to 36% and 27%, respectively, by the time of the survey (1–7 or more yr after starting their training; Goulden et al., 2009; Mason et al., 2009). Data from these and other studies (Fox and Stephan, 2001) showed that career preferences vary significantly across disciplines. Therefore, to understand and address career development needs in the basic biomedical sciences, we need to look deeply and specifically at the career preferences for students within this particular discipline.

In this paper, we show that large numbers of students in the basic biomedical sciences are considering career paths beyond academe—and even beyond research. This change in career preference occurs early in graduate school. We use the metaphor branching career pipeline to describe this substantial flow of PhD-trained scientists into various sectors of the workforce, and propose that this branching should be seen as a valuable opportunity for spreading science throughout society. To better support today’s branching science careers pipeline, we recommend that national standards for training and mentoring place more emphasis on career planning and professional skills development to ensure the success of PhD-level scientists as they contribute, in a variety of science-related career paths, to a broadly defined global scientific enterprise.

STUDENTS ARE CONSIDERING A RANGE OF CAREER OPTIONS

We surveyed all basic biomedical sciences doctoral students at University of California, San Francisco (UCSF) to determine what career paths they are strongly considering, whether these preferences are different from when they started their training, and, if so, why. UCSF is the only campus in the University of California system that focuses solely on graduate-level training. The PhD programs in the basic biomedical sciences are ranked among the top in the nation, and admission is highly competitive. Therefore, one would anticipate that these students have a high likelihood of success in research-focused career paths, and a bias toward choosing research-focused careers.

The survey, distributed in spring 2008, is available in Supplemental Material 1. Four hundred sixty-nine students responded, corresponding to 62.3% of all basic biomedical science graduate students at UCSF (Table 1).

Respondents initially identified all categories of careers they were strongly considering. As expected, the vast majority of students (92.3%, n = 432) were strongly considering careers in scientific research (i.e., in academia, industry, government; Figure 1). Seventy-two percent (n = 338) of students included a traditional academic career path (i.e., as faculty with a significant portion of their time spent on research) among the career paths they were considering. When asked to choose only a single career (Figure S1 in Supplemental Material 2), only 44.8% (n = 210) of student respondents selected a traditional academic career path; 26.9% (n = 126) selected another scientific research career path, such as a research-intensive career in biotechnology/ pharmace uticals, a research career in government, or a non–principal investigator (non-PI) research career in academia (collectively, these three categories are hereafter called “other research careers”). While these numbers represent a general preference for research careers, they leave over one-fourth of students choosing non–research careers.

Indeed, 71.2% (n = 333) of all respondents were “strongly considering” at least one career path not directly involving scientific research (hereafter called “non–research” careers; Figure 1), and 27.9% (n = 131) selected one of these non–research careers as their top preference (Figure S1 in Supplemental Material 2). Among these career paths, the most popular choices were business of science, teaching- or education-related, science policy, and writing career paths. Given that graduate training is focused almost entirely on research, it is surprising that so many students selected a non–research career path as their top choice.

The fact that students were considering such a broad range of career options suggests they also have low confidence in their current career choice. Indeed, only 15.8% (n = 73) of students self-identified as “very confident” in their
Table 1. Survey demographics and response rate

<table>
<thead>
<tr>
<th>Students</th>
<th>Enrolled at UCSF</th>
<th>Responded to survey</th>
<th>% Responded</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>138</td>
<td>84</td>
<td>60.9</td>
</tr>
<tr>
<td>2</td>
<td>121</td>
<td>80</td>
<td>66.1</td>
</tr>
<tr>
<td>3</td>
<td>142</td>
<td>88</td>
<td>62.0</td>
</tr>
<tr>
<td>4</td>
<td>105</td>
<td>69</td>
<td>65.7</td>
</tr>
<tr>
<td>5</td>
<td>114</td>
<td>58</td>
<td>50.9</td>
</tr>
<tr>
<td>6</td>
<td>78</td>
<td>49</td>
<td>62.8</td>
</tr>
<tr>
<td>7 or higher</td>
<td>55</td>
<td>21</td>
<td>38.2</td>
</tr>
<tr>
<td>Unreported</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Graduate program</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bioengineeringa</td>
<td>176a</td>
<td>67</td>
<td>38.1a</td>
</tr>
<tr>
<td>Biomedical Sciences</td>
<td>135</td>
<td>88</td>
<td>65.2</td>
</tr>
<tr>
<td>Biophysics</td>
<td>63</td>
<td>39</td>
<td>61.9</td>
</tr>
<tr>
<td>Biological and Medical Informatics</td>
<td>32</td>
<td>21</td>
<td>65.0</td>
</tr>
<tr>
<td>Chemistry and Chemical Biology</td>
<td>45</td>
<td>28</td>
<td>62.2</td>
</tr>
<tr>
<td>Neuroscience</td>
<td>91</td>
<td>64</td>
<td>70.3</td>
</tr>
<tr>
<td>Pharmaceutical Sciences and Pharmacogenomics</td>
<td>51</td>
<td>33</td>
<td>64.7</td>
</tr>
<tr>
<td>Tetrad</td>
<td>160</td>
<td>122</td>
<td>76.3</td>
</tr>
<tr>
<td>Cell Biology</td>
<td>—</td>
<td>39</td>
<td>—</td>
</tr>
<tr>
<td>Developmental Biology</td>
<td>—</td>
<td>7</td>
<td>—</td>
</tr>
<tr>
<td>Genetics</td>
<td>—</td>
<td>16</td>
<td>—</td>
</tr>
<tr>
<td>Biochemistry &amp; Molecular Biology</td>
<td>—</td>
<td>50</td>
<td>—</td>
</tr>
<tr>
<td>Tetrad—focus not yet determined</td>
<td>—</td>
<td>10</td>
<td>—</td>
</tr>
<tr>
<td>Unreported</td>
<td></td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>368</td>
<td>249</td>
<td>67.7</td>
</tr>
<tr>
<td>Male</td>
<td>385</td>
<td>205</td>
<td>53.2</td>
</tr>
<tr>
<td>Unreported</td>
<td></td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>Ethnicityb</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>American Indian or Alaskan Native</td>
<td>—</td>
<td>1</td>
<td>—</td>
</tr>
<tr>
<td>Asian</td>
<td>—</td>
<td>102</td>
<td>—</td>
</tr>
<tr>
<td>Black or African American</td>
<td>—</td>
<td>10</td>
<td>—</td>
</tr>
<tr>
<td>Hispanic or Latino</td>
<td>—</td>
<td>26</td>
<td>—</td>
</tr>
<tr>
<td>Pacific Islander</td>
<td>—</td>
<td>8</td>
<td>—</td>
</tr>
<tr>
<td>White</td>
<td>—</td>
<td>298</td>
<td>—</td>
</tr>
<tr>
<td>Other</td>
<td>—</td>
<td>9</td>
<td>—</td>
</tr>
<tr>
<td>Unreported</td>
<td></td>
<td>37</td>
<td></td>
</tr>
<tr>
<td>OVERALL</td>
<td>753</td>
<td>469</td>
<td>62.3</td>
</tr>
</tbody>
</table>

a All graduate programs returned a response rate of more than 60%, with the exception of Bioengineering, a joint graduate program with UC Berkeley. The response rate for Bioengineering students (38.1%) is likely lower because 62% of Bioengineering graduate students identify UC Berkeley as their home campus and many of these students may have ignored the survey request.

b Respondents could choose more than one category to describe their ethnicity. Response rates were not calculated because ethnic categories were defined differently in our study compared with those on file through university enrollments.

current career choice; 29.3% (n = 135) of students were “fairly confident”; and the majority of students (54.9%, n = 253) were “still considering a range of options.” There was no significant difference in confidence between students choosing research careers and students choosing non-research careers (p = 0.37). Although students choosing to become a PI at a research-intensive institution were more confident than students not considering that career path (p < 0.001), they were still not very confident, with only 23.7% (n = 31) “very confident” in that selection, 38.9% (n = 51) “fairly confident,” and 37.4% (n = 49) “still considering a range of options.”

CAREER PREFERENCES SHIFT MIDWAY THROUGH GRADUATE STUDIES

How do career choices and confidence levels change during the graduate school experience? Although longitudinal data would be ideal, our data provide a cross-sectional view. When asked to choose a single career path, confidence in the chosen career path depended on the stage of graduate training (p = 0.006). A large change in confidence occurred between the first and second year in graduate school, with the number of students “still considering a range of options” increasing from 48.8% (n = 40) to 66.7% (n = 52; see Figure S2 in Supplemental Material 2). Uncertainty in career choice remains high...
(61.4%, 61.8%, and 55.2% for third, fourth, and fifth years, respectively) until students approach the expected time of graduation (sixth or later years, 33.3%, respectively) until students approach the expected time of graduation (sixth or later years, 33.3%, respectively) until students approach the expected time of graduation (sixth or later years, 33.3%, respectively). Overall, 92% of students were strongly considering at least one category of research careers (represented by bars in dark blue), with 72% of all students strongly considering becoming a PI at a research-intensive academic institution and/or a PI with a balance of teaching and research. Seventy-one percent of all students were strongly considering at least one category of careers that typically do not directly involve performance of research (labeled as “non-research” careers and represented by bars in light blue). (B) As illustrated in this Venn diagram, many students (63%, n = 297) were strongly considering both research and non–research career paths.

Figure 1. Graduate students are strongly considering a range of career options. (A) The bar graph shows the percentage of all student respondents who chose each category as one of the career path categories they were strongly considering (respondents could choose more than one category). Overall, 92% of students were strongly considering at least one category of research careers (represented by bars in dark blue), with 72% of all students strongly considering becoming a PI at a research-intensive academic institution and/or a PI with a balance of teaching and research. Seventy-one percent of all students were strongly considering at least one category of careers that typically do not directly involve performance of research (labeled as “non-research” careers and represented by bars in light blue). (B) As illustrated in this Venn diagram, many students (63%, n = 297) were strongly considering both research and non–research career paths.

When interest in the research-intensive PI track decreases, one might expect a corresponding increase in interest for other research careers. However, student interest in other research career categories remained relatively stable (Figure 2B and Table 2). Instead, interest in non–research career paths increased (from 20 to 34% between the second- and third-year classes). Within that broad category, the percent of students choosing a career path in teaching or science education did not change significantly (Table 2). However, interest in non–research career paths outside of academia (the business of science, science writing, healthcare, science policy, law-related, and drug-approval and production) did increase, from 15.0 to 29.5% between the second- and third-year class (p = 0.006 when comparing early- and late-stage students). In summary, the shift in students’ career choices can be described most succinctly as a decreased interest for becoming a PI at a research-intensive university, and an increase of interest in non–academic, non–research career paths. These data are consistent with those described for PhD students across all disciplines in the University of California system (Mason et al., 2009).

In addition to a decreased interest in the traditional PI track as a top career choice, students also tended to eliminate the PI track from their inclusive list of career paths being “strongly considered.” While 89.3% (n = 75) of first-year students included the PI track (research-intensive or with a balance of teaching and research) as a career path they were strongly considering, this percentage decreased sharply between the second- (87.5%, n = 70) and third-year (67.8%, n = 59) classes. Within this broader PI career category, there was a more pronounced drop in consideration of PI positions at research-intensive institutions (from 75.0% in the second-year class to 55.2% in the third-year class) compared with the drop in consideration of PI positions with a more even balance of teaching and research (from 56.3 to 44.8%). Is this change in career preferences gender-specific? Interestingly, our analysis showed no significant difference in the percent of men (27.8%) or women (28.9%; p = 0.79) who selected a non–research career path as their first choice. There
Support a Branching Career Pipeline

Figure 2. Early in graduate school, some students lose interest in becoming a PI at a research-intensive academic institution. (A) The percentage of students in each cohort who currently would choose a research (dark blue circles) or non-research (light blue triangles) career. Between the second and third year, there is a steep drop in interest in research careers. (B) Within the broad category “research careers,” the only career choice that showed significant change was that of being a PI at a research-intensive academic institution (blue solid diamonds). Values and statistical analyses are given in Table 2.

Table 2. Current career choice as a function of year in graduate school

<table>
<thead>
<tr>
<th>Current career choice</th>
<th>Year 1</th>
<th>Year 2</th>
<th>Year 3</th>
<th>Year 4</th>
<th>Year 5 and higher</th>
<th>Year 1 and 2 (n=164)</th>
<th>Year 3+ (n=285)</th>
<th>p Valuea</th>
<th>%</th>
<th>n</th>
<th>%</th>
<th>n</th>
<th>%</th>
<th>n</th>
<th>%</th>
<th>n</th>
<th>p Valuea</th>
</tr>
</thead>
<tbody>
<tr>
<td>Research career</td>
<td>81.0</td>
<td>80.0</td>
<td>65.9</td>
<td>66.7</td>
<td>67.2</td>
<td>65.7</td>
<td>0.067</td>
<td>80.5</td>
<td>132</td>
<td>66.3</td>
<td>189</td>
<td>0.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non–research career</td>
<td>17.9</td>
<td>20.0</td>
<td>34.1</td>
<td>33.3</td>
<td>31.0</td>
<td>34.3</td>
<td>0.051</td>
<td>18.9</td>
<td>31</td>
<td>33.3</td>
<td>95</td>
<td>0.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PI at a research-intensive institution</td>
<td>41.7</td>
<td>35.0</td>
<td>25.0</td>
<td>23.2</td>
<td>22.4</td>
<td>25.7</td>
<td>0.047</td>
<td>38.4</td>
<td>63</td>
<td>24.2</td>
<td>69</td>
<td>0.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PI doing teaching and research</td>
<td>13.1</td>
<td>18.8</td>
<td>17.0</td>
<td>15.9</td>
<td>17.2</td>
<td>12.9</td>
<td>0.897</td>
<td>15.9</td>
<td>26</td>
<td>15.8</td>
<td>45</td>
<td>0.986</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other research careers</td>
<td>26.2</td>
<td>26.3</td>
<td>23.9</td>
<td>27.5</td>
<td>27.6</td>
<td>27.1</td>
<td>0.995</td>
<td>26.2</td>
<td>43</td>
<td>26.3</td>
<td>75</td>
<td>0.98</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teaching and education</td>
<td>2.4</td>
<td>5.0</td>
<td>4.5</td>
<td>8.7</td>
<td>8.6</td>
<td>7.1</td>
<td>0.51</td>
<td>3.7</td>
<td>6</td>
<td>7.0</td>
<td>20</td>
<td>0.14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other non–research careers</td>
<td>14.3</td>
<td>15.0</td>
<td>29.5</td>
<td>23.2</td>
<td>22.4</td>
<td>25.7</td>
<td>0.12</td>
<td>14.6</td>
<td>24</td>
<td>25.6</td>
<td>73</td>
<td>0.006</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

There is a trend toward significance for change of career choice from year to year (p = 0.067 and 0.051 for change in interest of research career paths and non-research career paths, respectively). Most of this change in career choice occurs between the second and third years of graduate school. This is evident when data are grouped for early (years 1 and 2) and later-stage (year 3 and above) students.

“Other non–research careers” includes careers in the business of science, science writing, healthcare, science policy, law, and drug approval and production. Some columns may not add to 100% because respondents who indicated “other science-related careers” were not included in the bottom half of the table.
was a significant difference, however, in interest for becoming a PI at a research-intensive institution (21.3% of women would choose this path, compared with 39.5% of men, \( p < 0.001 \)). The latter finding is consistent with prior studies of UC students across disciplines (Goulden et al., 2009; Mason et al., 2009). However, the decline over time in interest for this career path was not gender-specific; similar to both the Mason and Goulden studies, both genders in our study lost interest in this career path over time.

Do postdoctoral scholars show similar trends in career preference? We simultaneously surveyed postdoctoral scholars at UCSF (see Supplemental Material 3), and found postdoctoral scholars were more likely to prefer a research-focused career path (89%). Postdocs were generally more confident in their career choice than students, but still had low confidence (with 37% reporting they were “still considering a range of options”). Although 20% indicated their career preference had changed since beginning their postdoctoral training, this change was not apparent in our analysis of career choice data: There was no significant change in aggregate interest for each career category when tracked across years of postdoctoral training.

WHY DO STUDENTS MOVE AWAY FROM THE ACADEMIC PI TRACK?

What factors caused students to so radically change their career aspirations? Others have shown across disciplines that a variety of factors come into play, such as workload expectations, difficulty in getting research funding, competition within academia, low availability of jobs, loss of interest in basic research, and increased interest in other careers (Rice et al., 2000; Golde and Dore, 2001; Austin, 2002; Bakken et al., 2006; Goulden et al., 2009; Mason et al., 2009; Roach and Sauermann, 2010). In our study, 31.0% \( (n = 143) \) of students stated their career choice had changed since they began their graduate training. Seventy-nine students provided reasons for moving away from the academic PI track, most describing multiple reasons (hence, the percentages that follow do not add to 100%; see Supplemental Material 4 for examples). Of these, 91% \( (n = 72) \) described negative perceptions related to this career path. Thirty percent \( (n = 24) \) referenced inadequate quality-of-life or work-life balance, 22% \( (n = 17) \) referenced the competition or stress associated with trying to succeed within an academic position, and 24% \( (n = 19) \) anticipated difficulty in getting research funding. Nineteen percent \( (n = 15) \) mentioned the length of training required or competition to get academic jobs, and 11% \( (n = 9) \) mentioned low salary during training or in the job. Twenty-five percent \( (n = 20) \) wrote they disliked the tasks required of being an academic PI, such as grant writing and project management, and the slow pace of research. In contrast, only 24% \( (n = 19) \) of these students provided a positive reason for change, such as learning more about other career options or discovering a new skill or interest in the course of their graduate education.

These data suggest role modeling by faculty, and even postdoctoral scholars, had great impact on graduate students’ perceptions of academic careers, for better or worse. Positive and negative effects of role modeling have also been described for doctoral students in the physical sciences (Paglis et al., 2006) and across all disciplines (Austin, 2002; see dis-

Figure 3. Graduate student career preferences predict a branching career pipeline. This diagram illustrates the branching pipeline model, describing the career trajectory of PhD-level scientists. The central pipes represent graduate (light orange) and postdoctoral (darker orange) training. Black arrows represent the desired career paths of students in their third or later year of graduate school. According to our survey, the branched nature of this pipeline can be predicted as early as the third year in graduate school, with 40% of these later-stage students intending to become a principal investigator in academia, 26% intending to pursue other research-focused career paths, and 33% intending to pursue non–research career paths. Many of these students move on to postdoctoral training—including some students who prefer to pursue a non–research career path.

SUMMARY OF THE DATA

In conclusion, our data showed that UCSF PhD students were considering many different career paths, with most simultaneously considering both non–research and research career paths. Career path choices shifted during the first 3 yr of graduate school. This was primarily driven by decreased interest in becoming a PI at a research-intensive university. By the later years of graduate school, fully one-third of students stated they would choose a non–research career path (Figure 3and Table 2).

ARE THESE LOCAL EFFECTS OR A NATIONAL TREND?

Are these data unique to UCSF? Some characteristics of the university itself may have influenced the data. First, UCSF’s location in the San Francisco Bay Area, a geographic region rich in biotech companies, may attract prospective students who are already interested in diverse career paths. A second influencing factor might be the variety of resources historically available to assist UCSF students as they explore careers beyond academic research. These resources include seminars and individual consultations provided by the Office of Career and Professional Development, the Center for Bioentrepreneurship’s “Idea to IPO” (a course created through
a collaboration of UCSF faculty and leaders in nearby industry, and numerous alumni events hosted by student organizations. It is possible UCSF students experience a broader awareness of their career path options because these resources exist. However, the data suggest these influences play a minor role: only 24% of respondents described “positive” reasons for moving away from research-intensive faculty careers (for example, increased knowledge about a career path), compared with 91% of respondents who described negative aspects of the research-intensive academic career path.

The question remains, then: does the UCSF environment in some way cause students to move away from research-intensive academic careers? UCSF, an academic medical center, is itself a research-intensive graduate-level institution. Faculty members focus on research and have minimal teaching responsibilities. It is common for faculty salaries to be supported entirely by grant money, with limited “hard money” support. UCSF therefore naturally attracts and promotes faculty driven by a singular passion: research. Students observe their advisors and wonder, “Would I also enjoy this career path?”

Furthermore, all students have high expectations of success, as evidenced by high Graduate Record Examination (GRE) scores and prevalence of external funding (PhDs.org, 2011; based on NRC, 2010). As high achievers, these students were probably coached to pursue academic careers by their undergraduate advisors. They may have entered graduate school with especially high expectations of pursuing and succeeding along an academic career path. As a result, our data may exhibit a sharper drop in interest for academic careers among UCSF students compared with students at other institutions. An in-depth, cross-institutional study should be pursued to test how (or if) undergraduate academic success impacts students’ career preferences and their confidence in career choice, and how their career preferences change with time.

While UCSF may be among the nation’s most research-focused institutions, it is hardly unique. Many basic biomedical sciences doctoral programs are similarly housed in research-intensive academic medical centers. Many of these programs also attract highly talented students. We predict a national survey of similar institutions would reveal that graduate students’ career decisions follow a similar trend, with a drop in interest for research-intensive academic positions following a year or less of full-time experience at the bench. Preliminary discussions of the data with colleagues at other institutions support this prediction, but a formal cross-institutional study should be done.

DATA HIGHLIGHT DISCIPLINE-SPECIFIC INFLUENCES ON CAREER CHOICE

Most studies of doctoral student career preferences report data averaged across all disciplines (Rice et al., 2000; Golde and Dore, 2001; Goulden et al., 2009; Mason et al., 2009; among others). However, disciplinary culture—and institutional culture—can affect how and why students choose certain career paths. Here we compare our qualitative data with the data of others to highlight some important differences.

As described above (see “Why Do Students Move away from the Academic PI Track?”), our respondents cited several reasons for no longer considering the traditional academic track. Two of the major themes—anticipated competition/stress and insufficient work-life balance—were also two of the greatest concerns cited by doctoral students in cross-disciplinary studies (including the humanities and social sciences; Rice et al., 2000; Mason et al., 2009). Stress and lack of work–life balance are legitimate concerns: cross-disciplinary studies have shown that new faculty frequently use terms such as “stress, pressure, and uncertainty” to describe their role (Rice et al., 2000; Austin, 2002). What is more, longitudinal studies show that faculty stress intensifies over the first 5 yr (Olsen and Sorcinelli, 1992).

Although stress and lack of work–life balance seem to be themes independent of discipline, our study highlights distinct disciplinary differences in the underlying causes for these concerns. While doctoral students in broader studies noted that faculty lack time for research because of significant teaching responsibilities (Rice et al., 2000), this was not a theme echoed by our respondents. Instead, UCSF students’ concerns about future stress and work–life balance were often described in the context of the scientific research itself. In particular, difficulty in getting research funding was specifically described by 24% of respondents in our study (n = 19). A lack of concern about teaching is not surprising; UCSF faculty have few teaching responsibilities. Instead, faculty promotions—and even salary—are often contingent on research productivity and funding.

Another example of how cultural differences impact career perceptions stems from the level of collaboration within department, institution, and research field. The cross-disciplinary study “Heeding New Voices” identified “isolation”—an insufficient sense of community and collaboration—as a primary concern for aspiring and junior faculty (Rice et al., 2000). As expected, this was not a concern expressed by students in our study. Basic biomedical research is increasingly interdisciplinary and collaborative, a culture of sharing ideas and projects within and across lab groups, and even across institutions. Others have noted this cultural difference as well (Gardner [2007] and Golde [1998], as referenced in Gardner [2010]).

These analyses emphasize that there are distinct differences in the cultures and nature of work—and therefore in career decision making—that depend on discipline and institutional context. To best understand the needs of basic biomedical science trainees, many of whom are trained in institutions similar to UCSF, we need to pursue cross-institutional studies specific to students within this field and within the research-intensive context of health science campuses.

HOW SHOULD WE REACT TO THESE DATA?

Some in our scientific community argue that the purpose of graduate education is to train future academic research faculty. If this is indeed the purpose of graduate education, then the data presented in this report are troubling. The question ultimately arises: should we be more selective in graduate school admissions in the first place, admitting only students who are both highly likely to succeed in these careers and also likely to pursue them? Our study suggests such an
approach would be fruitless. Despite an already highly rigid structural selection process and prevalent initial interest in academic careers, 60% of UCSF students in their third year or later would choose a career path beyond the traditional academic path. As discussed above, this trend is likely echoed at other research-intensive graduate institutions.

Instead of turning talented students away from doctoral education based on their career preferences, we believe that we, as educators, should embrace the branching career pipeline and shift the current paradigm for graduate education toward a more inclusive curriculum capable of preparing doctoral students for a variety of scientific careers. Indeed, to maintain our science pipeline, we must ensure graduate-level training and career prospects after training are perceived as valuable and rewarding. To quote Bruce Alberts, former President of the National Academy of Sciences and current editor-in-chief of Science, “The entire enterprise will be jeopardized if students generally feel dissatisfied with their training” (personal communication). In combination, our data and data from prior studies (Golde and Dore, 2001; Aanerud et al., 2006; Mason et al., 2009) support three recommendations for how we as scientists, educators, and policy makers can strengthen graduate training, improve student wellness and satisfaction, and produce a more highly skilled national scientific workforce.

1. SHIFT ACADEMIC CULTURE TO EMBRACE THE “BRANCHING” SCIENCE CAREER PIPELINE

We believe the academic community should be supportive of individual PhD-level trainees interested in pursuing careers beyond the traditional academic path. Our data show that trainees do not make major career decisions lightly; respondents shared thoughtful reasons for shifting their career choices away from the traditional research-intensive PI track. The PI track, and the lifestyle, stressors, and lack of security currently associated with it, is not a fit for everyone. Moreover, there are not enough jobs in the academic sector for all PhD-trained life scientists, and this “supply–demand” gap is growing each year (Teitelbaum, 2008; Cyranoski et al., 2011; NRC, 2011). With only 14% of PhDs in the biological sciences entering tenure-track positions within 5–6 yr of earning their PhD (2006 data; Stephan, 2012), how can we continue to devalue other career paths? Finally, it is important for us to have PhD-trained scientists in roles that will benefit the scientific enterprise as a whole. They provide services critical to the advancement of science in today’s world, by developing and running research facilities, working with researchers to patent discoveries, bringing those discoveries to market, funding research, setting policies, and teaching future generations of scientists. As a scientific community and as individual mentors, we should be applauding PhD graduates who move on to become leaders in any science-related career path. When PhD-level scientists are distributed broadly throughout the workforce, we all benefit, because we are “creating the bridges needed for science to affect a wider society” (Alberts, 2008).

2. INTEGRATE CAREER DEVELOPMENT INTO THE GRADUATE CURRICULUM

Our national investment in graduate-level training will be optimized when trainees have a positive graduate experience, and then move on equipped to succeed in their future career paths.

Some institutions already offer career and professional development services tailored to graduate students and/or postdoctoral scholars in the basic biomedical sciences; examples are given in Table 3. However, such offerings are not the norm. Even institutions that do offer such services frequently emphasize preparation for academic careers, offering little support to students planning to pursue other types of career paths. We need to supplement graduate education with career development initiatives—including career planning and professional skills development—that will prepare our doctoral trainees for careers both within academia and beyond.

The branching nature of today’s biomedical sciences career path—and trainees’ low confidence in their career choices within the pipeline—underscores the need for structured career planning at the doctoral level, yet few science trainees are provided with career-planning assistance. A lack of career planning is likely one factor contributing to the high proportion of students who move on to postdoctoral training (80% of all biological sciences PhDs nationally [NRC, 2011]), even though this additional training is unnecessary for most students following a non–research career path.

Currently, career discussions between students and mentors often occur near the end of training, if at all. Our data emphasize that this is too late. Career education, guidance, and mentoring—tailored to the needs of students in the basic biomedical sciences and provided early in students’ graduate education—would help students make career decisions from a well-informed position. Students considering non–research career paths (or research career paths outside of academia) may greatly benefit from an opportunity to try out this new role through a short-term internship. This would help ensure career decisions are made based on realistic expectations.

Skills in areas such as interpersonal communication, presentation, leadership, management (Smith et al., 2002; Melin and Janson, 2006; Rudd et al., 2008), and networking are imperative for success in all careers. Teaching skills are also needed in many of the career choices. Yet, with our traditional emphasis on developing scientific knowledge and research skills in graduate education, few if any resources are dedicated to the broader professional development of graduate students and postdoctoral fellows. Graduate education should be supplemented with structured training and mentoring in these broader professional skills areas to prepare students for success in the broad range of traditional or nontraditional science-related careers. Students could each

---

4 As demonstrated in the Bridges to Independence report (NRC, 2005), the number of PhD-trained life scientists ages 35 and younger increased by 59% between 1993 and 2001 in the United States, but the number of these scientists in tenure-track positions increased by only 7%. The job market is perhaps tightest at research-focused universities: the number of tenure-track life scientists at Research I institutions decreased over this time by 12%. Clearly, there are not nearly enough tenure-track positions available for the number of life scientists being trained.

5 Many academic career development initiatives are based on the successful Preparing Future Faculty model (www.preparing-faculty.org; DeNeef, 2002).

6 Table 3 provides examples of centralized initiatives in this arena.
create an Individual Development Plan (IDP) and then discuss their own professional skills training with mentors to ensure such training is pursued in a time-efficient and productive manner (Lindstaedt, 2009; National Institute of General Medical Sciences [NIGMS], 2011).

Some will argue that encouraging students to explore career options and prepare for these careers will take away from the lab, and detract from research training. Graduate students and postdocs make up as much as 50% of the basic biomedical research workforce (NRC, 2011), and their tuition and stipends are increasingly funded by PIs' research grants (NIGMS, 2011). This creates an apparent conflict of interest for the PI: optimizing productivity of the lab as a whole, while being supportive of individual trainees within that lab as they pursue career-preparation activities (Benderly, 2010). One way to alleviate this conflict of interest is to give thesis committees, rather than individual PIs, the responsibility for overseeing student career development. It would be appropriate for thesis committees to participate in career-related mentoring, discuss with the student his/her IDP, and help the student and PI negotiate an appropriate level of time spent toward career-related activities.

Recent studies suggest career development activities do not negatively impact research training or productivity; in fact, the opposite may be true. Nationally, a recent report from NIGMS noted many investigators believe “training and laboratory productivity are synergistic” (NIGMS, 2011). Indeed, graduate students participating in the National Science Foundation’s GK–12 program (spending 15 h/wk developing teaching skills through training and in-classroom experience) reported that the experience improved their time-management skills and motivated them to complete their graduate degree. In fact, GK–12 fellows spent, on average, the same number of weekly hours on their doctoral research projects and ultimately completed their doctoral degree in the same amount of time as their non-GK–12 peers (Gamse et al., 2010). Other data take this a step further and suggest that career development activities can improve research productivity. Paglis et al. (2006) showed that doctoral students in the physical sciences who had received research mentoring submitted more abstracts, papers, and grants during their training. A national study funded by Sigma Xi found that postdocs who participated in career development-related activities reported better advisor relations, fewer conflicts, higher satisfaction, and, in some cases, more first-author papers and grants submitted (Davis, 2006). While further study is needed to determine how best to prepare students for a branching career pipeline—and how this training may or may not affect their productivity in the lab—the benefits and (lack of) risks already demonstrated by local and national initiatives strongly suggest career development should become a standard aspect of graduate training.

**3. TRANSFORM GRADUATE EDUCATION POLICY AT THE NATIONAL LEVEL**

Change in graduate education is often motivated by policies set at the national level. As such, it is important to consider how actions by national agencies might impact our view of the branching scientific pipeline and our ability to assist trainees in their career development.

Although the concept of the branching pipeline is becoming more broadly accepted at the institutional level and by individual faculty mentors, national funding agencies continue to use the traditional academic pathway as the formal definition of success. For example, in a ranking of graduate schools released in 2010 by the NRC, student career outcome was defined as the percent of PhDs “with definite plans for an academic position” (NRC, 2010). In addition, currently most, if not all, biomedical funding sources evaluate U.S. doctoral training programs based in part on the success of alumni, with many measures of success pointing to PI-level positions in academia. Funding agencies and review

---

**Table 3. Career development initiatives tailored to the needs of doctoral students and/or postdocs in the basic biomedical sciences**

<table>
<thead>
<tr>
<th>Institution/Program</th>
<th>Website</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fred Hutchinson Cancer Center, Office of Scientific Career Development</td>
<td><a href="http://www.fhcrc.org/science/education/grad_postdoc/oscd">www.fhcrc.org/science/education/grad_postdoc/oscd</a></td>
</tr>
<tr>
<td>Medical College of Wisconsin, Office of Postdoctoral Education (for graduate students also)</td>
<td><a href="http://www.mcw.edu/VirtualCareerCenter.htm">www.mcw.edu/VirtualCareerCenter.htm</a></td>
</tr>
<tr>
<td>National Institutes of Health, Office of Intramural Training and Education</td>
<td><a href="http://www.training.nih.gov/career_services">www.training.nih.gov/career_services</a></td>
</tr>
<tr>
<td>The Scripps Research Institute, Career and Postdoctoral Services Office</td>
<td><a href="http://www.scripps.edu/services/postdocs">www.scripps.edu/services/postdocs</a></td>
</tr>
<tr>
<td>Stanford University School of Medicine, Career Center</td>
<td><a href="http://med.stanford.edu/careercenter">http://med.stanford.edu/careercenter</a></td>
</tr>
<tr>
<td>UCSF, Office of Career and Professional Development</td>
<td><a href="http://career.ucsf.edu">http://career.ucsf.edu</a></td>
</tr>
<tr>
<td>UCSF, Graduate Student Internships for Career Exploration Program</td>
<td><a href="http://gsice.ucsf.edu">http://gsice.ucsf.edu</a></td>
</tr>
<tr>
<td>University of Pittsburgh School of Medicine, Office of Academic Career Development</td>
<td><a href="http://www.oacd.health.pitt.edu">www.oacd.health.pitt.edu</a></td>
</tr>
<tr>
<td>Vanderbilt University School of Medicine–Biomedical Research Education and Training, Office of Career Development and Outcomes Analysis</td>
<td><a href="http://bret.mc.vanderbilt.edu/career_development">http://bret.mc.vanderbilt.edu/career_development</a></td>
</tr>
</tbody>
</table>

*This list highlights offices and programs within research institutions whose primary purpose is to target the career development needs of graduate students and/or postdocs in the basic biomedical sciences. The list is intended to provide examples of the breadth of services that can be tailored to this population, and how they are implemented. This list is in no way exhaustive, or intended to be an endorsement of any particular program.*
committees should explicitly redefine the description of a “successful” PhD graduate as one whose contributions promote the scientific enterprise, including a variety of research and non–research career paths in both academic and nonacademic sectors. This would allow graduate schools to more freely support and encourage graduate students considering such career paths.

In addition to redefining a successful career outcome, funding agencies could urge institutions to incorporate career development components into all graduate programs. As discussed in recommendation #2 above, preparation of our future scientific leaders should include training beyond scientific knowledge and research skills. To promote this broader curriculum, funding agencies should define national expectations for mentoring, professional skills training, and career development for graduate and postdoctoral trainees, and provide funding to develop and implement these types of initiatives.

CONCLUDING REMARKS

Part of our responsibility as educators is to adequately prepare doctoral students for success in their upcoming careers. To achieve this, we will need to realign our goals in graduate education with the realities of today’s branching science career pipeline. Pursued simultaneously, the cultural, academic, and policy changes recommended in this and other reports will help us continue to develop talented, confident, and well-trained scientific professionals who will contribute directly to our research enterprise as trainees, then move on to diverse careers that will elevate the pace and quality of scientific discovery, improving the health of our nation and our world.

ACKNOWLEDGMENTS

We thank foremost the many UCSF students and postdoctoral scholars who participated in this study. Thank you to Bruce Alberts, whose advocacy around these issues served as inspiration for this study. C.N.F. thanks Bill Balke, Laurie Littlepage, Françoise Chanut, and Brian Kelch for discussions that clarified the story and Quincey Aanerud R, Homer L, Nerad M, Cerny J (2006). Paths and perceptions: assessing doctoral education using career path analysis. In: The Assessment of Doctoral Education: Emerging Criteria and New Models for Improving Outcomes, ed. PL Maki and NA Borkowski, Sterling, VA: Stylus, 109–141.


Austin AE (2002). Preparing the next generation of faculty: graduate school as socialization to the academic career. J Higher Educ 73, 94–122.


Essay

Teaching Biology through Statistics: Application of Statistical Methods in Genetics and Zoology Courses

Migdalisel Colon-Berlingeri and Patricia A. Burrowes

Department of Biology, University of Puerto Rico, San Juan, PR 00931

Submitted November 4, 2010; Revised June 8, 2011; Accepted June 16, 2011

Monitoring Editor: John R. Jungck

Incorporation of mathematics into biology curricula is critical to underscore for undergraduate students the relevance of mathematics to most fields of biology and the usefulness of developing quantitative process skills demanded in modern biology. At our institution, we have made significant changes to better integrate mathematics into the undergraduate biology curriculum. The curricular revision included changes in the suggested course sequence, addition of statistics and precalculus as prerequisites to core science courses, and incorporating interdisciplinary (math–biology) learning activities in genetics and zoology courses. In this article, we describe the activities developed for these two courses and the assessment tools used to measure the learning that took place with respect to biology and statistics. We distinguished the effectiveness of these learning opportunities in helping students improve their understanding of the math and statistical concepts addressed and, more importantly, their ability to apply them to solve a biological problem. We also identified areas that need emphasis in both biology and mathematics courses. In light of our observations, we recommend best practices that biology and mathematics academic departments can implement to train undergraduates for the demands of modern biology.

INTRODUCTION

Biology and mathematics have been interconnected for a long time. In fact, many biological processes are described by mathematical equations and certain mathematical concepts have arisen directly from the need to describe interactions, relationships, and processes in living systems (Jungck, 1997; Cohen, 2004). This has led to the concept of interdisciplines such as biophysics, bioinformatics, and bioinformatics, to mention a few. Modern technology allows researchers to rapidly generate vast amounts of data that are shared through virtual databases and, depending on how they are analyzed, can serve to answer a variety of questions. Examples are the field of genomics, which demands the need for mathematics and computer science to significantly contribute to modern biology (Ditty et al., 2010); or phylogenetic systematics, which requires mathematical algorithms and statistical tests to propose hypotheses regarding the evolutionary relationships of our biodiversity (Hedges et al., 2008). One specific example of the relevance of this integration is the work done by structural biologists (Tsai et al., 2007). These scientists use crystallography to study the three-dimensional structure of a protein, a process that applies geometry and physics. With the aid of computational models and statistics, structural biologists test several hypotheses to explain a protein’s function based on its structure. Furthermore, they can use bioinformatics to trace the evolutionary history of a protein. Thus, now more than ever, biologists need to be proficient in mathematics and computer science to be able to acquire, analyze, and understand the significance of data (Gross, 2004).

Despite the anticipated demands for greater math emphasis in biology education, curricular reforms have not complied with the need to integrate mathematics and computational sciences into undergraduate biology courses (Bialek and Botstein, 2004; Klymkowsky, 2005). Most science and mathematics courses are taught to undergraduate students as a set of facts isolated from related fields, resulting in

DOI: 10.1187/cbe.10-11-0137

The online version of this essay contains supplemental material at CBE-LSE online at www.lifescied.org.

Address correspondence to: Migdalisel Colon-Berlingeri (lisycolonberlingeri@yahoo.com).

Possible conflict of interest: The authors designed assessment instruments to evaluate the new biology curriculum.

© 2011 M. Colon-Berlingeri and P. A. Burrowes. CBE—Life Sciences Education © 2011 The American Society for Cell Biology. This article is distributed by The American Society for Cell Biology under license from the author(s). It is available to the public under an Attribution–Noncommercial–Share Alike 3.0 Unported Creative Commons License (http://creativecommons.org/licenses/by-nc-sa/3.0).

"ASCB®" and "The American Society for Cell Biology®" are registered trademarks of The American Society for Cell Biology.
students with a shortsighted view of their own disciplines. Undergraduate students majoring in biology often question why they are required to take statistics and other mathematics. The problem becomes more evident when we as researchers have students working in our labs and discover that even bright students, who passed these math and statistics requirements, are incapable of applying the concepts to solve a biological question or as tools for analysis of data they generate in the lab. Therefore, incorporation of mathematics into the biology curriculum is critical to underscore for our students the importance and usefulness of mathematics in most fields of biology (National Research Council, 2003; Marsteller et al., 2010). The challenge is to determine how this can be accomplished. How do we help our students see the relevance of mathematics and statistics in biology? How do we offer our students the tools to confront the new interdisciplinary problems in biology?

At our institution, we have made significant changes to better integrate mathematics into the undergraduate biology curriculum. Curricular revision, implemented in 2008, included changes in the suggested course sequence and the addition of statistics and precalculus as prerequisites to core science courses such as genetics and zoology. In response to these efforts and a National Institutes of Health Minority Access to Research Careers (NIH-MARC)-funded initiative to improve mathematical skills among undergraduate biology students, we decided to modify the way we taught our genetics and zoology courses (Colon, 2010). We provided opportunities to apply statistical methods to solve biological problems based on real data, expecting these opportunities would help our students increase their understanding of fundamental concepts in statistics, see how statistics helped interpret biological patterns, and develop scientific process skills. In this article, we describe the activities developed to attain these goals and the assessment tools employed to measure improved learning in biology and statistics. In light of the current need for the integration of science and mathematics, this work describes a feasible way in which students can be engaged not just in learning about these fields but in valuing the need for computational literacy in biology. We present our efforts in two very different courses, genetics and zoology, as examples of how this can be done with diverse subject matter. Hopefully, students who learn the applicability of math to biology in this way will be better prepared to succeed in graduate studies and will contribute to the development of new ways in which the fields of math and biology can enhance knowledge (Miller and Walston, 2010).

The time—now—and place—Puerto Rico—chosen to implement efforts to improve mathematical skills among biology students is just right for several reasons. The University of Puerto Rico (UPR) has a history of National Science Foundation (NSF)-funded initiatives to improve teacher preparation in the sciences and mathematics (Scope Sequence and Coordination; Systemic Statewide Initiatives; Collaborative for Excellence in Teacher Preparation; Math and Science Partnerships). Despite these efforts, teachers are still deficient in quantitative skills and limited in their ability to show their students the connections between these disciplines (Quintero, 2006). Future math and biology teachers learn their science content in our courses together with science majors. Therefore, the learning activities we provide in the classroom have the potential to influence students further down on the educational ladder. In addition, the curricular reform in our biology department calls for student involvement in undergraduate research. Development of quantitative skills with relevance to biological problems, as modeled herein, will help these students to see the application of mathematics to bioinformatics, biotechnology, rates of change, and data analysis. Finally, the Puerto Rican government has proposed to improve the local economy by stimulating the biotechnology industry in what they have referred to as an “economy of knowledge” (INDUNIV Research Consortium, 2010). As a result, several consortia between industry and UPR have been developed to promote student training in biotechnology and development of scientific process skills (Potera, 2007). Undergraduate teaching initiatives that promote interdisciplinarity between math and science are key to achieving this goal.

**STUDENT PROFILE AND TARGETED COURSES**

The student populations in the genetics and zoology courses are similar, as they have already completed a year of introductory biology; however, while statistics is a prerequisite for genetics, not all students in zoology have taken it.

**Description of Activities in the Genetics Course**

Genetics is a 1.5-h lecture course taught twice weekly independent of the genetics laboratory. Students usually take the lab the semester after they take the lecture course. Most of the activities that emphasized development of quantitative skills in genetics were conducted in the first 6 wk of the semester. These activities involved solving genetic crosses; analyzing data describing the distribution of quantitative traits, and allelic, genotypic, and phenotypic frequencies; and genetic mapping of eukaryotic chromosomes. Typically, students allocated 30–40 min of the class period to solving and discussing problem sets in cooperative groups. Additional quantitative exercises with supplementary explanation modules were provided for students to solve at home and discuss later in the week in scheduled peer-tutoring sessions and/or office hours with the professor. The Blackboard (2011) academic online platform was used to distribute supplementary modules, practice exercises, and weekly quizzes that enforced student preparation. The supplementary statistics unit addressed concepts relevant to the analysis of genetics data, such as the basic laws of probability, binomial and chi-square probability distributions, defining a null hypothesis, and the application of the chi-square as a “goodness of fit” test between observed and expected genetic outcomes. The peer-tutoring walk-in sessions were an important component of this course, because they provided another opportunity for students to get help on the assigned problem sets. These peer mentors were volunteer undergraduate students who had passed the genetics course in previous semesters with an A or B. The genetics tutors held weekly meetings with the professor (M.C.B.) to discuss the material to be covered in class, clarify their own doubts, and conceive ways in which they would work with the students. As part of their training, tutors also took a weekly quiz through Blackboard that included content-related questions, as well as math–genetics exercises.

With the purpose of measuring the understanding of basic statistics concepts that students brought into the genetics course, and later, assessing the knowledge gained from
curricular activities provided throughout the course, we implemented an assessment instrument in a pretest/posttest manner. For this, we translated and slightly modified a survey by Metz (2008) designed to measure knowledge of statistics among incoming students in general biology. The modified test instrument included two different surveys: a 12-item pretest and a 12-item posttest that shared only five identical questions. The other seven questions measured knowledge of the same concepts, but encouraged higher-level thinking skills by requiring application of the quantitative concepts discussed in class to solve genetics problems. The concepts assessed in this pretest/posttest manner included: basic probability, descriptive statistics, p value, r value, graph interpretations, and cause-and-effect relationships (Metz, 2008). Content validity was achieved by the critical examination of the test by three other genetics professors and two biology professors engaged in science education and math–biology integration. The pretest survey was administered on the first day of class, whereas the posttest was administered during the last week of the semester. Students were asked to voluntarily complete both surveys. Although the quantitative activities were offered to all the students in the course (n = 220), we report the learning results of only a small session taught by one of the authors (M.C.B.). Of 33 students registered for this session, only 16 students took both the pretest and the posttest. Analyses of the results for the 16 students taking the pretest and the posttest included a graphic view of the change in score distribution between the pretest and the posttest (Figure 1), and a chi-square test to determine whether the levels of achievement per concept were independent of the knowledge gained after instruction (pretest vs. posttest).

All the class and assessment materials, including the pretests and posttests, are available upon request in Spanish and English (see Supplemental Material 1 and 2 or contact the corresponding author).

**Description of Activities in the Zoology Course**

A total of 40 students were enrolled in the zoology course. During the semester, students undertook three projects requiring the application of some kind of computational skill to answer questions about animals. The first project aimed at developing simple bioinformatics skills by requiring consultation, analysis, and synthesis of data available in the International Union for Conservation of Nature website. The second project required the application of statistical methods to determine morphological patterns in bats. The third project involved the application of precalculus and calculus concepts to develop a model of sustainable yield for two valuable resources in an Amazonian forest. The projects were presented in class as biological problems, and the needed statistical or mathematical skills were discussed and explained in the context of the question. Students worked in pairs, and had 2 wk to turn in a five-page written report, which included appropriate tables and figures. Because we focus on the integration of statistics to biology in this paper, we limit our discussion to the second project, which dealt with bat morphology and statistics.

Students were given an Excel worksheet with data on the wingspan (cm), weight (g), and sex (male vs. female) of a neotropical bat, *Artibeus jamaicensis*, commonly known as the fruit bat because of its alimentary habits. The data were based on field observations (Rodriguez-Duran, 2005) but were artificially generated to ensure that they would meet all the assumptions of parametric statistics. In class, we discussed the ecology and reproductive biology of this bat, mentioning relevant details, such as the fact that Caribbean fruit bats reproduce twice a year and female fruit bats give birth to a single offspring, which they carry around, and lactate for 3–4 mo (Gannon et al., 2005). Students were expected to analyze the data by applying simple descriptive (mean, variance, SD, range) and parametric (t test and linear regression) statistics in order to answer the biological and statistical questions described in Table 1. Before analyzing the data, students were required to formulate a biological hypothesis for each question and to state the null and alternative hypotheses for each statistical test considered. Finally, students were expected to predict which sex was capable of carrying more weight during flight, and discuss what selective forces may have favored the evolution of the characteristics of body size and morphology observed in these bats.

To assess the heterogeneity of the student population, we used clickers (InterWrite Personal Response System, GTCO Calcomp, Columbia, MD; Smith et al., 2011) to ask questions about their gender, ethnicity, year at the university, courses approved, research experience, and, more

![Figure 1](https://example.com/figure1.png)
relevant to this project, their background in and attitude toward math/statistics. The effectiveness of the bat activity in promoting biological and statistical knowledge was assessed by their performance in their written reports, which was evaluated using a rubric that measured the level of achievement of the following goals: A) the biological hypotheses reflect knowledge on the biology, ecology, or physiology of the bats; B) the statistical hypotheses ($H_0$ and $H_a$) were correctly stated for every test applied; C) the $p$ value associated with each statistic was properly reported and correctly interpreted biologically; D) the data were graphed in a way that would properly describe a pattern; E) the biological interpretation of the graph was accurate; and F–J) the answers to statistical concept questions (1–5) as described in Table 1 were correct. Students could score within a range of 1–4 for each of these learning objectives, where 4 = completely achieved, 3 = somewhat achieved, 2 = poorly achieved, and 1 = not achieved (see Table 2).

### OUTCOMES

#### Genetics

The overall student achievement, evaluated by the distribution of student test scores, improved after instruction, in spite of the increase in complexity of some of the questions in the posttest (Figure 1A and Tables 3 and 4). When student performance was categorized by levels of achievement as low (0–5 points/12 available points), medium (6–8 points/12 points), or high (9–12 points/12 points), it was evident that the greatest change after instruction was among students improving from a low to a medium level (Figure 1B). In the pretest, close to 35% of the students scored poorly, while in the posttest more than one-half of these students improved to the medium level of achievement (Figure 1B).

By looking at student performance on each of the 12 items, we were able to identify specific skills where students showed improvement in the posttest. Table 3 summarizes student outcomes on questions that were identical in the pretest and the posttest, considerable improvement was shown on most items. These questions included statistical concepts of basic probability, probability applied to genetics, and data interpretation, which were discussed during the first 6 wk of the course relevant to the topics of Mendelian genetics, population genetics, and quantitative genetics. The one item that showed no change, but was well achieved even in the pretest, asked students to predict the gametes involved in a cross between individuals of known genotype. This suggests that students understood this basic Mendelian concept, perhaps from studying it in their general biology course. The item in this group with the lowest correct answer rate dealt with interpretation of data comparing three blood groups of two

### Table 1. Biological and statistical questions addressed by the students through the bat activity in zoology class

<table>
<thead>
<tr>
<th>Biological questions</th>
<th>Statistical understanding questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Do you expect to see differences in the wingspan and weight between male and female bats? Hypothesize.</td>
<td>1. Test for differences in wingspan and weight between sexes. State your null and alternative hypotheses. What does it mean to have statistical significance?</td>
</tr>
<tr>
<td>2. If there are differences, can they be appreciated graphically? Illustrate.</td>
<td>2. What probability value must be associated to a statistic in order to reject your null hypothesis? What does it mean to have a $p &gt; 0.05$?</td>
</tr>
<tr>
<td>3. Do you expect to see a relationship between wingspan and weight for all bats? Predict how this relationship might be and justify.</td>
<td>3. What does it mean to have a significant Pearson product-moment correlation coefficient ($r$) between two variables?</td>
</tr>
<tr>
<td>4. If present, can this relationship be illustrated graphically?</td>
<td>4. What is the difference between a positive or negative $r$?</td>
</tr>
<tr>
<td>5. Does this relationship hold for males? For females? Explain.</td>
<td>5. What additional information can you get from a linear regression if you have already determined a significant positive $r$?</td>
</tr>
</tbody>
</table>

#### Table 2. Results of rubric to assess learning gain after the zoology bat activity

<table>
<thead>
<tr>
<th>Learning objective</th>
<th>Achievement level score</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. The biological hypotheses reflect knowledge on the biology, ecology, or physiology of the bats</td>
<td>4 (25%) 6 (30%) 8 (40%) 1 (5%)</td>
</tr>
<tr>
<td>B. The statistical hypotheses ($H_0$ and $H_a$) were correctly stated for every test applied</td>
<td>6 (30%) 3 (15%) 9 (45%) 2 (10%)</td>
</tr>
<tr>
<td>C. The $p$ value associated with each statistic was properly reported and correctly interpreted biologically</td>
<td>3 (15%) 4 (20%) 6 (30%) 7 (35%)</td>
</tr>
<tr>
<td>D. The data were graphed in a way that would properly describe a pattern</td>
<td>14 (70%) 4 (20%) 2 (10%) 0 (0%)</td>
</tr>
<tr>
<td>E. The biological interpretation of the graph was accurate.</td>
<td>6 (30%) 6 (30%) 5 (25%) 3 (15%)</td>
</tr>
<tr>
<td>F. Understanding statistical significance</td>
<td>3 (15%) 4 (20%) 6 (30%) 7 (35%)</td>
</tr>
<tr>
<td>G. Understanding $p$ value</td>
<td>3 (15%) 4 (20%) 6 (30%) 7 (35%)</td>
</tr>
<tr>
<td>H. Understanding applicability of Pearson product-moment correlation coefficient</td>
<td>6 (30%) 6 (30%) 5 (25%) 3 (15%)</td>
</tr>
<tr>
<td>I. Differentiating between positive and negative correlations</td>
<td>6 (30%) 7 (35%) 5 (25%) 2 (10%)</td>
</tr>
<tr>
<td>J. Understanding the added mathematical value of a linear regression</td>
<td>3 (15%) 3 (15%) 7 (35%) 7 (35%)</td>
</tr>
</tbody>
</table>

* Students could score between 1 and 4 for each learning objective, where 4 = completely achieved, 3 = somewhat achieved, 2 = poorly achieved, and 1 = not achieved.
different populations (Table 3; 10 in the pretest, eight in the posttest). In spite of instruction, only 50% of the students were able to determine allelic frequencies for blood groups, given the phenotypic and genotypic frequencies for two populations. This question was at the highest level of Bloom’s taxonomy of learning domains (Bloom and Krathwohl, 1956), because it required comparison, evaluation, and formulation of a conclusion regarding the phenotypic, genotypic, and allelic frequencies of two populations. Overall, the identical pretest/posttest questions revealed that, while most students acquired knowledge in statistics that improved their ability to apply statistics to genetics, only a few students reached higher thinking levels requiring evaluation and synthesis (creation) of conclusions.

Table 4 summarizes student outcomes in the questions that were modified from the pretest and the posttest. All of these questions involved analysis of quantitative concepts in a biological context. Students improved mostly in skills that required interpretation of graphs and correlation coefficients. This outcome suggests that efforts to provide exercises where students had to discern scientific patterns from graphical data and calculate correlation coefficients in class proved beneficial (see item 12 in the pretest and 3 in the posttest in Table 4). However, the concept of probability did not show considerable improvement, perhaps because the questions were applied to difficult genetics concepts that required higher levels of analysis than the ones presented in the pretest. Questions dealing with interpretation of \( p \) value presented a decrease in correct responses in the posttest. The items testing knowledge and understanding of the \( p \) value (items 4 and 7) had different elements and distractors in the posttest that augmented the difficulty and the level of analysis required to answer the question. The decrease in scores observed for these items shows students had a very basic knowledge of the concept of \( p \) value before the genetics course, and our efforts to improve their understanding of it to a point where they could synthesize resulting data and create a conclusion were not enough.

**Zoology**

The population of students enrolled in this experimental zoology course was 100% Hispanic, 75% female, 65% upperclassmen (third year or greater), and, as far as their mathematical background, 100% had taken college precalculus and 65% had taken calculus, but only 52% of the students had taken a statistics course. When asked “Do you think that mathematics is very useful for biologists?” in a pre/postinstruction manner, the number of students agreeing increased at the end of the semester, and their responses were significantly associated with postinstruction (\( \chi^2 = 22.789, \text{ degrees of freedom} = 3, p = 0.000 \)). Table 2 and Figure 2 summarize the learning gain in biology and statistics obtained by zoology students after working on the bat activity. The results show that learning objectives B, C, F, G, and J were not achieved by students, regardless of class instruction and the opportunity to work on a relevant biological problem requiring the application of statistical methods. Misconceptions were noted in the inability to formulate hypotheses and understanding of the significance of the probability that statistical tests associate to these hypotheses. For example, when obtaining a \( p \) value ≤ 0.05 for an \( f \) statistic of a regression equation between two variables, many students did not know whether they should reject the null hypothesis (\( H_0: \) there is no relationship between the variables; slope = 0) and accept the alternative (\( H_1: \) there is a linear relationship between the variables; slope \( \neq 0 \)), or vice versa. This suggests that students confused type I (\( \alpha \)) and type II (\( \beta \)) errors and, as a consequence, often drew incorrect conclusions on their results. Although the concept of correlation between variables with sex was understood, as indicated by students being able to ascertain that bat weight and wingspan were associated with sex from calculating high Pearson product-moment correlation coefficients (\( r \)), students were unable to distinguish the added value of a regression analysis. Only 30% of the student reports showed complete or partial achievement of this objective (Table 2). This small group of students included at least two of the following critical elements in their answers: 1) a discussion of how regression can add information on rate of change by providing a slope for a line describing the relationship between the dependent (weight) and independent (wingspan) variables; 2) a description of how the linear model can serve to predict the wingspan of a bat with a particular weight; 3) an acknowledgment that regression might suggest a nonlinear association between the variables; and/or 4) a mention of the value of \( r^2 \) when discussing the fit of the data to the linear model. On the other hand, students were capable of stating informed biological hypotheses related to the physiology of the bats, had no trouble translating numbered data into graphs that

<table>
<thead>
<tr>
<th>Quantitative concept</th>
<th>Question</th>
<th>% Correct answers</th>
<th>Pretest</th>
<th>Posttest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basic probability</td>
<td>A bag has seven red balls and two white balls; what is the probability of drawing a white ball?</td>
<td>62.5</td>
<td>87.5</td>
<td></td>
</tr>
<tr>
<td>Probability applied to genetics</td>
<td>What combination of gametes is generated from an organism with genotype EeVv?</td>
<td>93.8</td>
<td>87.5</td>
<td></td>
</tr>
<tr>
<td>Probability applied to genetics</td>
<td>What percentage of EeVv organisms will result from the cross between EeVv and eevv?</td>
<td>56.3</td>
<td>75</td>
<td></td>
</tr>
<tr>
<td>Data interpretation</td>
<td>Given the number of individuals with blood types M, MN, and N in two different populations, determine the frequency of the M allele.</td>
<td>56.3</td>
<td>62.5</td>
<td></td>
</tr>
<tr>
<td>Data interpretation</td>
<td>Given the number of individuals with blood types M, MN, and N in two different populations, choose the best interpretation of the results.</td>
<td>37.5</td>
<td>50</td>
<td></td>
</tr>
</tbody>
</table>

\(^{a}\) The alternative choices for each of the questions can be obtained by requesting a copy of the tests from the authors.
Table 4. Item analysis of different questions in the pretest and posttest in genetics ($n = 16$)\(^a\)

<table>
<thead>
<tr>
<th>General concept</th>
<th>Pretest</th>
<th></th>
<th>Posttest</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Item number</td>
<td>Bloom’s taxonomy</td>
<td>Question</td>
<td>% of correct answers</td>
<td>Item number</td>
</tr>
<tr>
<td>Probability</td>
<td>2</td>
<td>Application</td>
<td>Two fertilization events are independent. What is the probability that a couple with two boys will have another boy?</td>
<td>81.3</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>Application</td>
<td>A bag has seven red balls and two white balls. What is the probability that when drawing two balls, both of them are white?</td>
<td>25</td>
<td>12</td>
</tr>
<tr>
<td>Graph</td>
<td>3</td>
<td>Analysis</td>
<td>Which type of graph is used to evaluate the relationship between two quantitative variables?</td>
<td>62.5</td>
<td>2</td>
</tr>
<tr>
<td>interpretation</td>
<td>4</td>
<td>Analysis</td>
<td>Which type of graph is used to evaluate the distribution of potato weights under certain conditions?</td>
<td>31.3</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>Evaluation</td>
<td>Given two different graphs that plot the number of individuals with three different genotypes vs. weight, choose the sentence that better describes the data.</td>
<td>18.8</td>
<td>9</td>
</tr>
<tr>
<td>Correlation</td>
<td>12</td>
<td>Evaluation</td>
<td>A table with three columns with data from cattle is shown: father weight (Z), mother weight (X), and progeny weight (Y). Interpret the $r$ value.</td>
<td>18.8</td>
<td>3</td>
</tr>
<tr>
<td>coefficient</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$p$ Value</td>
<td>5</td>
<td>Evaluation</td>
<td>Given the average peak size in two populations of <em>Geospiza fortis</em>, and the $p$ value, determine if the differences between the two populations is significant.</td>
<td>56.3</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Creating</td>
<td>The chi-square analysis of a dihybrid cross determines that the differences between the observed and the expected frequencies in F2 are not significant. If the critical value is 0.5, find the associated $p$ value.</td>
<td></td>
<td>4</td>
</tr>
</tbody>
</table>

\(^a\)The alternative choices for each of the questions can be obtained by requesting a copy of the tests from the authors.
CONCLUSIONS AND RECOMMENDATIONS

This paper presents the results of two independent projects aimed at integrating statistical concepts into undergraduate biology courses—genetics and zoology—in two very different ways. In general, we found an increase of understanding of and ability to apply statistical concepts of correlation and association between variables to biology. We also saw improvement in students’ ability to use graphs to describe and interpret patterns in numerical data. In addition, students expressed their understanding of the relevance of statistics as a tool to analyze biological data and understand its significance. However, we identified other concepts for which students had more difficulty demonstrating knowledge gain despite in-class instruction and educational activities. Interestingly, although the instructional approach and the assessment instruments used were different, these two courses yielded similar results. We will discuss the implications of these results in light of the challenges we face in order to further integrate math and biology, and we provide some suggestions to meet the need for interdisciplinary teaching in these fields.

Areas That Need Emphasis

The lack of understanding of probability distributions, interpretation of the $p$ value associated with statistics, and when to reject a null hypothesis was evident in both courses. This was observed in spite of the fact that more than 50% of our students had completed an introductory statistics course and received additional instruction in the statistical concepts relevant to biology in our classes. Mathematical concepts associated with linear relationships between two variables were also weak. Specific to the students in the zoology course was the inability to extrapolate the biological meaning of strong positive or negative slopes versus flat slopes describing the relationship between two variables, a concept that can be traced back to high school geometry and algebra. In both courses, we could see that students were able to distinguish the meaning of high or low correlation coefficients, but it was then difficult for them to synthesize a conclusion explaining the relationship between the variables involved, and to predict potential biological causation for the event described by the data.

How to Confront the Challenges

The challenges biology educators face today are multidimensional. First, biology courses generally do not emphasize the role of mathematical analysis in the description of biological data, a critical element in many major scientific discoveries. Second, most mathematics courses do not make a connection between mathematical concepts and applications to other fields of science (Robeva and Laubenbacher, 2009). The lack of connection between the general statistics course required for majors in biology and the content discussed in the biology courses is also evident (A’Brook and Weyers, 1996; Metz, 2008). Our observations underscore the need for interaction and collaboration to provide new alternatives to traditional math courses (Marsteller et al., 2010). It has been shown that a course integrating math and biology concepts does not hinder student learning in either of the two content areas, but rather may enhance interdisciplinary knowledge (Madlung et al., 2011). Thus, we suggest that professors of introductory courses in math, statistics, and biology convene to discuss these issues. Such collaboration could lead to new math courses or supplementary instruction involving biological examples in calculus and statistics courses. This can be done by sharing research data and/or discussing interdisciplinary papers dealing with computational biology and effectively planning ways to present this information in their classes (Robeva et al., 2010; Watkins, 2010). In addition, instructors may consider using or modifying the activities available from the BioQUEST Curriculum Consortium (2011) or resources from the National Institute of Mathematics and Biological Sciences (NIMBIOS; 2011).

The fact that students take math requirements at different points in the curriculum and come to biology courses with widely divergent quantitative skills is yet another challenge. To rectify this, we recommend that faculty members in biology departments reevaluate their undergraduate curricular sequence, so that students take calculus and basic statistics during their freshman year, when they are also taking their general biology courses. In this manner, they can move on to higher-level courses, such as genetics, microbiology, botany, zoology, and biotechnology, with basic concepts that will allow them to apply mathematical tools to data generated in laboratories, interpretation of graphics, and the understanding of modern scientific theory.

Student attitude is also a challenge, because students may resist learning more math than required for math courses and may tend to avoid the biology professors who incorporate math into their curricula. However, if a coordinated effort is developed among the departments, such that the true interdisciplinary nature between math and biology is highlighted in all courses across the curriculum, we would expect this feeling to be ameliorated with time, and students will value the opportunity to apply what they have learned throughout their career as they construct new knowledge. In fact, in the semesters following this study, we have observed a change in student attitude; students seem more receptive to the idea
that they will do “a lot of math” in our courses. Specifically, for genetics, we have observed an improvement in student grades for exams one and two, which are mostly quantitative, compared with previous semesters. We will evaluate whether this improvement in grades and attitude is due to the curricular changes, or because a subset of students who are more advanced in math and in the curriculum in general are concentrated in our sections.

Finally, we must admit that professors of mathematics and biology are not always capable of bridging between these two subject areas with the best examples. This might be resolved by a summer workshop for biology professors given by a biostatistician or an expert in bioinformatics or mathematical ecology who has experience and interest in the teaching/learning processes. The idea of the workshop would be to work out specific problems where statistical methods and mathematical algorithms contribute to better understanding of biological processes. NIMBIOS offers a variety of investigative workshops that may provide professional development in this area (NIMBIOS, 2011).

The incorporation of math in biology is now an irrefutable need. As mentioned earlier, modern fields of biology require mathematical and computational analysis of large amounts of data that help to predict models and describe the function of biological processes at the ecological, organismal, and cellular level (Elser and Hamilton, 2007). Although modern fields like bioinformatics and biostatistics reflect the interdisciplinary nature of biological sciences today, the issue of the lack of emphasis of mathematical concepts in biology education is still evident. In this paper, we present alternative ways to incorporate into biology curricula the application of mathematical tools, specifically statistics, that help students see the advantage of developing quantitative skills in order to analyze different kinds of biological data and discover the patterns and processes occurring in that data. We expect our work to serve as an example for professors in different fields of biology on providing learning opportunities to bridge the gap between math and biology, identifying specific areas that need to be improved, and devising alternative ways to address the gap. The materials and assessment instruments used in this study are available upon request in Spanish and English. These materials will also be very useful for high school teachers interested in using activities that link math and biology. In the future, as our curriculum becomes more interdisciplinary, and we put in practice some of our own recommendations, we expect to adopt a case-based approach, where students design their own experiments, as exemplified by the BioQUEST Curriculum. In this way, we will encourage the development of higher-order thinking skills, while illustrating the relevance of math and computational skills to understanding biological processes.

REFERENCES


ACKNOWLEDGMENTS

This work was supported in part by a curriculum improvement project funded by an NIH-MARC supplement to the UPR-Rio Piedras program 5T36GM078010-02 directed by Michelle Borrero. We are grateful to our colleagues Rosaura Ramirez and Michelle Borrero for motivation and helpful comments in the development of the activities. M.C.B. acknowledges the contribution of other genetics professors, Noemi Cintron, Tomas Hrbeck, and Jose L. Agosto in revising the assessment instruments implemented in the course.


Increased Course Structure Improves Performance in Introductory Biology

Scott Freeman, David Haak,* and Mary Pat Wenderoth

Department of Biology, University of Washington, Seattle, WA 98195

Submitted August 26, 2010; Revised February 24, 2011; Accepted February 25, 2011
Monitoring Editor: Diane Ebert-May

We tested the hypothesis that highly structured course designs, which implement reading quizzes and/or extensive in-class active-learning activities and weekly practice exams, can lower failure rates in an introductory biology course for majors, compared with low-structure course designs that are based on lecturing and a few high-risk assessments. We controlled for 1) instructor effects by analyzing data from quarters when the same instructor taught the course, 2) exam equivalence with new assessments called the Weighted Bloom’s Index and Predicted Exam Score, and 3) student equivalence using a regression-based Predicted Grade. We also tested the hypothesis that points from reading quizzes, clicker questions, and other “practice” assessments in highly structured courses inflate grades and confound comparisons with low-structure course designs. We found no evidence that points from active-learning exercises inflate grades or reduce the impact of exams on final grades. When we controlled for variation in student ability, failure rates were lower in a moderately structured course design and were dramatically lower in a highly structured course design. This result supports the hypothesis that active-learning exercises can make students more skilled learners and help bridge the gap between poorly prepared students and their better-prepared peers.

INTRODUCTION

In 1920, <4% of the U.S. population went to college (Ratcliff, 2010). Now, >55% of the general population over the age of 25 has at least some college experience (U.S. Census Bureau, 2009). In the United States, the democratization of higher education began with the founding of the land grant (public) universities in the 1860s, continued with the founding of community colleges in the early 1900s, accelerated with the G.I. Bill that passed in 1944, and culminated with the expansion of the women’s movement and a long-delayed end to exclusion based on race in the 1960s and 1970s (Eckel and King, 2006). Indeed, increased access to higher education is occurring internationally (e.g., Scott, 1995).

For faculty, the democratization of higher education means that an increasingly smaller percentage of students come from privileged social and economic backgrounds. Although faculty should celebrate this fact, it is common to hear instructors express concern about the downside of democratization: high variation in student ability and preparedness. Data from the ACT (2006), for example, suggest that 49% of high school students who took the ACT college entrance exam in 2005 were not ready for college-level reading.

How can we help underprepared but capable students succeed, while continuing to challenge better-prepared students? The issue is particularly acute in gateway courses—the large, introductory classes that undergraduates take in their first or second year. Across the science, technology, engineering, and mathematics (STEM) disciplines, failure rates in these courses can be high—even in moderately and highly selective schools, where students are “prescreened” on the basis of academic capability. Although we are not aware of a comprehensive review, it appears common for one-third of students to fail in STEM gateway courses (Table 1).

Failure has grave consequences (Wischusen and Wischusen, 2007). In addition to the emotional and financial toll that failing students bear, they may take longer to graduate, leave the STEM disciplines, or drop out of school entirely.
Analysts have been particularly concerned that underrepresented minorities (URMs) and other students who quit the STEM disciplines take valuable perspectives and creativity with them (Seymour and Hewitt, 1997). Students who repeat gateway courses also add pressure on enrollments—in some cases, denying places to students who want to take the course for the first time.

How can we reduce failure rates in STEM courses, without reducing rigor? Recent research suggests that changes in course design—specifically, the introduction of active-learning strategies—can help. For example, Beichner et al. (2007) reported that changing to workshop (or studio) models of instruction, which emphasize collaborative group work in class, cut failure rates by 40–60% in introductory physics across a range of institutions. Similarly, Lasry et al. (2008) found that the use of peer instruction with clickers (see Mazur, 1997) reduced the drop rate in introductory physics at a community college and at a research university by factors of two to three.

We hypothesized that intensive active learning, combined with frequent formative assessment, can lower failure rates in an introductory biology course for majors. Previously we reported an average increase of 3–4% on exam scores and a 30% reduction in failure rate when we introduced active-learning exercises that were graded (Freeman et al., 2007). Subsequently we have introduced additional active learning exercises in an attempt to increase exam performance and reduce the failure rate even further. The result is what we term a highly structured course design. Low-structure courses are based on traditional lecturing and high-stakes assessments—typically two or three midterms and a comprehensive final exam. In contrast, highly structured courses assign daily and weekly active-learning exercises with the goal of providing constant practice with the analytical skills required to do well on exams.

If highly structured courses lower failure rates, faculty might find their increasingly diverse student population a source of reward and inspiration, rather than frustration. But do highly structured courses work? Do they lead to greater success and fewer failures? Testing the highly structured course design hypothesis requires that instructor identity, exam difficulty, and student ability be controlled across the classes being studied.

**METHODS**

**Course Background**

This research focused on students in Biology 180, the first in a three-quarter introductory biology sequence designed for undergraduates intending to major in biology or related disciplines at the University of Washington (UW). The course introduces evolution, Mendelian genetics, diversity of life, and ecology. The course is offered every quarter; in the 2009–2010 academic year, total enrollment approached 2100.

Although the course is usually team-taught, the data analyzed here come from six quarters when one of the authors (S.F.) was the sole instructor of record. Throughout the study, the course was offered for five credits and included four 50-min class sessions and a 2- or 3-h laboratory each week. In every quarter there were 400 exam points possible; all exam questions were written—most were short-answer, but some involved graphing, computation, or labeling diagrams.

**Student Demographics**

During the study period, most students had to complete a chemistry prerequisite before registering for Biology 180; the majority were in their sophomore year. In the most recent quarter analyzed here, the course’s demographic makeup was 61.1% female and 38.9% male; 46.7% Caucasian, 38.5% Asian-American, and 7.4% URM (African-American, Hispanic, Native American, or Pacific Islander), with 7.4% of students declining to declare their ethnicity. In addition, 16.4% of the students were in the UW Educational Opportunity Program, meaning that they were identified as economically or educationally disadvantaged. These demographic data are typical across the quarters and years of the study.

**Course Design**

During the six quarters analyzed in this study, the instructor used various combinations of teaching strategies, detailed here in order of implementation.

---

**Table 1. Failure rates in some gateway STEM courses**

<table>
<thead>
<tr>
<th>Field</th>
<th>Course</th>
<th>Failure rate</th>
<th>Failure criterion</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Biology</td>
<td>Intro-majors</td>
<td>56%</td>
<td>Average proportion of Ds and Fs on exams</td>
<td>Burrowes, 2003</td>
</tr>
<tr>
<td></td>
<td>Intro-majors</td>
<td>&gt;25%</td>
<td>Course outcome: D, F, or drop</td>
<td>Wischusen and Wischusen, 2007</td>
</tr>
<tr>
<td></td>
<td>Intro-nonmajors</td>
<td>27%</td>
<td>Course outcome: D, F, or drop</td>
<td>Marrs and Chism, 2005</td>
</tr>
<tr>
<td></td>
<td>Biochemistry</td>
<td>85%</td>
<td>F on first exam</td>
<td>Peters, 2005</td>
</tr>
<tr>
<td></td>
<td>Medical Microbiology</td>
<td>30%</td>
<td>Course outcome: D or F</td>
<td>Margulies and Ghent, 2005</td>
</tr>
<tr>
<td>Chemistry</td>
<td>Intro-majors</td>
<td>~50%</td>
<td>Course outcome: D, F, or drop</td>
<td>Reardon et al., 2010</td>
</tr>
<tr>
<td></td>
<td>Intro-nonmajors</td>
<td>≥30%</td>
<td>Course outcome (“at most institutions”): fail or drop</td>
<td>Rowe, 1983</td>
</tr>
<tr>
<td>Computer science</td>
<td>Intro to programming</td>
<td>33%</td>
<td>Course outcome (international survey): F or drop</td>
<td>Bennedsen and Casperson, 2007</td>
</tr>
<tr>
<td>Engineering</td>
<td>Intro to chemical engineering</td>
<td>32%</td>
<td>Course outcome: D, F, or drop</td>
<td>Felder et al., 1998</td>
</tr>
<tr>
<td>Mathematics</td>
<td>First-year calculus</td>
<td>42%</td>
<td>Course outcome (U.S. national average): failure</td>
<td>Treisman, 1992</td>
</tr>
<tr>
<td>Physics</td>
<td>Intro-majors</td>
<td>33%</td>
<td>Course outcome: D, F, or drop</td>
<td>Marrs and Chism, 2005</td>
</tr>
</tbody>
</table>
- **Socratic lecturing** involved frequent use of questions posed to the class, with answers solicited from students who raised their hands. In addition to calling on “high responders,” wider participation was encouraged by use of think/pair/share (Lyman, 1981), asking for a response from students in a particular section of the room, or asking for a response from a student who had not contributed before. The intent of Socratic lecturing was to engage student attention and provide feedback to the instructor.

- **Ungraded active-learning exercises** encouraged active participation in class. The exercises used were minute papers (Mosteller, 1989; Angelo and Cross, 1993; Boyd, 2001), case studies with question sets completed by informal groups (Yadav et al., 2007), writing answers to exam-style questions followed by discussion, and in-class demonstrations with student participation (Milner-Bolotin et al., 2007). In most cases, each class session involved at least three such exercises. The intent of the ungraded exercises was to give students practice with the higher-order cognitive skills required to do well on exams.

- **Clicker questions** were multiple-choice questions presented in class; students responded with a personal response device or “clicker” (Caldwell, 2007). Clicker questions were implemented in a peer instruction format, where individuals answered on their own and then reanswered after discussion with one or more students seated next to them (Mazur, 1997; Smith et al., 2009). The instructor asked three to five clicker questions per class session; in most cases, a maximum of three clicker points was possible each class session, assigned for right/wrong responses (see Freeman et al., 2007). Typically, clicker questions summed to ~12% of the total course points. The intent of the clicker questions was to develop student thinking at the application and analysis level and encourage peer teaching.

- **Practice exams** were online, weekly, peer-graded exercises where students submitted written responses to exam-style questions. Students were given 35 min to respond to five short-answer questions. After answers were submitted, software developed in our department randomly and anonymously gave each student a set of answers to grade, along with sample answers and grading rubrics (Freeman et al., 2007; Freeman and Parks, 2010). At two points per question, there were 10 points possible each week—representing ~8% of the total course grade. The intent of the exercises was to give students practice with answering high-level, exam-style, written questions under time pressure, but in a low-stakes environment.

- **Class notes summaries** were weekly assignments that required students to state the three most important concepts introduced each day in lecture, along with a question based on the idea that they understood least well in that class session. The summaries were filled out online and were due each Monday morning. Students were given a course point per week for participation, with a five-point bonus for completing the exercise every week of the course—for a total of ~2% of total course points. The objectives of class notes summaries were to help students 1) organize and synthesize their course material, and 2) increase metacognition—specifically, the ability to identify which information is most important and which concepts are understood most poorly (Bransford et al., 2000).

- **Reading quizzes** opened every afternoon after class and closed the morning of the next class (Novak et al., 1999; Crouch and Mazur, 2001). They consisted of 5–10 multiple-choice questions, delivered and corrected via an online quizzing system, and tested understanding of basic vocabulary and concepts. The exercises were open-book and open-note; students were free to do them in groups or individually. Typically, the two-point reading quizzes summed to ~8% of total course points. The intent of the reading quizzes was to make students responsible for learning basic course content on their own and prepare them to work on higher-order cognitive skills in class.

- **In-class group exercises** involved informal groups of three or four students sitting adjacent to one another. The exercises consisted of one to three exam-style questions on the topic currently being discussed in class (Farrell et al., 1999; Eberlein et al., 2008). As students discussed the questions, graduate and peer teaching assistants (TAs) moved around the lecture hall to monitor the conversations and answer queries from students. Although no course points were awarded during these activities, participation was encouraged because the instructor closed the small-group discussions and then called on students, from a randomized class list, to solicit student responses in front of the entire class. Typically, a single 50-min class session included five or six group exercises, with 12–15 students called on each day. These class sessions, then, consisted of a series of 3- to 5-min mini-lectures that introduced or discussed either a clicker question or a group exercise. These sessions differed dramatically from the ungraded active-learning exercises introduced before, for two reasons: There were more than double the number of activities per class session, and participation—“enforced” by random-call versus calling on volunteers—appeared much higher. The intent of the group exercises was to help students develop higher-order cognitive skills, with peer and TA feedback, in a low-stakes environment.

Over the six quarters in the study—when the same instructor taught the course—the courses can be placed into three categories: 1) relatively low structure in Spring 2002 and Spring 2003; 2) moderate structure, due to the addition of clickers and practice exams, in Spring 2005 and Autumn 2005; and 3) high structure, due to the addition of reading quizzes, along with the substitution of in-class group exercises for Socratic lecturing, in Autumn 2007 and Autumn 2009 (Table 2).

Several other observations are noteworthy: 1) The Spring and Autumn 2005 sections were involved in experiments on the use of clickers versus cards and grading clicker questions for participation versus right/wrong (see Freeman et al., 2007); 2) the Spring 2005 and Autumn 2007 sections were involved in experiments on individual versus group work on practice exams; 3) course enrollment varied widely, from 173 students per section to 700; 4) the exam number changed in Autumn 2007 due to increased enrollment—two 100-point exams, spaced several weeks apart, replacing a comprehensive, 200-point final. The short-answer format for exams remained the same, however.
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Elements of course design</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Socratic lecturing</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Ungraded active learning</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clickers</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Practice exams</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reading quizzes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Class notes summaries</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>In-class group exercises</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exams</td>
<td>Two 100-point midterms, 200-point comprehensive final</td>
<td>Two 100-point midterms, 200-point comprehensive final</td>
<td>Two 100-point midterms, 200-point comprehensive final</td>
<td>Two 100-point midterms, 200-point comprehensive final</td>
<td>Two 100-point midterms, 200-point comprehensive final</td>
<td>Four 100-point exams</td>
</tr>
<tr>
<td>Total course points</td>
<td>550</td>
<td>550</td>
<td>720; 620</td>
<td>720</td>
<td>793</td>
<td>741</td>
</tr>
</tbody>
</table>

*The sections were taught back-to-back, with identical lecture notes. They took similar or identical midterms and an identical final exam.

*bOne section answered questions with clickers; one section answered identical questions with cards (see Freeman et al., 2007). Card responses were not graded.

*cIn one section, clicker questions were graded for participation only; in one section, identical clicker questions were graded right/wrong (see Freeman et al., 2007).

*dAt random, half the students did practice exams individually; half did the same exercise in a four-person group structured by predicted grade.

*eAll students did practice exams individually.
Computing Final Grades
Across the introductory biology series at UW, instructors strive to implement the following system for computing final grades: 1) total course points are summed for each student, 2) students with the top 5% of total course points earn a 4.0, 3) the instructor sets a level for passing the course (receiving a 0.7, and thus course credit toward graduation) that typically represents ~50% of total course points and 40–45% of total exam points, and 4) the range of course points between the 0.7 and 4.0 cutoffs is divided into equal bins and assigned 0.8, 0.9, 1.0, and so on, up to 3.9.

During the years included in this study, students had to receive a course grade of at least 1.5 on a 4.0 scale to register for the next course in the series. Thus, we define failure as a final course grade of <1.5. Except when noted, the analyses reported here include only those students who received a grade—meaning that students who dropped the course before the final exam were not included. The data sets also excluded a small number of students who had been caught cheating.

Exam Equivalence across Quarters
In a longitudinal study that evaluates changes in failure rates, it is critical to test the hypothesis that changes in failure rates were due to changes in exam difficulty. In the Freeman et al. (2007) study, this hypothesis was tested by comparing student performance on an identical midterm in two quarters that differed in course design. Because we wanted to avoid giving an identical exam again, we created two methods for evaluating exam equivalence across the six quarters.

Weighted Bloom’s Index. Bloom’s taxonomy of learning (Bloom et al., 1956; Krathwohl, 2002) identifies six levels of understanding on any topic. Bloom’s framework has been applied in an array of contexts in undergraduate biology education (Crowe et al., 2008), including characterizing exams (Zheng et al., 2008). We used the six levels to create a Weighted Bloom’s Index, which summarizes the average Bloom’s level of exam questions weighted by the points possible:

\[ \text{Weighted Bloom’s Index} = \left( \frac{\sum_{i=1}^{n} P \cdot B}{T \cdot 6} \right) \times 100, \]

where \( n \) is the number of questions, \( P \) is points/question, \( B \) = Bloom’s rank (1–6) for that question, \( T \) is total points possible, and 6 is the maximum Bloom’s score. To help interpret the index, note that Level 1 and 2 questions test lower-order cognitive skills, Level 3–6 questions assess higher-order cognitive skills (Bloom et al., 1956; Krathwohl et al., 2002) and specific Weighted Bloom’s Index values (Figure 1) conform to each Bloom’s level, as follows:

16.7 is an exam with exclusively recall questions (Bloom’s Level 1);
33.3 is an exam with exclusively conceptual understanding questions (Bloom’s Level 2);
50.0 is an exam with exclusively application questions (Bloom’s Level 3);
66.7 is an exam with exclusively analysis questions (Bloom’s Level 4);
83.3 is an exam with exclusively synthesis questions (Bloom’s Level 5); and
100 is an exam with exclusively evaluation questions (Bloom’s Level 6).

We calculated a Weighted Bloom’s Index for every exam in the study by recruiting three experienced TAs to assign a Bloom’s level to every exam question given in each quarter. The raters were trained in “Blooming” exams by one of the authors (M.P.W.) and assessed questions that were presented in a common format and in random order. Although raters knew that they were assessing Biology 180 exam questions, they were blind to the quarter and year. They were also blind to the study’s intent and to the hypothesis being tested by the ratings. Point values for each question were omitted to avoid any bias introduced by high- versus low-point-value questions. Because multipart questions are common on these exams, each rater gave a Bloom’s rating to each question part. Each rater assessed 295 exam questions and assigned a total of 724 Bloom’s levels. (Including the identical exam, there were a total of 310 exam questions and 750 Bloom’s rankings in the study.)

We used the decision rules published by Zheng et al. (2008) to arrive at a single Bloom’s level for each question part. Briefly, the three raters achieved consensus on 53 rankings during “norming sessions,” where questions were discussed as a group after being rated individually. None of the subsequent questions were discussed as a group to resolve conflicts in ratings. Instead, we assigned the consensus rating when all three raters agreed and the majority-rule rating when two of three raters agreed. 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. To assess the degree of agreement among the multiple raters, we calculated Krippendorff’s alpha—an appropriate measure for ordinal coding data from multiple raters (Hayes and Krippendorff, 2007)—and the intra-class \( r \).

The Weighted Bloom’s Index should accurately summarize the average Bloom’s ranking of exams, facilitating comparisons across courses or even institutions. In addition, it should contain information on exam difficulty because students typically perform much better on lower-level versus higher-level questions (e.g., Knecht, 2001; Freeman and Parks,

Vol. 10, Highlights of 2011
Predicted Exam Score. As an alternative method for assessing exam difficulty, we created a Predicted Exam Score (PES) by recruiting three experienced TAs—different from the individuals who did the Bloom’s ratings—to predict the average number of points that students would receive on each part of each exam question in the study. These individuals were experienced graders: In addition to running labs and attending every class, each TA in Biology 180 spends ~40 h grading exams. Thus, the PES raters judged the difficulty of exam questions based on 1) an understanding of what had been presented in the textbook and introduced in class and 2) extensive grading experience that made them alert to wording, context, level, and other issues that cause confusion or difficulty. All three PES raters were peer TAs—senior undergraduate biology majors. We hypothesized that peer TAs might be more attuned to how undergraduates read and respond to exam questions than faculty are.

The PES raters used the same randomized list of 295 identically formatted exam questions as did the Bloom’s raters, except that point values for each of the 724 question parts were indicated. Like the Bloom’s raters, the PES raters were blind to the study’s intent and the hypothesis being tested with the predicted-points data.

Work on the PES began with a norming session. This meeting started with each of the three individuals assigning a predicted-average-points value to 25 questions—a total of 58 question-parts—on his or her own. The three then met to discuss their predictions for average points on each question-part until they arrived at a consensus value. Subsequent questions were assessed individually but not discussed. To arrive at a single predicted-average-points value for each of the exam-question parts after the norming session, we computed the arithmetic average of the three ratings submitted.

The goal of the PES was to test the hypothesis that changes in failure rates were due to changes in the difficulty of exams, independent of changes in course design. The metric should be useful because it is reported in units of average expected points on exams and because exam points predict most of the variation in final grade (see Results).

Note that because the Weighted Bloom’s Index and the PES were computed from data on all exam questions in the study, there was no sampling involved. As a result, there is no sample variance on the data reported here, and it was not possible to use statistical approaches to assess “significance” when comparing data from different exams. In cases like this, significance is a judgment about the relevance of observed differences to the hypothesis being tested.

Exam Impact on Final Course Grade

A proponent of low-structure course designs could argue that failure rates decline in highly structured courses not because of increased student learning, but because the “practice” exercises (reading quizzes, clicker questions, etc.) are easier than actual exam questions and thus inflate student performance. In addition, even in the most highly structured course designs, our intent was for exams to remain the primary determinant of final course grade—because exams represent the most controlled type of assessment possible and because exams are the major type of assessment in subsequent courses.

To test the point-inflation hypothesis, we performed simple linear regressions with each student’s total exam points in each quarter as the predictor variable and his or her final grade as the response variable. In addition, we computed the 1.5 (failing) cutoff predicted by the regressions. These 1.5 cutoffs represented the average total exam points required to progress in the major, each quarter.

If exams are the major determinant of final grades irrespective of degree of course structure—even when many nonexam points are possible—then $R^2$ values for the regressions should be uniformly high, the slopes and intercepts of the regression lines should be indistinguishable, and the 1.5 cutoffs should be similar across quarters. It was not possible to use analysis of covariance (ANCOVA) to test for heterogeneity of regression slopes and intercepts across quarters because final grades were not normally distributed (Shapiro-Wilks normality test $W = 0.9361, p \ll 0.001$). Rather, the best-fit distribution followed a binomial error distribution. Consequently, we used analysis of deviance to formally test the hypothesis that the slopes and intercepts of the regressions did not vary across quarters, using a Generalized Linear Model (GLM) framework (Crawley, 2007, p. 516). More specifically, we compared the fit among three linear models 1) incorporating only the covariate exam points, 2) additionally incorporating quarter as a fixed effect, and 3) the full model with an interaction term. If the additional explanatory variable and the interaction term failed to improve the fit of the GLM, it provides assurance that slopes and intercepts are homogeneous across quarters (are not significantly different, using a likelihood ratio test [LRT]).

Student Equivalence across Quarters

In a longitudinal study of student performance in courses, it is critical to test the hypothesis that changes in failure rates are due to changes in the academic ability and preparedness of the student population at the time of entering each course. To test this hypothesis, we used the Predicted Grade model introduced by Freeman et al. (2007). Briefly, we predicted a final course grade for each student in each quarter by using a regression model based on UW grade point average (GPA) and SAT-verbal scores (Predicted grade = (0.00291 × SATverbal) + (1.134 × UWGPA) – 2.663). Students who lacked an SAT-verbal score were assigned the average SAT-verbal score from that quarter; on average, 10% of students had a missing SAT-verbal score each quarter. Because the percentage of missing values was low and the loading of SAT-verbal in the regression model is small, the missing values should have a minimal impact on the analysis. In addition, analysis of variance (ANOVA) showed no heterogeneity among quarters in the UW GPAs of students with missing SAT-verbal scores ($F = 0.46, df = 5, 228, p = 0.81$). Thus, the substitution for missing values should not bias tests for differences in predicted grade across quarters. Because UW GPA has a large impact in the regression model, grades of students who lacked a UW GPA at the start of the course were dropped from the predicted grade analysis.

We performed linear regressions to compare the Predicted Grade with Actual Course Grade in each quarter of the study, and then used analysis of deviance to assess the robustness
Table 3. Exam equivalence analyses

<table>
<thead>
<tr>
<th></th>
<th>Discussed-consensus</th>
<th>All three agree</th>
<th>Two of three agree</th>
<th>Sequential ratings</th>
<th>Nonsequential ratings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percentage of total ratings</td>
<td>7.3</td>
<td>26.4</td>
<td>51.1</td>
<td>9.9</td>
<td>5.2</td>
</tr>
</tbody>
</table>

"Discussed-consensus" means that questions were rated independently and then discussed to reach a consensus; "Independent ratings" were not discussed among raters. "Sequential ratings" were questions that received three ratings that differed by one Bloom’s level (e.g., a 2, 3, and 4); “Nonsequential ratings” were questions that received three ratings that differed by more than one Bloom’s level (e.g., a 2, 3, and 5).

b. Weighted Bloom Indices

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Midterm 1</td>
<td>50.8</td>
<td>48.5</td>
<td>45.3</td>
<td>58.9</td>
<td>54.4</td>
<td>53.7</td>
</tr>
<tr>
<td>Midterm 2</td>
<td>36.1</td>
<td>51.6</td>
<td>51.6</td>
<td>46.8</td>
<td>50.8</td>
<td>54.7</td>
</tr>
<tr>
<td>Midterm 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>50.2</td>
</tr>
<tr>
<td>Final (or Midterm 4)</td>
<td>48.1</td>
<td>54.3</td>
<td>45.4</td>
<td>51.6</td>
<td>51.6</td>
<td>55.3</td>
</tr>
<tr>
<td>Course average</td>
<td>45.8</td>
<td>52.1</td>
<td>46.9</td>
<td>52.2</td>
<td>52.1</td>
<td>53.5</td>
</tr>
</tbody>
</table>

c. PES values (predicted percent correct)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Midterm 1</td>
<td>70.8</td>
<td>73.0</td>
<td>71.9</td>
<td>67.8</td>
<td>64.9</td>
<td>66.0</td>
</tr>
<tr>
<td>Midterm 2</td>
<td>73.0</td>
<td>68.0</td>
<td>68.0</td>
<td>72.1</td>
<td>67.7</td>
<td>67.0</td>
</tr>
<tr>
<td>Midterm 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>68.5</td>
</tr>
<tr>
<td>Final (or Midterm 4)</td>
<td>69.4</td>
<td>70.0</td>
<td>71.8</td>
<td>71.0</td>
<td>69.6</td>
<td>68.6</td>
</tr>
<tr>
<td>Course average</td>
<td>70.6</td>
<td>70.2</td>
<td>70.9</td>
<td>70.5</td>
<td>68.0</td>
<td>67.5</td>
</tr>
</tbody>
</table>

a. Identical exams.
b. Course averages were computed from data on all exam questions from that quarter. (They are not the averages of the indices from each exam.)

RESULTS

Exam Equivalence

The TAs who ranked questions on Bloom’s taxonomy showed a high level of agreement. At least two of the three raters agreed on the identical Bloom’s level for 85% of the questions rated. Using the original values assigned for the norming session questions—not the consensus values—the Krippendorf’s alpha among raters was 0.481 for the entire data set; the intra-class $r$ was 0.481 ($F = 4.07, df = 668, p \ll 0.0001$).

There was also a high level of agreement among the three TAs who evaluated exam questions for the PES values. Using the original values assigned for the norming session questions—not the consensus values—the intra-class $r$ for the entire data set was 0.84 ($F = 24.3, df = 725, p \ll 0.0001$).

There is a strong association between the Weighted Bloom’s Indices (Table 3b) and the PES values (Table 3c): The $r$ for the 18 exams in the study (excluding one iteration of the repeated exam) is –0.72, and a regression analysis shows that variation in the Weighted Bloom’s Index explains 51% of the variation in the PES (Figure 2; $F = 16.7, df = 1,16, p < 0.001$).

Exam Impact on Course Grade

$R^2$ values indicate that, even in the most highly structured versions of the course, when exam points represented as little as 55% of the total course points, total exam points still explained 89% of the total variation in final grade (Table 4a). Across quarters in the study, the regressions of total exam points on final grade were similar, and the average number of exam points required to get a 1.5 varied little—the range was only nine points (Table 4b). In addition, analysis of deviance (Table 4b) indicated no statistically significant variation among quarters in how well total exam points predicted final grade.
Table 4. Regression analyses: Total exam points as a predictor of final course grade. Data from the two sections in Spring 2005 were analyzed separately because the clickers and card sections in that quarter (see Materials and Methods and Table 2) had different total course points and thus a different scale for computing final grade.

a. Regression statistics

<table>
<thead>
<tr>
<th></th>
<th>Low structure</th>
<th>Moderate structure</th>
<th>High structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adjusted $R^2$</td>
<td>0.96</td>
<td>0.95</td>
<td>0.97</td>
</tr>
<tr>
<td>$y$-intercept</td>
<td>$-2.29^{a}$</td>
<td>$-2.35^{a}$</td>
<td>$-2.12$</td>
</tr>
<tr>
<td>$\beta$</td>
<td>0.0184$^{b}$</td>
<td>0.0186$^{b}$</td>
<td>0.0172</td>
</tr>
<tr>
<td>1.5 cutoff predicted by regression</td>
<td>206.4</td>
<td>206.6</td>
<td>210.1</td>
</tr>
<tr>
<td>$n$</td>
<td>323</td>
<td>335</td>
<td>174</td>
</tr>
</tbody>
</table>

b. ANCOVA fit by GLMs incorporating the effect of exam points, quarter, and the interaction term exam points by quarter; the response variable is actual grade (GLM with binomial error distribution). Analysis of deviance shows that slope and intercept do not significantly vary across quarter.

<table>
<thead>
<tr>
<th>Model</th>
<th>Residual df</th>
<th>Residual Deviance</th>
<th>df</th>
<th>Deviance</th>
<th>$\chi^2$</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exam points only</td>
<td>2303</td>
<td>82.840</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exam points + quarter</td>
<td>2298</td>
<td>79.366</td>
<td>5</td>
<td>3.473</td>
<td>0.6274</td>
<td></td>
</tr>
<tr>
<td>Exam points × quarter</td>
<td>2293</td>
<td>78.475</td>
<td>5</td>
<td>0.892</td>
<td>0.9708</td>
<td></td>
</tr>
</tbody>
</table>

Values have 95% confidence intervals that overlap.

bValues have 95% confidence intervals that overlap.

Student Equivalence across Quarters

Across all six quarters of the study, there was a strong relationship between the Predicted Grade and actual grade for each student. In regression analyses, Predicted Grade explained between 51 and 64% of the variation in final grade; slopes ranged from 0.88 to 1.13 (complete data not shown). Furthermore, analysis of deviance indicates no significant differences among the regressions of Predicted Grade on actual grade across quarters (Table 5). These data support the earlier claim by Freeman et al. (2007) that the Predicted Grade model is a robust index of student ability and preparedness across quarters and years.

An ANOVA indicates significant heterogeneity in Predicted Grades across quarters (Table 6; $F = 13.2$, $df = 5$, $2351$, $p < 0.001$). Post-hoc Tukey’s Honestly Significant Difference tests demonstrate that this heterogeneity was distributed across quarters in the study: Of the 15 pairwise tests, only six were not significantly different, and these crossed levels of course structure (data not shown).

Evaluating the Drop Rate

Among five quarters from Spring 2002 through Autumn 2007, drop rates varied from 1.8 to 3.6%. A $\chi^2$ analysis indicates no heterogeneity in drop rate as course design changed from low to medium to high structure ($\chi^2 = 3.9$, $df = 4$; $p = 0.42$). Significant heterogeneity occurs when the drop rate of 6.1% in Autumn 2009 is added to the analysis, however ($\chi^2 = 21.5$, $df = 5$; $p < 0.001$). This increase is probably due to 1) a change in enrollment from 350 to 700, and 2) a change in departmental policy that loosened restrictions on repeating the course.

Changes in Failure Rate in Biology 180

A $\chi^2$ analysis of the number of students who took the final and received a grade, and were above and below the 1.5 threshold, confirms significant differences across quarters (Table 7; $\chi^2 = 44.7$, $df = 5$; $p < 0.001$).

This result is confounded, however, by changes in student academic ability across quarters, reported earlier in the text. To control for student academic characteristics, we constructed a generalized linear mixed-model (GLMM) to test the hypothesis that level of course structure plays a significant role in explaining the proportion of students failing each quarter. Specifically, we analyzed the decline in proportion of students failing as a function of those predicted to fail in

Figure 2. Weighted Bloom’s Indices and PES values are negatively correlated. The Weighted Bloom’s Index summarizes the average Bloom’s level per point on an exam; the PES summarizes expert-grader predictions for average points that a class will receive on an exam. Regression statistics are reported in the text.
Table 5. Robustness of the Predicted Grade model ANCOVA fit by GLMs incorporating the effect of predicted grade, quarter, and the interaction term predicted grade by quarter; the response variable is actual grade (GLM with binomial error distribution). Analysis of deviance shows that slope and intercept do not significantly vary across quarter.

<table>
<thead>
<tr>
<th>Model</th>
<th>Residual df</th>
<th>Residual deviance</th>
<th>df</th>
<th>Deviance</th>
<th>χ²</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predicted grade only</td>
<td>2276</td>
<td>306.616</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Predicted grade + quarter</td>
<td>2271</td>
<td>298.432</td>
<td>5</td>
<td>8.184</td>
<td>0.1464</td>
<td></td>
</tr>
<tr>
<td>Predicted grade × quarter</td>
<td>2266</td>
<td>297.386</td>
<td>5</td>
<td>1.047</td>
<td>0.9587</td>
<td></td>
</tr>
</tbody>
</table>

Table 6. Average predicted grades across quarters

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean ± SD</td>
<td>2.46 ± 0.72</td>
<td>2.57 ± 0.73</td>
<td>2.64 ± 0.70</td>
<td>2.67 ± 0.60</td>
<td>2.85 ± 0.66</td>
<td>2.70 ± 0.61</td>
</tr>
<tr>
<td>n</td>
<td>327</td>
<td>338</td>
<td>334</td>
<td>328</td>
<td>339</td>
<td>691</td>
</tr>
</tbody>
</table>

DISCUSSION

The negative association between the Weighted Bloom’s Indices and the PES values supports the claim that questions that are higher on Bloom’s taxonomy of learning are harder, and both methods of assessing exam equivalence suggest that exam difficulty increased in highly structured versions of the course (Table 3, b and c). The 2–3% drop in PES in the highly structured courses, relative to the PES values in the low- and medium-structure courses, represents 8–12 total exam points over the quarter—enough to drop students’ final grades 0.1–0.2, on average, in the highly structured courses. This difference, and the uptick in the Weighted Bloom’s Index in Autumn 2009, supports the instructors’ impression—that he wrote harder exams because the reading quizzes implemented in those quarters had already presented lower-level questions.

It is important to note that the exams analyzed here appear rigorous. Although the Weighted Bloom’s Index and PES will become more informative if and when values are reported for

Table 7. Failure rates across quarters

<table>
<thead>
<tr>
<th></th>
<th>Low structure</th>
<th>Moderate structure</th>
<th>High structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percentage of students &lt; 1.5</td>
<td>18.2</td>
<td>15.8</td>
<td>10.9</td>
</tr>
<tr>
<td>n</td>
<td>324</td>
<td>333</td>
<td>330</td>
</tr>
</tbody>
</table>

Table 8. MMI: Models and comparison criteria. Best-fit models are recognized by 1) the conservative LRT p value of the lowest AIC model or 2) ΔAIC > 2. Note that the LRTs are hierarchical: The p value reported on each row is from a test comparing the model in that row with the model in the row below it.

<table>
<thead>
<tr>
<th>Model</th>
<th>df</th>
<th>AIC</th>
<th>ΔAIC</th>
<th>ω</th>
<th>logLikelihood</th>
<th>LRT (p value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Structure + predicted</td>
<td>5</td>
<td>1610.27</td>
<td></td>
<td>0.47</td>
<td>−800.14</td>
<td>0.027</td>
</tr>
<tr>
<td>Predicted</td>
<td>3</td>
<td>1613.5</td>
<td>3.22</td>
<td>0.09</td>
<td>−803.75</td>
<td>0.098</td>
</tr>
<tr>
<td>Structure × predicted</td>
<td>7</td>
<td>1613.68</td>
<td>3.4</td>
<td>0.43</td>
<td>−799.84</td>
<td>2.2e-16</td>
</tr>
<tr>
<td>Structure</td>
<td>4</td>
<td>1903.27</td>
<td>293</td>
<td>0</td>
<td>−947.64</td>
<td>0.0015</td>
</tr>
<tr>
<td>Null</td>
<td>2</td>
<td>1912.3</td>
<td>302.02</td>
<td>0.01</td>
<td>−954.15</td>
<td></td>
</tr>
</tbody>
</table>
other courses, instructors, and institutions, the data indicate that an average Biology 180 exam question in this study is at the application level of Bloom’s taxonomy (Table 3b).

Finally, the data reported in Table 4 support the hypothesis that the large number of “practice points” available in highly structured versions of the course did not inflate grades, and thus did not affect changes in the failure rate.

Although student academic ability and preparedness varied among quarters, it did not vary systematically with changes in course structure. The results of the GLMM, which controlled for heterogeneity in student preparedness and formative assessment, and extensive active learning (no lecturing) and formative assessment, respectively. The difference between the proportion of students predicted to fail and the actual proportion failing decreases with increasing structure (GLMM, binomial error $n = 2267$, $p = 0.06$, **$p = 0.0004$).

Thus, the data presented here support the hypothesis that increasing course structure can help reduce failure rates in an introductory biology course—from 18.2 to 6.3% (Table 7). They are consistent with data from other STEM disciplines suggesting that intensive use of active-learning exercises can help capable but underprepared students succeed in gateway courses (e.g., Beichner et al., 2007), and identify essential components of successful, highly structured course designs—were much lower than failure rates for gateway courses at many other institutions and other STEM disciplines (Table 1). We propose that this pattern is due to enrollments in Biology 180 consisting primarily of sophomores who had already completed a three-quarter, introductory chemistry sequence for majors. On our campus, the initial course in the chemistry series enrolls ~2500 students annually—approximately half of a typical freshman class. Only ~1500 students per year take the last course in the introductory chemistry series, however. Thus, it is likely that many under-prepared students who might have taken Biology 180 were ineligible, due to a failure to complete the chemistry prerequisite.

Would highly structured course designs reduce failure rates more or less if most students were freshmen—before they had been “screened” by a chemistry prerequisite? The experiment to answer this question is underway. Our department has now removed the chemistry prerequisite for Biology 180—a change that became effective in Winter 2010—and recent quarters have enrolled up to 50% first-year students. It will be interesting to test whether highly structured course designs analyzed here have an impact on this increasingly younger student population.

The Role of Reading Quizzes
If the benefit of highly structured courses is to help students gain higher-order cognitive skills, what role do reading quizzes play? By design, these exercises focus on Levels 1 and 2 of Bloom’s taxonomy—where active learning may not help. We concur with the originators of reading quizzes (Crouch and Mazur, 2001): Their purpose is to free time in class for active learning exercises that challenge students to apply concepts, analyze data, propose experimental designs, or evaluate conflicting pieces of evidence.

As a result, reading quizzes solve one of the standard objections to active learning—that content coverage has to be drastically reduced. Reading quizzes shift the burden of learning the “easy stuff”—the vocabulary and basic ideas—to the student. The premise is that this information can be acquired by reading and quizzing as well as it is by listening to a lecture.

Without reading quizzes or other structured exercises that focus on acquiring information, it is not likely that informal-group, in-class activities or peer instruction with clickers will be maximally effective. This is because Bloom’s taxonomy is hierarchical (Bloom et al., 1956). It is not possible to work at the application or analysis level without knowing the basic vocabulary and concepts. We see reading quizzes as an essential component of successful, highly structured course designs.

New Tools for Assessing Exam Equivalence
This study introduces two new methods for assessing the equivalence of exams across quarters or courses: the Weighted Bloom’s Index based on Bloom’s taxonomy of learning and the PES based on predictions of average
performance made by experienced graders. These approaches add to the existing array of techniques for controlling for exam difficulty in STEM education research, including use of identical exams (Mazur, 1997; Freeman et al., 2007); concept inventories or other standardized, third-party tests (e.g., Hestenes et al., 1992); and isomorphic or “formally equivalent” questions (e.g., Smith et al., 2009).

The Weighted Bloom’s Index also has the potential to quantify the degree to which various courses test students on higher-order cognitive skills. In addition to assessing Weighted Bloom’s Indices for similar courses across institutions, it would be interesting to compare Weighted Bloom’s Indices at different course levels at the same institution—to test the hypothesis that upper-division courses primarily assess the higher-order thinking skills required for success in graduate school, professional school, or the workplace.

The analyses reported here were designed to control for the effects of variation in the instructors, students, and assessments. More remains to be done to develop techniques for evaluating exam equivalence and student equivalence. With adequate controls in place, however, discipline-based research in STEM education has the potential to identify course designs that benefit an increasingly diverse undergraduate population. In the case reported here, failure rates were reduced by a factor of three. If further research confirms the efficacy of highly structured course designs in reducing failure rates in gateway courses, the promise of educational democracy may come a few steps closer to being fulfilled.

ACKNOWLEDGMENTS

David Hurley wrote and supported the practice exam software, Tiffany Kwan did proof-of-concept work on the PES, John Parks and Matthew Cunningham helped organize the raw data on course performance from the six quarters, and Janneke Hille Ris Lambers contributed ideas to the statistical analyses. Bloom’s taxonomy rankings for the Weighted Bloom’s Indices were done by Micah Horwith, Hannah Jordt, and Alan Sunada. PES ratings were done by Eunice Lau, Ignatius Lau, and Chelsea Mann. This work was funded by a grant from the University of Washington’s College of Arts and Sciences as part of the Foundations Course program, and by a grant from the University of Washington-Howard Hughes Medical Institute Undergraduate Science Education Programs (Grant 52003841). D.H. was supported by a National Science Foundation Predoctoral Fellowship. The analysis was conducted under Human Subjects Division Review #38125.

REFERENCES

ACT (2006). Reading Between the Lines: What the ACT Reveals About College Readiness in Reading, Iowa City.


Bloom BS, Engelhart MD, Furst EJ, Hill WH, Krathwohl DR (1956). Taxonomy of Educational Objectives: The Classification of Educa-


Vol. 10, Highlights of 2011 73


Culturally Relevant Inquiry-Based Laboratory Module Implementations in Upper-Division Genetics and Cell Biology Teaching Laboratories

Dimuth Siritunga, María Montero-Rojas, Katherine Carrero, Gladys Toro, Ana Vélez, and Franklin A. Carrero-Martínez

Department of Biology, University of Puerto Rico, Mayagüez, PR 00680

Submitted April 11, 2011; Revised June 28, 2011; Accepted June 28, 2011
Monitoring Editor: Robert L. DeHaan

Today, more minority students are entering undergraduate programs than ever before, but they earn only 6% of all science or engineering PhDs awarded in the United States. Many studies suggest that hands-on research activities enhance students’ interest in pursuing a research career. In this paper, we present a model for the implementation of laboratory research in the undergraduate teaching laboratory using a culturally relevant approach to engage students. Laboratory modules were implemented in upper-division genetics and cell biology courses using cassava as the central theme. Students were asked to bring cassava samples from their respective towns, which allowed them to compare their field-collected samples against known lineages from agricultural stations at the end of the implementation. Assessment of content and learning perceptions revealed that our novel approach allowed students to learn while engaged in characterizing Puerto Rican cassava. In two semesters, based on the percentage of students who answered correctly in the premodule assessment for content knowledge, there was an overall improvement of 66% and 55% at the end in the genetics course and 24% and 15% in the cell biology course. Our proposed pedagogical model enhances students’ professional competitiveness by providing students with valuable research skills as they work on a problem to which they can relate.

INTRODUCTION

Allowing undergraduate students to feel the excitement and self-investment that is related to the discovery of new knowledge is now thought to be a better way to educate students in science (National Research Council, 2000, 2003; Lord et al., 2007; Prince and Felder, 2007; Shaffer et al., 2010). Laboratory modules that foster research-oriented or inquiry-based exercises are replacing or supplementing the traditional “cookbook”-style lab modules in many disciplines. Multi-week research or inquiry-based laboratory exercises have been shown to enhance critical thinking, retention of knowledge, technical skills gained, and ability to interpret data, as well as to increase students’ interest in science and improve their preparation for postgraduate education (Hathaway et al., 2002; Seymour et al., 2004; Hunter et al., 2007; Russell et al., 2007). One key obstacle when implementing true research-oriented laboratory modules is the availability of tested procedures/techniques; this becomes particularly problematic when the lab exercises are designed for students in very large classes.

In this paper, we describe the development of laboratory modules for a general genetics course (~600 students per year) and a cell biology course (~200 students per year), which, respectively, primarily enroll junior- and senior-level students. The main goal of our modules was to incorporate culturally relevant content and high-quality, research-oriented instruction in undergraduate courses, while increasing the students’ ability to understand and use modern
molecular biology tools. Because cassava (Manihot esculenta) is an integral part of the diet of Puerto Ricans (18.4 million metric tons were consumed in 2007 in the form of mofongo, pasteles, fritas, chips, or boiled cassava), we decided to implement modules based on the study of cassava diversity in Puerto Rico.

Module implementation strategies varied in order to enhance both students' understanding of the science behind the module and hands-on experience. Implementation in the genetics course was spread throughout the semester within four sessions, while the cell biology module was completed in two sessions. This strategy was based mostly on the fact that students generally take the genetics course before the cell biology course.

Our goal for the modules was to allow the untrained undergraduate students to experience the excitement and challenges of original research, while enhancing their understanding of the multidisciplinary nature of modern experimentation. Within this primary goal, we also were attempting to increase the students' ability to learn about and use modern tools of molecular and cellular biology. Students in both courses were asked to bring cassava samples from their hometowns. These samples were assigned unique identifiers and analyzed in the genetics course. To assess genetic diversity of samples brought from the field, students in the genetics course used molecular markers, while students in the cell biology course assessed starch content and visualized other cellular structures using light microscopy. Our specific objectives were divided into research and educational objectives:

**Education Objectives**
- Equip biology students with experience in handling modern molecular and cellular techniques and equipment within the scope of an original research project
- Promote students' ability to search for solutions to common scientific problems within a team research effort
- Incorporate and institutionalize modern laboratory exercises into the current course schedule
- Update the laboratory manuals in the genetics and cell biology courses
- Develop skills, such as critical thinking, awareness of contemporary issues, and problem solving, in students

**Research Objectives**
- Relate the unknown cassava accessions in Puerto Rico to the varieties of the germplasm collection
- Identify cassava varieties not in the germplasm collection
- Identify and clarify duplicate nomenclature
- Identify the type of starch—amylose and amylpectin—of known cassava varieties
- Identify the type and number of cells in known cassava varieties

**Research Project Background**
Knowledge of germplasm diversity and its genetic characterization is an invaluable asset in crop improvement strategies, as well as in conservation strategies. Though cassava has been an important crop and present in the Caribbean since the fifteenth century, the origin of today's cassava in the Caribbean is poorly understood. Evaluation of the Puerto Rican cassava population for its diversity and genetic characterization is an invaluable asset in the improvement strategies of cassava. Due to the continuation of traditional farming practices in Puerto Rico, evaluation of its cassava population can potentially lead to the discovery of new varieties. Since maintenance of cassava varieties is conducted in the field by farmers, correct assessment of genetic diversity of those farmer-held varieties can be invaluable for the crop's conservation and the identification of new combinations with maximum genetic variability. These new varieties can then be used for further selection and introgression of desirable genes from diverse germplasm into the available genetic base (Smith, 1984; Cox et al., 1986; Mohammad and Prasanna, 2003).

A number of molecular methods, such as restriction fragment length polymorphism, random amplification of polymorphic DNA, amplified fragment length polymorphism, single nucleotide polymorphism, and single sequence repeat (SSR) markers (also known as single tandem repeat or microsatellites) have been used to study cassava diversity (Fregene et al., 2003; Kizito et al., 2005). Overall, SSR molecular markers have been the method of choice, since they can be easily adapted for classification and identification of cassava, and are particularly useful for studying the variation in allelic frequency of unlinked loci, which is the preferred way of assessing genetic differentiation. They also exhibit high levels of polymorphisms, are somatically stable, are inherited in a codominant Mendelian manner, and are conducive to automation (Morgante and Olivieri, 1993; Fregene et al., 2003). Fregene et al. (2003) developed a set of SSR markers that are useful for diversity analysis in cassava. Using 33 of these SSR markers, we attempted to assess the diversity of cassava in Puerto Rico. To accomplish this research goal we obtained cassava leaf samples from different townships in Puerto Rico with the help of the ~600 genetics students. Undergraduate students brought in a sufficient number of samples (n = 162) to permit a thorough assessment of diversity of cassava in Puerto Rico. The same samples were also used in the laboratory module for the genetics course, as described below in DNA Extraction, by undergraduate students enrolled in all teaching laboratory sections. This generated a strong sense of student ownership of the work and responsibility for the data, and also led to more enthusiasm for the laboratory exercises. Furthermore, the student analysis of these samples resulted in a recent publication (Montero-Rojas et al., 2011) and a graduate thesis project.

**Course and Lab Module Context**
The first part of the implementation took place in the genetics course, which normally includes ~30% sophomores, 40% juniors, and 30% seniors. The students who enroll in this course are not only from the Department of Biology but also from other departments, such as Agronomy and Soils, Horticulture, Food Technology, and Engineering. The course format was 2 h of lecture and 3 h of laboratory per week for 15 wk. The lecture component of the genetics course was taught by one of four professors, and the number of students per classroom varied between 30 and 60 students. The laboratory component consisted of sections of 18–24 students, and was taught by graduate teaching assistants (TAs) of diverse backgrounds. Because the lecture section of the course covers a broad range of topics, ranging from classical Mendelian
genetics to molecular genetics, the lab exercises proceeded independently of the lecture section of the course.

The multi-week genetics laboratory module began on week 1 and ended on week 15. The new lab exercises replaced more “cook-book”-style, stand-alone exercises. Comparatively, the new lab exercises mimicked an actual research project. To achieve the final goal, the students had to work continuously on the same project over the course of several weeks. Throughout the genetics module, the students worked in groups of four, and the students were introduced to the cassava module in week 1.

The second part of the implementation took part in the cell biology course (1% sophomores, 21% juniors, 78% seniors). Traditionally, most students (~80%) follow the suggested curricular sequence, where the genetics course is taken the semester before cell biology. Thus, the majority of students were already familiar with the project, as they were likely to have completed the genetic characterization of cassava. The purpose of the cell biology component was to allow students to gather quantitative and qualitative data from cassava samples. Prior to this module implementation, no plant cellular structures were studied as part of the cell biology laboratory, even though they are covered in the lecture component of the course.

Both modules were initially test-piloted in one laboratory section (~22 students), and compared with a control laboratory section (~22 students) taught by the same TA. The rationale behind this was to minimize any TA-related impact. Furthermore, the TA chosen for the pilot implementation had prior experience teaching the traditional genetics or cell biology laboratory.

METHODS

Genetics Module

Sample Collection. The key objective when collecting cassava leaf samples from different townships in Puerto Rico was to make sure that the samples were itemized coherently. With ~600 students per year enrolling in the genetics course, a clear and resourceful “sample collection form” was prepared (see Supplemental Material). This “sample collection form” was part of a handout developed for this module and was handed to each student during the first week of laboratories. The main objective here was to be able to backtrack any particular sample to the farm from which it originated. Thus, key facts, such as the location of and contact information of the owner of the cassava plant was collected. Furthermore, to ease the archiving of all the samples in the principal investigator’s (PI’s) research laboratory, an internal coding system was utilized. For example, Fa10-066-1 would represent sample 1 of the lab section 066 during the Fall 2010 semester. Each undergraduate student was asked to collect cassava leaf samples from his or her township; the samples were placed in ziplock plastic bags, which were stored at 4°C until their transportation to the genetics laboratory.

DNA Extraction. During the semester prior to the pilot implementation of the lab module, the DNA extraction protocol described by Dellaporta et al. (1983) was modified to fit 1 h of lab time (see Supplemental Material). An undergraduate research student and a senior graduate student in the PI’s research laboratory were assigned the task of modifying the protocol. Once optimized to fit a 1-h time duration, the modified protocol was subsequently tested in the research laboratory by two untrained undergraduate students enrolled in the genetics course during that particular semester under the supervision of the trained undergraduate research student.

During the pilot and full implementation of this lab module, each group of four students extracted DNA from a sample brought by one of its members. With a lab section having four to five groups, DNA from at least four cassava samples was extracted per section. Any unused leaf material was returned back to the marked ziplock bag, and subsequently analyzed in the research laboratory as part of the master’s degree thesis of a graduate student, who subjected the samples to a more thorough evaluation using 33 SSR markers (see Assessment of Cassava Diversity in Puerto Rico).

SSR Marker Amplification. Due to the vast number of students enrolled in the genetics course per semester, it was impossible to facilitate each student preparing a polymerase chain reaction (PCR) by adding each component individually. Therefore, the lab coordinator prepared three master mixes prior to the laboratory exercise, and each student added 8 μl of Mix1 (MgCl2/dNTPs/H2O), 5 μl of Mix2 (Reaction buffer/Taq polymerase/H2O), 5 μl of Mix3 (M13 primers/SSR primers/TE buffer), and 7 μl of cassava DNA to obtain a final PCR containing 10 mM Tris-HCl (pH 8.3), 1 U of Taq polymerase, 2 mM MgCl2, 0.2 mM dNTP mix, 0.5 pmoles of the M13 primer, and 100 pmoles of the SSR primers.

Each group was given DNA samples from the other groups of the lab section to amplify with a single SSR marker. Thus, a particular group would set up four to five PCRs with a single SSR marker. Once all the groups had set up their respective PCRs, the following amplification cycle was utilized: 95°C for 5 min, followed by 34 cycles of 94°C for 30 s, 55°C for 45 s, and 72°C for 1 min, with a final extension of 5 min at 72°C.

Gel Electrophoresis. Though the scoring of SSR markers cannot be performed on normal agarose gel electrophoresis, size difference of some alleles can be visualized if the difference is more than 10–15 base pairs. In addition, the difference between a homozygote sample and heterozygote sample at a particular SSR locus can be observed on an agarose gel. For these reasons, as well as to allow the students the opportunity to gain experience utilizing DNA gel electrophoresis, each group made a standard 1.5% agarose gel, and ran their PCR samples along with a molecular marker sample.

SSR Marker Evaluation. The addition of the M13 tail sequence to the 5′ end of each forward SSR primer allows the amplicons to be visualized by the fluorescence emitted by the “fluorophore” bound to the M13 primer (third primer added to the PCR). For better resolution, the amplicons were visualized on 6.5% denaturing polyacrylamide gels on a LI-COR automated DNA sequencer. The molecular weight of each band was assessed by running a 50–350 base pair molecular-size ladder (LI-COR Biosciences, Lincoln, NE) in each gel. Though the lab technician performed this step, the logistics of the technique were shared with the undergraduate students as part of the lab module. The results of the
SSR marker evaluation on the denaturing polyacrylamide gels were shared with the undergraduate students in a subsequent lab session.

Cell Biology Module

Light Microscopy. Students were also asked to view root cuttings, which helped in visualization of the cellular and physiological significance of cell-specific nutrient storage. Students cut thin sections of selected root portions and stained them with potassium iodine, followed by several washes to remove excess stain. Samples were then mounted for light microscopy visualization. Students also viewed leaf structures under the microscope. The modified procedure asked the student to cut 5 mm × 5 mm pieces of leaf, which were then fixed in paraformaldehyde/glutaraldehyde solution. Following incubation and washes in phosphate-buffered saline (PBS), the leaf cuttings were mounted and observed. These results were made available to the PI’s research laboratory, which determined whether the samples varied from the unknown cassava samples.

Starch Assessment. On the basis of a modified protocol (Cabral and Carvalho, 2001), students cut 0.5 g of cassava and macerated it into a paste that was then suspended in 80% ethanol and centrifuged. Lugol’s solution was added to the sample, and a dilution series was performed. Absorbance was measured at 480 nm using a Genesis 10 UV Scanning spectrophotometer (Thermo Electron Corporation, Waltham, MA; Hovenkamp-Hermelink et al., 1988). Students were then asked to use the Beer-Lambert’s equation to determine the starch concentration present in their cassava samples compared with known controls. As in the genetics module, a student handout was developed for the cell biology module.

Assessment Tools

Assessment tools focused on evaluating two important aspects: gain of content knowledge and impact on reported self-confidence. The gain-of-content knowledge assessment tool consisted of questions both broad and specific in nature. For example, “What is the use of RNase in nucleotide extraction?” was used to assess specific knowledge gained, while “Why is cassava an important crop in this world?” was used to assess broad knowledge gained. To minimize students’ tendency to guess, all questions in both pre- and postmodule tests had “I don’t know” as an option. In addition, the postmodule assessment also contained open-ended questions. These questions allow for the opportunity to provide relevant comments regarding their experiences with the modules. As part of the assessment cycle, these suggestions were evaluated and incorporated wherever appropriate.

During the first two semesters of implementation, the TAs and the lab coordinator/technician were also assessed, using a similar assessment tool (unpublished data). More importantly, the postmodule assessment also contained questions that gave the TAs, lab coordinator, and lab technician a chance to comment on the content and implementation of the new lab exercise. The faculty member in charge of the genetics and cell biology courses met with laboratory coordinators, technicians, and TAs before, during, and after each implementation. This allowed for timely intervention, leading to successful implementation of the modules. These interactions with the others involved in the implementation of the modules enhanced our capacity to improve the next implementation cycle.

RESULTS AND DISCUSSION

Sample Collection

During the course of implementations (one pilot and two full implementations) we received 162 cassava leaf samples collected by students enrolled in the genetics course at the University of Puerto Rico Mayagüez campus. We developed a sample collection form as part of the modules; this form was made available to the students at the beginning of each semester. From these forms, we determined that collected samples came from 59% of Puerto Rican municipalities, with higher representation from areas where most of the agricultural land is located. Specifically, a higher number of samples (67) came from the northwestern region, which has the most arable, well-drained soil suitable for cassava growth. Additionally, the majority of the samples (95%) were collected from home gardens or subsistence farmers. This is to be expected, due to the lack of significant local commercial cassava production and the fact that ~90% of the cassava consumed in Puerto Rico is imported (Goenaga et al., 2002). The strategy of engaging students in large-scale sample collection allowed us to properly analyze the diversity present in Puerto Rican cassava (see Assessment of Cassava Diversity in Puerto Rico).

DNA Extraction, SSR Marker Amplification, and Gel Electrophoresis

Each group of students performed their own DNA extraction, set up PCRs with one of the SSR markers (their own plus other groups’ DNA samples), and analyzed the banding pattern on agarose gel electrophoresis. Figure 1A shows the quality of the DNA extracted from cassava leaves by the students of a single lab section (five groups). After completion of the DNA extraction, 2 μl of the sample was separated by gel electrophoresis for visualization of genomic DNA. During the three semesters, almost all sections were able to extract DNA from cassava leaves using the modified Dellaporta et al. (1983) method described previously. Approximately 75% of the time, the fluorescent PCRs set up by individual students succeeded. With each group setting up four to five PCRs, every student had the opportunity to set up the PCR for at least one sample. Figure 1B shows gel electrophoresis of such PCR amplicons, including the molecular marker ladder. The gels were made and run completely by the undergraduate students. On completion of the gels, the students were asked to assess the gels and state the likely sizes of the bands, the homozygous versus heterozygous nature of the samples, and the allelic differences that existed between samples.

In this technique, the allelic differences or similarities between DNA samples are best visualized through the fluorescence emitted by the M13 primer; this was accomplished by the separation of the amplicons on 6.5% denaturing polyacrylamide gels on a LI-COR 4300 automated DNA sequencer (LI-COR Biosciences). This step was performed by the lab technician, due to the cost of the machine and the
requirement of trained personnel to operate the software and the equipment. Each lab section was brought in front of the LI-COR 4300, and the process of electrophoresis was explained and pertinent information was provided verbally. After completion of the electrophoresis, each group of students was provided the gel image to analyze the allelic diversity for the five cassava samples they ran based on one SSR marker (Figure 1C). Subsequently, all the student groups in a particular lab section evaluated one another’s samples. Thus, the students were able to identify and understand the differences between SSR markers, as well as the difference between cassava samples based on each SSR marker.

**Starch Quantification and Stomatal and Root Observations**

Each group of students (two students per group) performed its own cassava root sample preparation for the starch-quantification portion of the exercise. Each student group was able to complete sample preparation and preparation of a dilution series for analysis using Beer-Lambert’s equation to determine the starch concentration present in their cassava samples, using known standards as controls. On completion of their analysis, students were asked to compare their results with other groups in search of differences. To enhance the laboratory experience and maximize available time (3 h maximum), each TA received a handout before the laboratory session and a set of slides with specific time-management instructions. For instance, during long incubation steps, each student performed other module-related activities, such as leaf surface area estimates. These measurements were collected and subsequently analyzed in the research lab and correlated with relevant variables (growth conditions, soil type, etc.). Even though no significant trend emerged from the data, it provided an opportunity for students to apply math concepts in the context of the module.

**Assessment of Student Learning**

The objective of these laboratory module implementations was to introduce students to research and allow them to meaningfully contribute to an actual research project. We used a culturally relevant problem to engage the largest possible number of students in this activity. In this context, we aimed to teach 800+ undergraduate students the principles and techniques for proper sample collection, DNA extraction, PCR, SSR molecular marker assessment, gel electrophoresis, and basic light microscopy, as well as quantitative and qualitative starch assessment. Through hands-on practical experience, we also sought to improve students’ confidence in learning, understanding, and performing experimental inquiry-based science. Knowledge and confidence tools were used to assess students’ learning and confidence building before and after module implementation. The gain-of-knowledge assessment tool was initially used during the pilot implementation in the genetics course. Taught by the same TA, the students in the pilot section (including the new cassava module) showed higher gain of knowledge between the pre- and postmodule assessments compared with the control section (excluding the new cassava module; see Supplemental Material). In both sections, the overall content knowledge in the premodule assessment was ∼42%. In the postmodule assessments, the students in the pilot section reported an overall content knowledge increase up to 86%, while the control section students reported only a modest increase up to 54%. In the cell biology pilot implementation, due to a miscommunication with the TA involved, proper assessment could not be completed. During the subsequent full-implementation semesters, 284 and 264 students in the genetics course, as well as 99 and 120 students in the cell biology course, were exposed to the cassava laboratory modules described here. Previous exposure
to theoretical aspects of cell and molecular biology offered in introductory biology courses explains the high premodule knowledge. For instance, in the premodule assessment for the genetics course administered during the two semesters of implementation, 50% and 75% of the students reported that they understood the definition of molecular biology (GCQ2), while 73% and 75% reported having prior knowledge in the uses of gel electrophoresis (GCQ3; Figure 2A). Students enrolled in the cell biology module reported 62.5% and 60% knowledge of “cell biology” (CCQ2) and 80% and 78.8% of techniques associated with the field (CCQ4), before each semesters’ module implementations (Figure 2B). However, there was a slight postmodule increase in GCQ2 to 69% and 80%, and in GCQ3 to 80% and 92%, during the two semesters (Figure 2A). Similar results were observed for the cell biology module (Figure 2B; CCQ2 and CCQ4). In general, during the first full implementation of the genetics module, premodule assessment tools reported >50% of students chose the correct answer for only one out of seven questions (Figure 2A), while postmodule assessment showed that >50% of students in that cohort correctly answered six out of the seven questions. During the second implementation semester, >50% of students correctly answered three of seven questions prior to the module, and improved to seven out of seven after module implementation (Figure 2A). Similar results were observed with students enrolled in the cell biology module (Figure 2B).

Overall, there was a significant increase in students’ ability to correctly answer postmodule content questions compared with premodule questions. This was particularly true for questions that were more specific to the module (Figure 2A; GCQ5–GCQ7) for students enrolled in the genetics course. In the cell biology module, we did not see a trend of greater improvement in the module-specific questions compared with general questions but observed improvement in both categories of questions during the semesters (Figure 2B). Based on our assessments, the student handout was revised, and a new TA handout and presentation were created.

Normalizing the results of the postmodule assessment for each question, based on the percentage of students who answered correctly in the premodule assessment, shows an overall improvement of 66% and 55% in the first and second semesters, respectively, in the genetics course (Figure 2C). There was an improvement in all the questions, with GCQ7 (“What is the use of RNase in nucleotide extractions?”) showing the highest increase of 155% and 177% during the first and second semesters, respectively. This translates to having >2.5 times more students learning and understanding the use of RNase in molecular biology experiments, compared with the beginning of the semester. A high percentage of student improvement was also observed through GCQ5 (“Which of the following techniques is not associated with molecular biology?”) and GCQ6 (“What is the enzyme used in the amplification of DNA in PCR technique?”), with the number of students answering correctly almost or actually doubling in both semesters. Students in the cell biology module showed similar improvement, though not as strongly as in the genetics course. In particular, students showed increased learning and understanding in specific module knowledge, for instance “What is the compound used in starch quantification?” (Figure 2D; CCQ5). After normalizing the results based on the premodule assessment, we found students in the cell biology course had an overall improvement of 24% and 15% in the two semesters.

One of the disappointments was the student response to question GCQ1 (“Why is cassava an important crop in this world?”; Figure 2A). Cassava being an important crop in Third World countries, and also being part of local cuisine, we expected a higher correct response rate in both pre- and the postmodule assessment tools. Even though there was a twofold and a 1.5-fold increase in the number of genetics students answering correctly in the two semesters (Figure 2C), only 50% and 57% of the students answered this question correctly in the postmodule (Figure 2A). In the cell biology course, assessment tools report a similar trend (twofold and 1.75-fold increases, respectively; Figure 2D). Nevertheless, a higher proportion of students (91% and 63%) answered correctly in the cell biology postmodule compared with their performance in the genetics course (Figure 2B). A possible explanation could be related to the TAs’ discussion and emphasis of the overall importance of cassava. Thus, in the subsequent implementations, more emphasis was given to the overall value of cassava as a source of food and nutrition during the end-of-semester, TA-led discussion. Interestingly, the fact that most students take genetics prior to cell biology may explain the higher correct response rate in the latter course.

Assessment of Student Confidence

To assess whether our module had an impact on student confidence, we developed a postmodule survey using a 5-point Likert scale (Figure 3). When students were asked whether they were confident “designing experiments to test a hypothesis” in both a cell (CPQ2) and genetics (GPQ3) laboratory, our assessment tools report that roughly 50% of the students report themselves as “very confident” or “confident.” With the exception of that question, ~60% or more students in the genetics course perceived themselves as “very confident” or “confident” in all other categories (Figure 3A). In the cell biology module, the majority of students (60% or more) reported high confidence in all questions, except those related to designing experiments and scheduling activities related to those experiments. In those questions, 34.4% (CPQ2) and 41.3% (CPQ3) self-reported as “very confident” or “confident,” respectively.

Given the large number of students enrolled in both courses, and considering their diverse backgrounds, we believe that the self-reported confidence gain met our expectations. Of significance is that students in both cell and genetics laboratories report an increased confidence in their abilities to “properly and safely use lab equipment” (>85% and >73%, in GPQ5 and CPQ4, respectively; Figure 3) and “observe and collect data” (>90% and >72%, in GPQ6 and CPQ5, respectively; Figure 3) through the implemented modules. From a perspective of cell and molecular biology training, 68–72% of students enrolled in the genetics teaching laboratory agreed they are “very confident” or “confident” in “learning theoretical aspects of modern molecular biology techniques” (GPQ9), while 65–69% students in cell biology reported similar results in terms of cell biology techniques (CPQ7). When students in the genetics laboratory were asked whether they were “very confident” or “confident” in “learning the practical uses of
Laboratory Modules of Cultural Relevance

### Figure 2.

Gains in content knowledge. Assessment of knowledge gained by students through pre- and postmodule assessments consisting of seven questions. Percentage correct per question in the premodule (gray bars) and postmodule (black bars) assessments during Fall 2009 and Spring 2010 semester in the genetics (A) and cell biology (B) courses. The last horizontal bars show the overall percentages. Percentage increase of the number of students obtaining the correct answer per question as a result of the implementation of the lab module (postmodule data normalized for the premodule data) in the genetics (C) and cell biology (D) courses. GCQ and CCQ refer to "genetics content question" and "cell biology content question," respectively. The overall scores in the posttests were higher (*p < 0.01; **p < 0.05; ***p < 0.1) as determined using an unpaired t test. (Continued)

#### (A) Genetics Content Questions

<table>
<thead>
<tr>
<th>Question</th>
<th>Pre-Module</th>
<th>Post-Module</th>
</tr>
</thead>
<tbody>
<tr>
<td>Why is cassava an important crop in this world?</td>
<td>Fall09</td>
<td>Sp10</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>What is molecular biology?</td>
<td>Fall09</td>
<td>Sp10</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Which of the following can be accomplished by gel electrophoresis?</td>
<td>Fall09</td>
<td>Sp10</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Which of the following research projects require molecular biology?</td>
<td>Fall09</td>
<td>Sp10</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Which of the following technique is not associated with molecular biology?</td>
<td>Fall09</td>
<td>Sp10</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>What is the enzyme used in the amplification of DNA in PCR technique?</td>
<td>Fall09</td>
<td>Sp10</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>What is the use of RNase in nucleotide extraction?</td>
<td>Fall09</td>
<td>Sp10</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fall09 semester n=284; Sp10 semester n=264

#### (B) Cell Biology Content Questions

<table>
<thead>
<tr>
<th>Question</th>
<th>Pre-Module</th>
<th>Post-Module</th>
</tr>
</thead>
<tbody>
<tr>
<td>Why is cassava an important crop in this world?</td>
<td>Fall09</td>
<td>Sp10</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>What is cell biology?</td>
<td>Fall09</td>
<td>Sp10</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Which of the following is associated with cell biology?</td>
<td>Fall09</td>
<td>Sp10</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Which of the following is correct regarding cell biology techniques?</td>
<td>Fall09</td>
<td>Sp10</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>What is the compound used in starch quantification?</td>
<td>Fall09</td>
<td>Sp10</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Which of the following would you use to determine starch concentration?</td>
<td>Fall09</td>
<td>Sp10</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Which of the following can be done with basic microscopy techniques?</td>
<td>Fall09</td>
<td>Sp10</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fall09 semester n=99; Sp10 semester n=120

(The diagram is not fully visible, but it shows bar graphs for each question with percentages marked.)
molecular biology techniques in research projects” (GPQ10) and “performing well in a scientific research project using PCR” (GPQ13) as a result of their participation in the lab module, 72–74% and 65–66%, respectively, agreed (Figure 3A). Students in cell biology reported similar perception gains (Figure 3B). As a result of using cassava as the platform in this lab module, 61% and 63% of the students during the two semesters agreed that they are “very confident” or “confident” in “comprehending topics in agricultural biotechnology” (GPQ11), which is an emerging field of immense importance (Figure 3A).

**Assessment of Cassava Diversity in Puerto Rico**

Through this lab module, we were able to achieve our research objective by thoroughly assessing the genetic diversity of cassava in Puerto Rico. This is the first comprehensive evaluation of cassava in the Caribbean region. The undergraduate
Gains in confidence. Assessment of self-confidence postmodule measured using a 5-point Likert scale (5 = very confident, 4 = confident, 3 = somewhat confident, 2 = not confident, and 1 = not at all confident) in genetics (A) and cell biology (B). GPQ and CPQ refer to “genetics perception question” and “cell biology perception question,” respectively. (Continued)

Students in the genetics course used five SSR markers with the highest polymorphism information content value in their analysis. The same samples were subjected to a more vigorous assessment of diversity using 33 SSR markers as part of a thesis project of a graduate student in the PI’s research laboratory (Montero-Rojas et al., 2011). These samples of unknown genetic background were also compared with the Puerto Rican cassava germplasm, which consists of 23 different cultivars and accessions (Montero-Rojas et al., 2011).

Beyond achieving both our research and education objectives, our assessments demonstrate that our module implementation had a positive impact on students. Our approach differs significantly from the traditional “recipe-based” teaching laboratory, in which students can easily anticipate expected results, and moves to engage them in a research-oriented experience. It allows the student to be actively engaged from the very beginning (sample collection) to the end, in a laboratory sequence that spans two courses over two
semesters and uses a variety of scientific tools to analyze various aspects of a research question. Thus, this learning experience closely resembles the multidisciplinary nature of modern research. The selection of cassava, a common food source in Puerto Rico, as a model to develop and implement these modules further enhances student understanding of the experience, as it relates to a part of their daily activities. When thoughtfully conceived, local adaptations provide models for increased relevance to students, regardless of age and background (Fields, 2010). This is of particular relevance amid calls for increasing diversity in science, technology, engineering, and mathematics (STEM) careers (National Academy of Sciences, 2010; National Institute of General Medical Sciences, 2011), in particular for Latinos (Malcom et al., 2010), and its benefits to academia (Smith, 1997). We believe our modules represent a successful strategy for increasing participation of underrepresented minority students in STEM careers.

ACKNOWLEDGMENTS

We thank Lorraine Rodríguez-Bonilla and Jean Seda-López for help in developing the laboratory modules. Special thanks also to Terry Woodin (National Science Foundation, Arlington, VA) for critical review of this manuscript. We also thank the TAs in the genetics and cell biology courses during Spring 2009, Fall 2009, and Spring 2010 semesters for their efforts and willingness to participate in the implementation of these modules. This work was funded by a grant from the Course Curriculum and Laboratory Improvement program of the National Science Foundation (DUE-736727).

REFERENCES


Shaffer CD et al. (2010). The genomics education partnership: successful integration of research into laboratory classes at a diverse group of undergraduate institutions. CBE Life Sci Educ 9, 55–69.


Evaluating the Role of Faculty with Education Specialties in Science Education

Seth D. Bush,†† Nancy J. Pelaez,†‡ James A. Rudd,†§ Michael T. Stevens,†|| Kimberly D. Tanner,†¶ and Kathy S. Williams†#

†California Polytechnic State University, San Luis Obispo, CA 93407; ‡Purdue University, West Lafayette, IN 47907; §California State University, Los Angeles, CA 90032; ||Utah Valley University, Orem, UT 84058; ¶San Francisco State University, San Francisco, CA 94132; and #San Diego State University, San Diego, CA 92182

Submitted August 26, 2010; Revised November 22, 2010; Accepted November 23, 2010
Monitoring Editor: Deborah Allen

Efforts to improve science education include university science departments hiring Science Faculty with Education Specialties (SFES), scientists who take on specialized roles in science education within their discipline. Although these positions have existed for decades and may be growing more common, few reports have investigated the SFES approach to improving science education. We present comprehensive data on the SFES in the California State University (CSU) system, the largest university system in the United States. We found that CSU SFES were engaged in three key arenas including K–12 science education, undergraduate science education, and discipline-based science education research. As such, CSU SFES appeared to be well-positioned to have an impact on science education from within science departments. However, there appeared to be a lack of clarity and agreement about the purpose of these SFES positions. In addition, formal training in science education among CSU SFES was limited. Although over 75% of CSU SFES were fulfilled by their teaching, scholarship, and service, our results revealed that almost 40% of CSU SFES were seriously considering leaving their positions. Our data suggest that science departments would likely benefit from explicit discussions about the role of SFES and strategies for supporting their professional activities.

INTRODUCTION

In the United States there is strong and growing interest in improving science education (National Academy of Sciences [NAS], 2005; NAS et al., 2007; National Research Council, 2007; National Science Board, 2007; U.S. Congress House of Representatives, 2007). Three arenas in science education are particularly key: 1) the preparation and support of sufficient numbers of quality K–12 science teachers, 2) the improvement of undergraduate science education, and 3) the expansion of the science education research base in specific science disciplines. Most scientists who are situated in a college or university are engaged in undergraduate teaching, either in classrooms or laboratories, and as such are continually called to join in undergraduate science education reform. In addition, the call for scientists to be involved in K–12 science education has been apparent since the 1980s in the following ways: within professional scientific organizations (e.g., American Association for the Advancement of Science, 1993), as part of school university partnership programs (e.g., Alberts, 1994; Bower, 2001; Pelaez and Gonzalez, 2002; Dolan and Tanner, 2005), in the context of graduate training programs (e.g., National Science Foundation [NSF] GK–12), within science departments (e.g., NSF, 1996; McWilliam et al., 2001; NSF GK–12, 2003).
Investigation of SFES

2008; Woodin et al., 2010), and at the level of broader impacts within individual NSF research grants. Basic research scientists are apparently being asked to engage in discussions between the sciences and a range of disciplines informing the field of science education, such as cognitive science, developmental psychology, cultural diversity, and education. However, science faculty largely lack formal training in the teaching and learning of their discipline, deep knowledge of the culture and parameters of K–12 schools, and/or professional incentives to strongly engage science education. Perhaps unsurprisingly, attempts to involve already busy science faculty in additional science education efforts—outside of their formal training and research focus and without the professional reward structure associated with their basic research efforts—have been challenging and have met with limited impact (Sunal et al., 2001; Walczyk et al., 2007). Compounding the challenges in K–12 and undergraduate science education, there is an additional need for more research on how students learn science within specific science disciplines and on the effectiveness of science teaching strategies in those disciplines at all educational levels.

One emergent approach to stimulating change in all science education arenas is the seeding of university science departments with Science Faculty with Education Specialties (SFES), scientists who take on specialized roles within science education in their discipline (American Physical Society, 1999; Bush et al., 2008, 2010). Although a plethora of innovative curricular and pedagogical approaches to science education have been developed and investigated, the translation and implementation of these findings have been weak (Woodin et al., 2010). Inclusion of a science education specialist in science departments may provide a human bridge between the often isolated efforts in science and in education. SFES appear to be nucleating science education activities on campuses outside of Colleges of Education and may indicate a jurisdictional shift in science education, with science departments explicitly taking on the improvement of science education as part of their charge. The hypothesized impact of SFES includes, but is not limited to, increased articulation between K–12 and undergraduate science education, support for faculty development and nucleation of pedagogical innovation in undergraduate science, and research on teaching and learning specific to a science discipline within that discipline itself. Even though SFES occupy a pivotal role at the interface of key arenas in science education, there has been little formal discussion or systematic investigation of the purposes, structures, or outcomes of these SFES positions or strategies for the hiring, retention, or promotion of SFES (California State University, 2006; Bush et al., 2006, 2008, 2010). As such, SFES appear to be a phenomenon of national interest that is understudied and surrounded by assumptions that are untested.

The investment by science departments in a faculty-level academic position focused on science education—with its accompanying status and intellectual freedom—is substantial and has been described. In 2006, we reported on the challenges associated with hiring SFES (Bush et al., 2006). There we presented a hiring guide for departments interested in SFES and for SFES looking for employment. In 2008, we presented preliminary findings about SFES in the nation’s largest university system (annual enrollment ~450,000), the 23-campus California State University (CSU) system. We found that SFES were present throughout the CSU and included two distinct subpopulations: those specifically hired as SFES and those who transitioned to SFES roles from their initial faculty roles (Bush et al., 2008). Strikingly, we found that nearly 40% of CSU SFES were “seriously considering leaving” their current positions.

Here we more thoroughly report findings from our research on CSU SFES. The CSU’s primary mission is undergraduate and master’s-level graduate education, including K–12 teacher education. CSU undergraduates come from the top one-third of their high school graduating classes (University of California Office of the President, 2007). The 23 campuses include institutions that differ substantially in their founding dates, settings, student populations, enrollment sizes, and levels of research orientation, and as such our findings are potentially predictive of the characteristics of the SFES model at a variety of institution types.

The purpose of this study was to identify the extent to which SFES exist in the largest university system in the United States, as well as to examine the nature of SFES professional activities and SFES perceptions of their specialized positions. Results are intended to aid a broad audience of stakeholders—including higher education administrators, state and national policy makers, funding agencies, science departments in colleges and universities, and individual scientists considering SFES career pathways—in conceptualizing, structuring, and supporting SFES positions. Results from investigating the CSU SFES phenomenon may prove useful for framing discussions about the purpose of SFES positions, their potential impact on science education from within science departments, and strategies for maximizing the SFES impact.

METHODS

A survey instrument was designed to collect information about SFES demographics, position structure, and other issues, such as what SFES are doing and perceptions of how SFES positions are structured. In addition, the instrument collected attitudinal information relevant to SFES perceptions of job expectations relative to non-SFES peers, issues of professional satisfaction, pathways to SFES positions, and other information that is primarily of a descriptive nature (e.g., hire date, nature of formal training). SFES professional activities were probed with respect to teaching, scholarly activity, and service since this framework is used in the evaluation of CSU faculty for retention, tenure, and promotion. Although scholarly activity can be broadly defined, use here is in accordance with the CSU definition that includes research, scholarly, and creative activities. As part of face validation, the survey was piloted using non-CSU faculty. Preliminary research results from this study have already been published (Bush et al., 2008). This study constitutes a descriptive study of CSU SFES and is not intended to be a direct comparison of SFES and non-SFES science faculty. However, some survey questions asked SFES to consider their experiences relative to non-SFES in their department.

SFES were identified for this study by 1) soliciting names of potential SFES from CSU College of Science Deans, 2) examining all CSU science department websites in search of SFES profiles, and 3) prompting initial survey respondents to provide names of additional SFES on their campuses. A total
of 156 CSU faculty were invited to complete a 111-question, online survey, and 103 of the invitees responded to the survey between December 2007 and January 2008 (66% response rate). We used snowball sampling. The initial survey respondents provided 66 names of likely SFES, of which only 7 had not yet been invited, suggesting that our SFES search was comprehensive. Research participants represented 20 of the 23 campuses and received a $20 gift card in compensation for their participation.

Data were collected anonymously, such that individual responses were not associated with a particular CSU campus. Surveys that were incomplete (n = 12), not submitted by tenure/tenure-track science faculty (n = 10), or lacking informed consent (n = 3) were excluded from this analysis. Of the remaining 78 survey respondents, 59 individuals self-identified as SFES, whereas 19 self-identified as not SFES. Analyses presented in this paper are based on data from the 59 individuals who self-identified as SFES. The only exception appears later in Figure 13, which includes data from the 19 faculty who self-identified as not SFES. In a previous publication of preliminary findings, we excluded individuals located in science education centers from analyses, but in this report, we have included these individuals as part of all SFES. We have not displayed their disaggregated data because of their low number (n = 2).

Results from the survey instrument are generally reported as descriptive statistics, and comparisons are not statistically significant unless explicitly stated. When appropriate, Pearson χ² tests (independence and goodness-of-fit) were used to compare subpopulations of SFES (e.g., Biology SFES versus SFES from other science disciplines, or faculty considering leaving their position versus those intending to stay) and to infer at the p < 0.05-level that these differences in subpopulations would likely be found at comparable institutions (Plackett, 1983). Yates's correction was used for the χ² test when the contingency table involved cells with small numbers to prevent overestimation of statistical significance (Yates, 1934).

Open-ended survey questions were analyzed using grounded theory analysis (Glaser and Strauss, 1967). Two researchers examined all responses for each open-ended question, determined emergent themes independently, and then agreed upon a common set of thematic coding categories. Each researcher independently coded responses into these categories and calculated a percentage of respondents who offered evidence in each category. Categories presented in the results are those that included comments coded from more than 10% of respondents. Interrater reliability (IRR) was calculated by dividing the number of scoring agreements by the total number of scoring decisions. The number of scoring decisions was calculated by multiplying the number of respondents for each question by the number of thematic coding categories.

Although this CSU SFES study provides new and valuable information about SFES, a limitation is that the study was confined to SFES from one U.S. state university system, largely composed of BS- and MS-granting institutions. However, these results can be used as a basis for discussing the assumptions about faculty-level academic positions focused on science education within the science disciplines.

RESULTS

The results and figures are organized with a description and display of data for all SFES (including Biology, Chemistry, Geoscience, and Physics SFES) followed by disaggregation by discipline department. The disaggregation was provided because of the different histories of the emergence of SFES in each discipline and because we anticipated readers may be interested in SFES in a particular discipline.

Who Are SFES?

SFES (n = 59 respondents) were found in tenure/tenure-track faculty positions in the variety of institution types in the CSU and across a range of science departments, including Biology.
Investigation of SFES

Figure 2. SFES hire date, rank, and tenure status in relation to SFES status. Reported hire date (A), current rank (B), and tenure status (C) prior to becoming SFES for all SFES and disaggregated by science discipline.

SFES in our study had hire dates from 1970 through 2007 (Figure 2A). The greatest number of SFES were recent hires, hired between 2000 and 2007. This pattern was observed across all departments except Geoscience, where the greatest numbers of SFES were instead hired between 1990 and 1999. SFES represented tenure-track positions across all faculty ranks (31% assistant, 32% associate, 37% full professors; Figure 2B). The proportions of full professors were similar across departments. Biology and Geoscience departments tended to have smaller proportions of assistant professors and higher proportions of associate professors than did Chemistry and Physics departments. The majority (71%) of SFES as a group did not have tenure when they adopted their roles as SFES (Figure 2C). This pattern was similar across the disciplines (Biology 65%, Chemistry 79%, Geoscience 75%, Physics 73%).

(n = 20), Chemistry (n = 14), Geoscience (n = 8), Physics (n = 15), and science education centers in Colleges of Science (n = 2). In each figure in the Results, the n-value represents the actual number of respondents. Roughly equal numbers of female (48%) and male (52%) SFES responded to our survey. However, there were some differences in gender distributions across departments (Figure 1A). Approximately two-thirds (65%) of SFES responding from Biology departments were female, whereas over half of SFES responding from Chemistry, Geoscience, and Physics departments were male (57%, 57%, and 77%, respectively). In terms of age distribution, SFES in our study spanned a range of age categories (Figure 1B). The two youngest age categories combined (30–39- and 40–49-yr-olds) comprised a majority (65%) of the overall sample and represented over half of each discipline subsample (Biology 60%, Chemistry 57%, Geoscience 86%, and Physics 72%).
What Do SFES Do?

As a population, SFES appeared to be engaged in a variety of teaching, scholarly, and service activities rather than specializing in one of these areas. With regard to teaching, most SFES in our sample reported teaching courses for majors [lower division (78%), upper division (73%), and/or electives (51%)] and nonmajors (86%) (Figure 3A). Smaller proportions of SFES (15%) taught courses for teaching credential students, whereas over 50% taught precredential courses. As scholars, over half of SFES reported seeking funding to support science education research (58%), basic science research (61%), curriculum development (59%), and K–12 teacher development (68%) (Figure 3B). A smaller proportion of SFES reported seeking funding to support university teacher development (29%). All SFES reported involvement in departmental service with 92% serving Colleges of Science and 49% providing service for Colleges of Education (Figure 3C). Typically, the respondents in our sample reported being engaged in service activities for their university (80%) and involved in teacher preparation/recruitment (71%), in-service teacher support (76%), and assessment (88%).

SFES in the CSU procured external funding to support their professional activities. Over 40% of SFES in our sample had garnered more than $500,000 in their current position (Figure 4). Over half of Biology and Geoscience SFES had obtained over $500,000 in grant funding (55% and 63%, respectively), whereas smaller proportions of Chemistry and Physics SFES had procured that level of funding (21% and 20%, respectively).

SFES activities appeared to reflect their reasons for adopting an SFES role. The survey asked respondents to elaborate on their response choices in several open-ended questions. Table 1 describes six categories of reasons offered by respondents for taking an SFES position (n = 56, IRR = 97%). The categories that included responses from more than one-third of the respondents were their interest in science education across the three arenas (48%), specific nature of the faculty position (36%), and flexibility in balancing teaching and...
Table 1. Reasons for taking an SFES position in response to the following question: “Briefly, what were your original reasons for taking your current position?” (n = 56)

<table>
<thead>
<tr>
<th>Category</th>
<th>Sample quotes</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interest in science education</td>
<td>Influencing K–12 science education</td>
<td>48</td>
</tr>
<tr>
<td></td>
<td>I was interested and excited about the possibility of teaching teachers (both pre-service and in-service).</td>
<td></td>
</tr>
<tr>
<td></td>
<td>I wanted to positively impact the K–12 education system in the state of California.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>As a new parent, and seeing the state of science education in the country, I felt that it was imperative that people with backgrounds in science have an impact on the K–12 education process.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Improving college and university science teaching</td>
<td></td>
</tr>
<tr>
<td></td>
<td>I was excited about the opportunity to become involved in the curricular changes that were planned by my department, and I had an interest in science education/curricular innovation/assessment.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>I was interested in trying to encourage reform of university-level science teaching.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>I was excited to join a growing department where I would be able to shape the program development of the Physics major.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Engaging in science education research</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Interest in having a laboratory of my own where I could return to conducting science education research.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Desire to work in a Chemistry department and conduct research in chemical education.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ability to do Physics education–related research.</td>
<td></td>
</tr>
<tr>
<td>Specific nature of the faculty position</td>
<td>Geographic location of institution</td>
<td>36</td>
</tr>
<tr>
<td></td>
<td>I wanted to move my family back to California to be close to grandparents.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>I was looking for a job in this geographical area because of personal reasons.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Location in California is close to my field area (basic scientific research).</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Characteristics of student population</td>
<td></td>
</tr>
<tr>
<td></td>
<td>I was attracted strongly to working at an institution that serves a large population of ethnic minorities and first-time college students because of my personal background.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>I wanted to work at a CSU because I am committed to public education and access.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Reputation of the institution</td>
<td></td>
</tr>
<tr>
<td></td>
<td>The opportunity to teach and engage in research and scholarly activities at my alma mater.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>I already felt a sense of community and ability to relate to the student body.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>I wanted to be affiliated with an excellent institution and department with programs that I respected.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Flexibility in balancing teaching and research</td>
<td>36</td>
</tr>
<tr>
<td></td>
<td>I wanted a university position that was an even balance of teaching and research. The available position at this university offered that compromise.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>The position melded all my areas of expertise, genetics, teaching, science education, etc.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>I was interested in doing science teaching and research where both would be valued.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Desire to teach at the undergraduate level</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>After teaching at the high school level for several years, I wanted to teach at the university level.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>The reason I took my position in the first place was that I wanted to teach undergraduate science.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>I love science and wanted to transmit my passion for it through teaching.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>I wanted to change the way students think about the world around them, teaching, and learning.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Collegial environment, sometimes including presence of other SFES</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>Attracted to the collegial nature of faculty in my department and college.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Having another SFES faculty member in the department (and others across the college) was another important consideration.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[The institution] offered a strong, collegial department with a commitment to teaching, basic research, and K–12 teacher prep that allows me to pursue all my interests quite freely.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>It was rare to find a place that said that the science education specialist (SFES) would be treated just like a non-SFES in terms of research, teaching, and service.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Need for a job or more job security</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>I really needed a job! ...I viewed it as an opportunity to obtain tenure, without having to uproot my family.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>I was employed here as a lecturer and wanted a more stable position... and the security a tenure track position carried with it that I didn’t have as a Lecturer.</td>
<td></td>
</tr>
</tbody>
</table>

Investigation of SFES

research (36%). Additional reasons included a desire to teach at the undergraduate level (23%), the presence of a collegial environment, sometimes including other SFES (18%), or simply a need for a job or more job security (11%).

What Is Their Professional Training?

Virtually all SFES in all disciplines had formal training in science, whereas small percentages of Biology and Chemistry SFES did not (5% and 7%, respectively). Most SFES over all disciplines (88%) had earned science PhDs, whereas many had completed science postdocs (37%) and/or science master’s degrees (48%). The patterns of formal training across the disciplines were similar.

In stark contrast, only 32% of SFES overall reported having any type of formal postbaccalaureate training in science education (Figure 5C). Chemistry SFES had the greatest (43%) proportion of faculty with any formal science education training; proportions for Biology (30%), Geoscience (25%), and
Figure 5. SFES professional training. 
Pie graphs describe the proportions of SFES with any formal postbaccalaureate training in science (A) and science education (C). Bar graphs describe the types of formal postbaccalaureate training SFES report in science (B) and science education (D) for all SFES and disaggregated by science discipline.

Figure 6. SFES perceptions of time spent on professional activities compared with non-SFES. Perceptions of teaching (A), scholarship (B), and service (C) relative to non-SFES for all SFES and disaggregated by science discipline.
Physics (27%) were lower. The patterns of the types of formal training in science education varied among disciplines (Figure 5D). For example, 20% of Biology SFES had earned teaching credentials, 10% had conducted postdoctoral work in science education, and only 5% had earned science education doctorates. SFES in both Chemistry (29%) and Physics (20%) showed relatively large proportions with science education postdoctoral experiences, smaller proportions with science education doctoral degrees (21% and 13%, respectively), and even smaller proportions with teaching credentials (14% and 7%, respectively). In the small sample of Geoscience SFES, 25% held science education doctorates, 13% had teaching credentials, and none had postdoctoral experience in science education.

Perceptions of Activities and Professional Satisfaction

Given the many contributions SFES could make to science education, it is worth noting how SFES perceived the demands on their time for teaching, scholarly activities, and service (Figure 6). A large proportion of SFES in our sample (69%) reported spending “about the same” time on teaching as non-SFES. From here forward, we define non-SFES as other science faculty in the same department who are not specializing in science education. In terms of scholarly activities, there was less agreement about whether they spent “more” (30%), “about the same” (47%), or “less” (23%) time on their scholarly activities compared with non-SFES (Figure 6B). Though the sample size was small (n = 7), over 70% of Geoscience SFES reported spending more time on scholarly activities than their non-SFES Geoscience colleagues, whereas proportions for Biology (26%), Chemistry (14%), and Physics (27%) were lower. Regarding service activities, SFES across all disciplines (69%) perceived spending more time on service than non-SFES (Figure 6C) with none reporting less time spent on service. The extreme was represented by Biology SFES, with 89% reporting that they perceived spending more time on service activities than their non-SFES Biology colleagues. SFES perceived their department and college service activities as being similar to that of non-SFES. At least 50% of SFES reported having specific responsibilities in K–12 teacher preparation, K–12 teacher professional development, and other College of Education collaborations (Figure 3). SFES reported that these additional activities were not expected of non-SFES.

Although SFES were engaged in diverse activities in their positions (Figure 3), there was a high level of agreement among SFES that they were doing the job they were hired to do (Figure 7A). In addition, levels of fulfillment among SFES were very high with regard to their SFES position in general (Figure 7B), with more than 75% reporting being fulfilled by each of these activities.

That fulfillment appeared to be reflected in SFES responses to open-ended questions. Table 2 describes the reasons offered by respondents for continuing to stay in their SFES position (n = 52, IRR = 96%). Half of the SFES who responded reported that they remain in their position because they enjoy the challenge, freedom, and activities of their position. In addition, over one-third of SFES respondents expressed that they stayed because they valued relationships with colleagues and collaborators, including other SFES (37%) and/or they perceived that they were making improvements in science education (35%). Finally, SFES valued their impact on and relationship with students (25%), enjoyed teaching (19%), and/or were geographically tied to their position (15%).

Perceptions of Support and Access to Resources

Although SFES participated in a great variety of scholarly activities (Figure 3), many consistently perceived a lack of institutional support for those activities, as compared with the support they perceived non-SFES received. Most SFES felt their starting and current salaries were similar to those of non-SFES (Figure 8, A and B) but many SFES reported that, upon hiring, they perceived that they received less start-up funding and less laboratory space compared with non-SFES (45% and 53%, respectively; Figure 8, C and D). In addition, most (78%) SFES with departmental graduate programs reported having less access to graduate student researchers to support their scholarly activities as compared with non-SFES (Figure 8E).
Consistent with the perception of less access to resources, generally fewer than 25% of SFES reported being members of science departments with an academic infrastructure—including undergraduate or graduate courses and degree programs in science education—equivalent to that available for basic science training (Figure 9). For example, fewer than 10% of SFES had access to programs offering courses or degrees in science education research for either undergraduate or graduate students (Figure 9, A and B). Generally less than 25% of SFES reported having access to programs offering undergraduate courses or degree programs in science teaching, and this result was consistent across all science disciplines (Figure 9A). SFES as a group reported even less access to graduate programs and courses in science education research or science teaching (Figure 9B).

**Biology SFES Perceptions of Service Demands**

SFES commonly reported that they had greater demands on their time for service activities compared with non-SFES (Figure 6C). A notable difference appeared for Biology SFES, who consistently reported that their efforts were not valued or understood by their university to a degree not seen across other disciplines examined. Smaller proportions of Biology SFES felt that those service activities were understood by their department than did non-Biology SFES (Figure 10A; $\chi^2 = 4.4, p = 0.036$). Similarly, smaller proportions of Biology SFES felt that their service expectations matched those of their university when compared with non-Biology SFES (Figure 10B; $\chi^2 = 5.4, p = 0.020$). That pattern was similar for the perceptions of value for their service activities (Figure 10C).

**Comparison of Perceptions between SFES Who Are and Are Not Seriously Considering Leaving Positions**

Almost 40% of SFES surveyed were seriously considering leaving their current job. Statistical analyses ($\chi^2$ goodness-of-fit) indicated differences between “Stayers,” SFES who were not “seriously considering leaving” their particular position, and “Leavers,” SFES who were “seriously considering leaving” their positions. The terms are merely labels, which may or may not reflect reality; Stayers may choose to leave their positions, whereas Leavers may in fact remain in their positions. The largest differences reflected perceived discrepancies in time spent performing professional activities and access to various academic resources that would help them accomplish those activities, as compared with non-SFES. Recognizing that SFES generally perceived having greater demands on their time for service activities (Figure 6C), almost half (47%) of the Stayers felt they were spending about the same time on service activities, whereas only 11% of Leavers felt that way (Figure 11A; $\chi^2 = 7.1, p = 0.008$). Regarding starting salaries, 87% of the Stayers reported their perceptions of starting with similar or even larger starting salaries as non-SFES in their department, compared with only 58% of Leavers reporting starting with similar or even larger
Investigation of SFES

Figure 8. SFES perceptions of access to resources compared with non-SFES. Reported relative starting salary (A), current salary (B), start-up package (C), lab space (D), and access to graduate students (for SFES in departments with graduate programs; E) for all SFES and disaggregated by science discipline.

Starting salaries than non-SFES (Figure 11B; $\chi^2 = 5.0, p = 0.026$). Perceived current salaries followed the same trend, with 82% of Stayers reporting salaries similar to or greater than those of non-SFES in their department, whereas less than half as many Leavers (40%) perceived having salaries similar to or greater than non-SFES (Figure 11B; $\chi^2 = 9.9, p = 0.002$). Compared with Stayers, a higher proportion of Leavers perceived their start-up packages were less or much less than those of non-SFES (Figure 11B; $\chi^2 = 10.5, p = 0.001$). Finally, regarding levels of academic freedom for developing research projects, less than 3% of Stayers perceived limitations to their academic freedom relative to non-SFES, whereas 20% of Leavers perceived having less academic freedom than non-SFES (Figure 11B).

SFES who self-identified as Stayers and Leavers also differed in their perceptions of their positions (Figure 12, A and C) and professional fulfillment related to their activities (Figure 12B). Compared with Leavers, a higher proportion of Stayers perceived that their overall job expectations were similar to expectations for non-SFES (Figure 12A; $\chi^2 = 8.2, p = 0.004$). Looking at specific professional activities, only 9% of Stayers felt that their teaching expectations were different from those of non-SFES, whereas 40% of Leavers held that perception (Figure 12A; $\chi^2 = 5.7, p = 0.017$). Similarly, only 6% of Stayers felt that their scholarly expectations were different than those of non-SFES, whereas 35% of Leavers held that perception (Figure 12A; $\chi^2 = 5.5, p = 0.019$). Another discrepancy was seen in the perception of service expectations, with only 38% of Stayers perceiving that the service expectations placed on them were different from those placed on their non-SFES peers, whereas 79% of Leavers had that perception (Figure 12A; $\chi^2 = 6.6, p = 0.010$). Regarding service activities, 91% of Stayers but only 70% of Leavers reported being fulfilled by those activities (Figure 12B; $\chi^2 = 2.9, p = 0.091$). Finally, when comparing perceptions of whether they were doing the job they were hired to do, fully 97% of Stayers reported that they felt they were doing the job they were hired to do, whereas only 75% of Leavers shared that perception (Figure 12C; $\chi^2 = 4.5, p = 0.034$).

Reasons for considering leaving were offered by the subset of respondents who agreed that they were seriously considering leaving their current SFES position ($n = 20$ Leavers, IRR = 98%) and are described in nine categories in Table 3. The most prevalent reason for seriously considering leaving cited by SFES Leavers was that science education was not supported, valued, or understood by their department and/or university (35%). In addition, SFES Leavers offered experiences of being overworked, burned out, and under-appreciated (25%), of feeling that their professional values were not aligned with their department or university (25%), and generally feeling that they were not doing what they wanted to be doing (20%). Some Leavers cited issues related to resources—including salaries that were too low (20%) and a lack of resources to support scholarly activities in science education (15%), as well as issues related to workload, including teaching loads (20%) and service loads (15%) that were too high. Finally, 35% of Leavers offered a variety of other reasons for seriously considering leaving that did not fit into the previous categories.

SFES Self-Identification and Pathways

Of our total survey respondents ($n = 78$), 25% did not identify themselves as SFES (“Non-Identifier”; Figure 13). Higher proportions of chemists and physicists identified themselves as SFES than did biologists and geoscientists. Of those who did identify as SFES, there were two apparent pathways to becoming an SFES: those who were “Hired” as SFES and those who “Transitioned” to SFES roles from their initial faculty roles (Figure 13). Of these faculty, over half of Biology (65%) and Chemistry (64%) SFES indicated that they were hired as SFES, and 40% of Physics and 38% of Geoscience SFES reported being hired as SFES. The remaining individuals in
S. D. Bush et al.

Figure 9. Academic infrastructure for SFES scholarship in science education. Reported undergraduate curriculum elements comparable to basic science (A) and graduate curriculum elements comparable to basic science (B) for all SFES and disaggregated by science discipline.

each discipline department transitioned to SFES roles from their initial faculty roles.

SFES Advice about Beginning SFES Positions

SFES were asked to offer advice to beginning SFES. Table 4 describes nine categories of advice offered by respondents for beginning SFES (n = 50, IRR = 97%). The most prevalent pieces of advice offered by SFES to beginning SFES were to obtain clear position expectations from their department and college (44%) and to find colleagues, mentors, and advocates both within and outside their institution (42%). In addition, SFES respondents encouraged others to obtain funding for and publish their work (28%) and reduce their commitments and learn to say no (28%). Finally, several categories of advice offered strategies for navigating SFES positions, such as informing, educating, and highlighting your efforts among stakeholders at your institution (18%); making sure that your department and college value and reward science education activities (16%); and expecting to be treated equitably and just like non-SFES. Finally, having a thick skin and remaining confident and passionate about your work (18%) and having a clear vision of your professional interests (10%) were advised.

DISCUSSION

This research study represents the first systematic collection and analysis of data on the nature of SFES positions, the individuals who occupy these positions, and the state of satisfaction of these individuals with their positions. Results show that SFES positions exist across all science disciplines studied, across all faculty ranks, and at all CSU campuses represented in the study. Thus, the CSU SFES phenomenon is concentrated neither in a single discipline nor on a subset of campuses. Although CSU SFES appear to have been hired into such positions since 1970, more SFES have been hired since 2000 than in the cumulative history of SFES, which is consistent with either a recent expansion of the CSU SFES phenomenon or a long history of SFES hiring and simultaneous attrition.

Although many assume that SFES positions can potentially influence science education in a variety of ways, the SFES phenomenon appears also to have significant challenges. Results from this study suggest that there may be lack of clarity about the role of SFES positions within science departments. In addition, data reveal that 90% of CSU SFES have formal training in science and only 32% have formal training in science education. Results describe motivations for taking an SFES position as well as potential issues that may prompt almost 40% of SFES to seriously considering leaving their position. Although these results are an intriguing profile of CSU SFES, a national study of SFES is needed to investigate the nature and impact of the SFES phenomenon more broadly and to understand how this phenomenon may or may not vary across science disciplines and academic institutions. The results of this research may have implications for individual SFES and also for college and university science administrators hiring them, policy makers and funding agencies, and
non-SFES attempting to understand, support, and evaluate the efforts of SFES.

**Lack of Clarity about the Role of SFES Positions within Science Departments**

There appears to be a lack of clarity about the purpose of SFES positions, both in terms of the academic community’s general perceptions of SFES and more specifically among SFES and stakeholders at their institutions. In the academic community, multiple ideas about the purpose of SFES positions have been expressed informally. Some may perceive that SFES positions are primarily teaching positions within science departments. Our results do not appear to support this idea, because CSU SFES report teaching about the same amount as their non-SFES peers and a majority are engaged in multiple areas of scholarship. Similarly, others may assume that SFES positions are typically introductory course instructor and/or coordinator positions within science departments. However, CSU SFES were teaching courses that span general education through electives for science majors, as well as precredential and credential courses for teaching majors. Finally, others might consider that SFES positions are designed as science education research positions within science departments. Our results, however, do not support this perception either. Approximately the same proportion of CSU SFES reported being engaged in science education research (58%) as reported being engaged in basic science research (61%). As such, our data do not support any singular perception of the purpose of an SFES position, suggesting that there are a variety of different conceptualizations of these positions operating within the CSU and perhaps across the country.

In addition, several lines of evidence from this study suggest a lack of clarity between SFES and stakeholders at their institutions about the role of SFES within science departments. In fact, some potential CSU SFES chose to reject identification as SFES, despite being identified as such by their administrators. Of note, the most prevalent piece of advice offered by study respondents to beginning SFES was to obtain clear position expectations in writing from department and college administrators. The highly varied nature of SFES professional activities is perhaps evidence that the purpose of SFES within individual departments was not explicitly articulated, possibly leading to SFES engaging in a broader set of professional activities than they may have envisioned upon being hired. In particular, SFES indicated that they perceived both a lack of understanding of SFES
scholarly activities in their departments as well as a perception of greater service expectations as compared with non-SFES. More specifically, higher proportions of Biology SFES perceived misunderstandings between themselves and their administrators regarding service expectations and activities, as compared with SFES from other disciplines. Higher proportions of SFES were considering leaving their positions if they did not feel they were doing what they were hired to do, and as reasons they stated science education was not valued, understood, or supported by their university. One possible explanation for these frustrations is that departmental administrators and non-SFES may conflate SFES scholarly activities and service activities due to a lack of understanding of science education efforts. SFES advise new SFES to obtain clear expectations, educate colleagues about the nature of SFES work, and make sure that the department and college value and reward science education activities. SFES advice was consistent with a general lack of specificity of SFES roles and reasons for their hiring.

Results from this study suggest that departments interested in establishing SFES positions need to have explicit conversations and written expectations about the nature of SFES positions (Bush et al., 2006). Involvement in all science education arenas—K–12 science education, undergraduate science education, and science education research—may extend beyond the formal training and capacity of any individual SFES, yet SFES may be looked to as experts in all these arenas. Departments could choose to hire SFES to specialize in a particular arena (e.g., K–12 science teacher preparation or discipline-based science education research),

**Figure 11.** SFES Stayers’ versus SFES Leavers’ perceptions of time spent on professional activities and access to resources compared with non-SFES. (A) Perceptions of time spent on teaching, scholarship, and service relative to non-SFES disaggregated by those who are not “seriously considering leaving their current position” (Stayer) and those who are “seriously considering leaving their current position” (Leaver). (B) Perceptions of access to resources relative to non-SFES disaggregated by Leaver and Stayer. *Indicates statistically significant differences (p < 0.05). ‡Indicates statistically significant differences (p < 0.05) when comparing the joined category of “more or much more” and “about the same” with “less or much less.”

**SFES Formal Professional Training More Extensive in Science Than in Science Education**

Broadly speaking, CSU SFES appear to be similar to other university science faculty in their basic science training. Nearly 90% completed science PhDs, and nearly 40% had a postdoctoral position in science. Although more SFES who were hired into these positions have formal training in science education compared with SFES who have transitioned into an SFES role, formal training in science education is minimal for both groups (Bush et al., 2008). These data are consistent with the idea that science departments may prefer to hire individuals trained as basic research scientists into their departments to address science education issues. Because a majority of CSU SFES are working in multiple arenas of science education, they may be engaged in a wide variety of professional activities that require knowledge and skills that depending on local departmental needs. Alternatively, departments may choose to hire an SFES just as they would any other faculty member, with the expectation that the SFES’s scholarly activities would be of their own choice and generally in one of the arenas of science education. Finally, departments may need to articulate explicitly what will constitute SFES scholarly activities—whether these will be focused in basic science research, science education research, or both—and how SFES scholarly activities are distinct from service activities. Clarity and agreement among SFES, their administrators, and their non-SFES colleagues about the purpose of the position within the department appears to be critical.
Investigation of SFES

Figure 12. SFES Stayers’ versus SFES Leavers’ position expectations and professional fulfillment. Perceptions of professional expectations relative to non-SFES (A), fulfillment by professional activities (B), and engagement in professional activities hired to do (C) disaggregated by those who are not “seriously considering leaving their current position” (Stayer) and those who are “seriously considering leaving their current position” (Leaver). *Indicates statistically significant differences ($p < 0.05$).

Figure 13. Pathways to SFES positions. Percentages of respondents who were hired into SFES positions, transitioned into SFES positions, or did not self-identify as SFES for all respondents and disaggregated by science discipline.

extend beyond those acquired in their formal academic training. Whereas many science faculty pursue professional activities outside their original area of training, SFES may differ in that they are more likely to be working in areas traditionally considered domains of social science. Although CSU SFES as a group reported high levels of fulfillment in their position and overall job activities, scholarly activities had the lowest proportion (ca. 75%) of reported fulfillment. Potentially, SFES may find that their basic science training is inadequate for them to engage in multiple arenas of science education. At a minimum, SFES working in new areas outside of their formal training may need additional time to develop new
Table 3. Reasons offered for seriously considering leaving current SFES position in response to the following question: “What are the primary reasons you are seriously considering leaving?” (n = 20)

<table>
<thead>
<tr>
<th>Category</th>
<th>Sample quotes</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Science education is not supported, valued, or understood by department and/or university</td>
<td>• Lack of any real support from the college or university for science education. • Lack of understanding about what science education is and its value among the non-SFES faculty in my department. • Lack of understanding of SFES research and how to evaluate it.</td>
<td>35</td>
</tr>
<tr>
<td>Overworked, burned out, and unappreciated</td>
<td>• I feel burned out. • I’m underpaid and overworked. • I see new faculty who are only rewarded for traditional research and who care very little about educating students. • Anyone who does not fit this very traditional research mold is a second-class citizen. • There is a lack of acknowledgment for hard work. • I struggle coping with the stress.</td>
<td>25</td>
</tr>
<tr>
<td>Professional values are not aligned with department and/or university</td>
<td>• My professional activities are not valued by my colleagues. • I am not sure that my department and the CSU have the values that I thought I saw when I started this job. • In networking with other SFES faculty, I realize that there are more problems in this field than I would have as a scientist so the choices are to leave the field or the CSU system since what I do is not valued on any level.</td>
<td>25</td>
</tr>
<tr>
<td>I am not doing what I want to be doing</td>
<td>• I am not spending my time how I want to spend it. • I am unhappy with the quantity of work I am expected to do and the nature of much of this work. • When I was hired, it was made explicit that I would be able to choose research areas, including scholarship of teaching. This was a lie.</td>
<td>20</td>
</tr>
<tr>
<td>Salary is too low</td>
<td>• Salary is too low and cost of living here is too high. • Poor salary, high housing costs. • The financial support provided through the CSU system as an Associate Professor combined with the high cost of living in urban southern California led me to consider alternate positions.</td>
<td>20</td>
</tr>
<tr>
<td>Teaching load is too high</td>
<td>• How can you have scholarly activity with a 12-unit teaching load? The CSU has outrageous teaching loads as compared with University of California and private institutions. • The teaching expectations at my university are too high (24 units/year.)</td>
<td>20</td>
</tr>
<tr>
<td>Lack of resources to support scholarly activities in science education</td>
<td>• I have no problem mixing service with scholarly activity, but without a graduate level program that I can connect with, there is not much that is available in the way of scholarly activity grants available. • They take a “bean counting” approach to evaluate your research program, plus personnel committee usually does not have the background to evaluate education research.</td>
<td>15</td>
</tr>
<tr>
<td>Service load is too high</td>
<td>• This position is all service, which is not what I wanted, nor signed up for. • Although I am somehow juggling both, I do not feel that it is appropriate to expect me to be heavily involved in the teacher preparation activities—for which I have had zero training—and to build from the ground up an education research program. • Because I have somewhat more breadth in my training than a classical scientist, I am over-burdened with service expectations re: teacher prep, student care and recruitment, evaluation, etc.</td>
<td>15</td>
</tr>
<tr>
<td>Other</td>
<td>• I am on an administrative track and, if I decide to move up, I will likely have to move on. • Quality of life issues in my geographic area. • I’m retiring.</td>
<td>35</td>
</tr>
</tbody>
</table>
specialize in K–12 science education efforts (Bush et al., 2010). In addition, it is not yet clear whether individuals completing these types of postdoctoral training will be hired as full members of a science department, that is, full-time, tenure-track faculty, or as academic specialists within the department, such as lecturers and/or curriculum coordinators. Expectations for SFES to be full faculty members and/or involved in multiple arenas of science education may be supported by the development of a greater variety of training pathways at the graduate, postdoctoral, and even faculty levels. Furthermore, the multidisciplinary nature of SFES work suggests that successful SFES training pathways may need...

Table 4. Advice offered to beginning SFES in response to the following question: “What are the three most important pieces of advice you would offer to a beginning Science Faculty with an Education Specialty?” (n = 50)

<table>
<thead>
<tr>
<th>Category</th>
<th>Sample quotes</th>
<th>%</th>
</tr>
</thead>
</table>
| Obtain clear expectations from department and college | • Go over and/or negotiate the RTP requirements for your position with both your Dean and your Department Chair in the same room, including how you will be supported to accomplish those requirements.  
• Clear expectations, clear expectations, clear expectations.  
• Know all the expectations your department and college have for your position in terms of service and research/professional development.  
• Make sure that the department fully supports the position by writing its personnel document to align with scholarly activities that SFES perform. | 44 |
| Find colleagues, mentors, and advocates | • Establish a network of colleagues within and outside the department, for research support (collaborations) and for personal/professional support.  
• Ask SFES at your institution and outside your institution for insight, advice, mentoring, and collaborations.  
• It is extremely valuable to have senior faculty on your campus with whom you work, who will publicly state the value of the educational work, which you perform. They are your advocates when there is doubt among your peers.  
• Be sure there is at least one other science education specialist in your department or another science department. | 42 |
| Obtain funding and publish your work | • Apply for grants, big and small, often! Money talks and will buy you lots of cultural currency in the university.  
• Develop innovative approaches to science education that involve students and can be published.  
• Apply for external funding and contribute to the literature base.  
• Aggressively pursue external funding for your scholarly activities so SFES can be as productive as possible during early years in the position. | 28 |
| Reduce your commitments and learn to say no | • Avoid over-committing to departmental or university service.  
• Limit what you take on, even if other people refuse to do the work.  
• Know how and when to say no or have someone mentoring you who can give you hints.  
• Protect your time.  
• The world won’t fall apart if you don’t do everything. | 26 |
| Inform, educate, and highlight your efforts among your faculty colleagues and administrators | • Take advantage of every opportunity to educate your colleagues about what you do.  
• Be good at PR on your service work…let everyone know precisely all that you do.  
• Toot your own horn: publicize your achievements widely and loudly. | 18 |
| Have a thick skin, remain confident, and stay passionate about your work | • Be confident that your expertise is valued even if it is not obvious.  
• Some science faculty even have a hard time valuing plants.  
• Be prepared to be underappreciated/misunderstood by your colleagues. …(have) a thick skin and a warm wit.  
• Humor and patience will win over more individuals than aggression and frustration. | 18 |
| Make sure department and college value and reward science education activities | • Make sure that the Department will appreciate science education research on an equal footing with scientific research.  
• Be sure the department and college do value education research. | 16 |
| Expect to be treated like non-SFES faculty | • Request a reduced teaching load, travel funds and a comparable start-up package as non-SFES faculty.  
• Be sure you have the same opportunities for internal support for scholarship as non-SFES.  
• Make sure that the “assigned or expected” duties are in line with everyone else within the department.  
• Don’t get cornered into doing things that non-SFES faculty would not be expected to do.  
• When things go to hell, remember the passion that brought you here and focus on that aspect of your work.  
• Be strong! You are the future of science in the CSU! | 10 |
| Have a clear vision of and follow your professional interests | • Have a vision of what you’re trying to do and why you took this position.  
• Figure out the things that are important and effective to do.  
• More likely than not you will define your position. Define success you can live with.  
• Be sure your research is driven by your interests, not totally by departmental needs. | 10 |
to convene experts and mentors for future SFES from a variety of disciplinary fields beyond the scientific disciplines, including individuals who have occupied SFES positions themselves. Existing programs with the potential to train future SFES appear to have been developed primarily by traditionally trained scientists and with little to no influence from social scientists, science education researchers, or individuals with expertise in teacher preparation and K–12 education (Wieman, 2009). The establishment of multidisciplinary SFES training pathways may prove to be an important approach in sustaining the SFES model for building expertise and effective efforts in science education within college and university science departments.

Reasons Offered for Taking an SFES Position and Considerations about Leaving

This study portrays the majority of CSU SFES as individuals who are enthusiastic about their position. CSU SFES most often expressed that interest in science education—in the arenas of K–12 science education, undergraduate science education, and/or science education research—was the reason they originally took an SFES position. Beyond this, the motivating factors for individuals to take an SFES position did not appear to be specific to the nature of SFES work. In addition, CSU SFES report that they stay in their current positions because of the challenge, freedom, and flexibility of their jobs, as well as because of the positive relationships they have with colleagues, collaborators, and students. Importantly, many CSU SFES perceive that they are making a difference in science education.

Strikingly and despite these motivations, almost 40% of CSU SFES expressed that they were seriously considering leaving their current positions, but not the field. Because this study was an initial description of the SFES phenomenon, the uniqueness of SFES considerations of leaving compared with non-SFES is unclear. However, there were numerous statistically significant differences between those SFES who were considering leaving and those who were not. Data presented here show that SFES who were considering leaving have perceptions of unequal treatment compared with non-SFES colleagues. In addition, the most common reason offered by SFES considering leaving was that their work in science education was not supported, valued, and/or understood. CSU SFES perceived differential access to resources to support their scholarly activities compared with non-SFES, especially in terms of start-up funds, laboratory space, and access to graduate students and existing academic infrastructure, such as course and degree offerings in their field. In addition, the majority of SFES perceived that they spend more time on service activities than non-SFES, and high service loads were reported explicitly as a reason for seriously considering leaving. These perceptions of less scholarly support and elevated service were more prevalent among those CSU SFES who were seriously considering leaving. Some also reported a misalignment between their professional values and those of their colleagues and a more general feeling of being overworked, burned out, and unappreciated in their institutions.

To capitalize on the value that SFES could add to science departments, the structure and purpose of these positions need to be explicitly articulated and the material, intellectual, and time resources allocated be sufficient to meet SFES responsibilities. In particular, attention by administrators to SFES perceptions of inequality compared with non-SFES, especially in terms of access to resources to support scholarly activities and service expectations, may be key to SFES satisfaction and retention. For example, some SFES may need time to devise and implement an academic infrastructure to support training of students to participate in their scholarly activities. Other SFES may need time to gain expertise in a science education arena that is new to them. Finally, greater clarity about SFES job expectations and more and varied opportunities for SFES training, as described above, may also contribute to reducing potential SFES dissatisfaction and increasing SFES retention.

Need for National Study on the Nature and Impact of the SFES Phenomenon

Although the results of this CSU SFES research study are informative, a national study of the SFES phenomenon would provide a stronger basis for understanding this model for transforming science education. Some of the results presented here may represent the general SFES phenomenon, whereas others may only be characteristic of the CSU system. As an example, CSU SFES positions are not primarily teaching positions, but SFES positions at other types of academic institutions may be. Some may consider SFES positions as primarily discipline-based science education research positions, similar to many physics education research groups within physics departments, but CSU SFES positions are not. Rather, CSU SFES appear to be heavily involved in K–12 teacher preparation, which may be specific to the CSU and its primary role in California teacher preparation. To clarify the extent to which these findings about CSU SFES are predictive of the SFES phenomenon more generally, a national study of SFES across multiple types of academic institutions is essential. In addition, a national study of larger numbers of SFES would provide sufficient statistical power to investigate discipline-specific differences, if any, among SFES. Finally, a national study of the characteristics of SFES, their institutional contexts, and their perceived impact could shed light on the extent to which SFES are a transient phenomenon or represent a larger jurisdictional shift of science education into university science departments.

CONCLUSIONS

Our research study investigated the nature of the SFES phenomenon in the United States. We found that CSU SFES were engaged in the three key arenas of science education (K–12 science education, undergraduate science education, and discipline-based science education research), as well as basic science research. As such, CSU SFES appeared to be well-positioned to have an impact on science education from within science departments. However, there appeared to be a lack of clarity and agreement about the purpose of these CSU SFES positions, which could hinder both the effectiveness of SFES efforts and the growth of the SFES phenomenon. In addition, formal training in science education among SFES was limited, suggesting a need for formal training opportunities in science education for future and current SFES. Whereas over 75% of SFES were fulfilled by their teaching,
scholarship, and service, our results revealed that almost 40% of CSU SFES were seriously considering leaving their positions but not the field. A key statistically significant difference between those seriously considering leaving and those not seriously considering leaving was the perception of inequitable access to resources as compared with non-SFES. Our data suggest that science departments may benefit from explicit discussions about the role of SFES and strategies for supporting their professional activities. Our research findings have the potential to inform a national study of SFES, promote a greater SFES community, and strengthen the conversation among science education stakeholders about the purposes of SFES positions and strategies to maximize SFES impact.

ACKNOWLEDGMENTS

We thank the CSU Chancellor’s Office for funding and Dr. Beverly Marcum (CSU Chico) and Dr. Ed Walton (Cal Poly Pomona) for organizing the 2006 CSU-wide Science Education Colloquium in Pomona, CA, where the authors of this study initially met. We also thank the Deans of Colleges of Science for their assistance in identifying CSU SFES. Funding for this research was generously provided by the following CSU campus administrators: Dean Phil Bailey (Cal Poly San Luis Obispo), Dean Gary Novak (CSU Los Angeles), Interim Dean Jane Bruner (CSU Stanislaus), Dean Stanley Malloy (San Diego State University), and Dean Sheldon Axler (San Francisco State University). S.B. would like to acknowledge the Cal Poly San Luis Obispo, the National Science Foundation (Grant No. 0629887), and the Cal Poly Pomona for their support. N.P. was supported by the NSF Independent Research/Development Program while on leave from CSU Fullerton in 2007 to serve as Program Director in the Division of Undergraduate Education (EHR/DUE) of the NSF; any opinion, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the NSF. We thank Dean Steve Murray (CSU Fullerton) for providing space to meet, Dr. Barbara Gonzalez (CSU Fullerton) for use of her laboratory, and SFES outside the CSU for their participation in validation of the survey instrument. Finally, we thank the CSU SFES who participated in this research and our families for their ongoing patience and support.

REFERENCES


A Study Assessing the Potential of Negative Effects in Interdisciplinary Math–Biology Instruction

Andreas Madlung,* Martina Bremer,† Edward Himelblau,‡ and Alexa Tullis*

*Department of Biology, University of Puget Sound, Tacoma, WA 98416; †San Jose State University, San Jose, CA 95192; and ‡Biological Science, California Polytechnic State University, San Luis Obispo, CA 93407

Submitted August 23, 2010; Revised November 1, 2010; Accepted November 29, 2010

Monitoring Editor: John R. Jungck

There is increasing enthusiasm for teaching approaches that combine mathematics and biology. The call for integrating more quantitative work in biology education has led to new teaching tools that improve quantitative skills. Little is known, however, about whether increasing interdisciplinary work can lead to adverse effects, such as the development of broader but shallower skills or the possibility that math anxiety causes some students to disengage in the classroom, or, paradoxically, to focus so much on the mathematics that they lose sight of its application for the biological concepts in the center of the unit at hand. We have developed and assessed an integrative learning module and found disciplinary learning gains to be equally strong in first-year students who actively engaged in embedded quantitative calculations as in those students who were merely presented with quantitative data in the context of interpreting biological and biostatistical results. When presented to advanced biology students, our quantitative learning tool increased test performance significantly. We conclude from our study that the addition of mathematical calculations to the first year and advanced biology curricula did not hinder overall student learning, and may increase disciplinary learning and data interpretation skills in advanced students.

INTRODUCTION

In his autobiography, biologist Charles Darwin wrote: “I have deeply regretted that I did not proceed far enough at least to understand something of the great leading principles of mathematics, for men thus endowed seem to have an extra sense” (Darwin, 1995). Modern biology requires an ever-increasing set of computational and statistical skills. Thus, Darwin’s sentiment finds resonance not only among modern science educators (Kilpatrick et al., 2001; Richland et al., 2007), but also with future employers and funding agencies, which have introduced programs specifically to support education at the interface between biology and mathematics (National Science Foundation, 2009).

Different strategies have been proposed that educators and institutions could use to increase the mathematical ability of biology students. These strategies include a greater integration of hands-on mathematical problems into science classes (National Research Council [NRC], 2003; Hodgson et al., 2005), the use of more biological examples in traditional mathematics courses (NRC, 2003; Robeva and Laubenbacher, 2009), the development of introductory biology textbooks that use quantitative problems or computational exercises (Jungck, 2005), joint teaching of existing courses by faculty from biology and mathematics (Katz, 2003), and the development of entirely new curricula, integrating rigorous coursework between biology and mathematics (Bialek and Botstein, 2004).

Arguably, the quickest way to increase the mathematical ability of biology students might be to require more math coursework from biology students or to simply add

DOI: 10.1187/cbe.10-08-0102
Address correspondence to: Andreas Madlung (amadlung@pugetsound.edu).

Potential conflict of interest: Three of the four authors of this assessment study (A.M., M.B., E.H.) were also authors of the curriculum or (noncommercial) product to be assessed. No promotion of a particular product to the exclusion of other similar products should be construed from this article.

© 2011 A. Madlung et al. CBE—Life Sciences Education © 2011 The American Society for Cell Biology. This article is distributed by The American Society for Cell Biology under license from the author(s). It is available to the public under an Attribution–Noncommercial–Share Alike 3.0 Unported Creative Commons License (http://creativecommons.org/licenses/by-nc-sa/3.0).

"ASCB®" and "The American Society for Cell Biology®" are registered trademarks of The American Society for Cell Biology.

104
computations and mathematical applications to existing biology classes. Requiring more course work in mathematics is often not feasible due to limits on the total number of compulsory courses for any major in undergraduate college programs (Ares, 2004). Moreover, although the approach of adding more mathematical computations into existing biology courses is comparatively straightforward and cost effective, there are also potential problems associated with this strategy. For example, some biology instructors worry that too much disciplinary material would have to be sacrificed to accommodate the greater inclusion of mathematics (Gross, 2004), or that instruction in mathematical skills, in addition to the typical workload, could become overwhelming for the students (Brent, 2004).

Much of the focus of the last decade in this discussion has been to develop new integrative tools that would help students both learn quantitative skills and directly apply them to the material covered in science classes. Assessments of such learning tools have focused on determining if including a quantitative component leads to an increase in the quantitative ability of their students. In fact, many studies have demonstrated that integrating more mathematics, bioinformatics, or statistics into biology coursework can lead to an increase in the students’ quantitative or analytical ability (Campbell et al., 2006; Metz, 2008; Arnett and Van Horn, 2009; McEwen et al., 2009; Pursell, 2009). By contrast, comparatively less attention has been directed toward testing how integrating mathematics into the biology curriculum affects students’ comprehension of the biological concepts underlying the mathematical examples. Our interest in this question stemmed from a mixed lecture and assignment-based learning module we had developed that was designed to integrate concepts of basic statistics into a unit focusing on modern genetic and biotechnological tools (Bremer et al., 2010). The goal of the module was to demonstrate to students that learning statistics is vital for biologists and that modern biological techniques are increasingly dependent on mathematical tools. We were concerned that introducing quantitative concepts in an introductory biology class could also have adverse affects. For example, did including a substantial introduction to mathematical probabilities in the unit negatively affect our students’ ability to learn the biological principles and/or hamper their ability to interpret mathematical and statistical results in a biological framework? Even if such a unit increased the ability of our students to perform computations, the possibility of negative effects on the ability to interpret statistical results in a biological context may make faculty less likely to invest the time and effort in increasing the quantitative components of courses.

The notion that integrating a substantial amount of computation into a traditional biology class could negatively affect a student’s ability to apply mathematical and statistical results to biological concepts is not without basis. For example, simply the fear of having to do mathematical calculations, known to school psychologists as math anxiety, is a real problem for many students (Tobias, 1987; Bessant, 1995; Ashcraft, 2002), and could lead to adverse learning results. Math anxiety not only could hinder students’ progress in learning the quantitative tools added to the biology curriculum, but also could cause students to perform poorly in sections of the class in which they might otherwise have excelled. Unless math anxiety can be overcome with practice or through intervention (lossi, 2007), this anxiety could potentially lead to a loss of interest or self-confidence and may cause students to abandon a career in the life sciences or any field requiring mathematical aptitude or analysis (Ashcraft and Krause, 2007). Math anxiety is widespread, as evidenced by studies in which as many as 85% of students in introductory math classes reported to have at least mild math anxiety (Perry, 2004). Given these findings and our own anecdotal experience, we wondered if anxiety about quantitative work could also lead students to 1) focus so much on correctly performing the statistical and mathematical calculations that they lose sight of why they are performing the calculations in the first place, and 2) fail to realize that, within the context of biology, mathematics and statistics are simply tools used to gain a deeper understanding of the biology, and not an end in themselves.

In the present study, we assessed how the increased use of mathematics and statistics during a class unit on microarray technology affects student learning. Over the past decade, the use of microarrays has become widespread in many areas of biological research ranging from physiological ecology to genetics to medicine. The utility of microarrays rests with the fact that they enable researchers to investigate the activity of thousands of genes in a single experiment. Because analyzing such a large data set requires quantitative skills, this topic provides an excellent vehicle with which to integrate statistical tools and biological concepts in a meaningful context. Using this topic as a way to introduce statistical comparison of means was even more important in light of the fact that BIOL 111, the first course in a two-semester introductory biology sequence, does not otherwise introduce students to statistical tools during its other course topics. Moreover, the widespread use of microarrays in both medicine and basic research should help capture and hold the interest of beginning and advanced students. Specifically, we wanted to test what impact integrating mathematical and statistical techniques into biology curricula has on the ability of students to 1) learn biological concepts that could be taught with or without the introduction of computational skills, and 2) understand the biological relevance of statistical results. To this end, we developed two versions of our learning module (Bremer et al., 2010), both of which familiarized students with the use of microarrays in biological research. The module consisted of two lectures and associated assignments. The first lecture was about microarray technology and introduced examples of applications of the technique in biological experiments. The second lecture contained a statistical component, which was intended to help students understand microarray data analysis and to introduce them to a suite of statistical terms and tests relevant to all biological disciplines. To test the hypothesis that additional mathematical computation tasks would decrease overall student learning in a biology class, we designed two customized versions of this module. One version required computing statistical measures, whereas the other required interpreting statistics in the context of biological applications.

When designing our study, we realized that the impact of overlaying a substantial amount of computational work onto biological instruction might differ between beginning and advanced students. In this context, we reasoned that students in advanced biology classes are likely to have not only more background in biology, mathematics, and statistics, but also a greater confidence in their academic abilities than are
students in introductory classes. This difference could render more advanced students less vulnerable to any negative effects. Indeed, a study focusing on students’ anxiety about learning statistics showed a correlation between class level and anxiety (Sutarso, 1992). Additional studies showed that students with more experience in mathematics generally had a more positive attitude toward quantitative subjects, such as statistics (Perney and Ravid, 1990; Brown and Brown 1995; Mills, 2004), and that students who had experienced the need of quantitative tools in their particular field of study showed a greater overall interest in learning more about them (Evans, 2007). In light of these findings, we hypothesized that the effect of additional math on student performance in biology courses depends on the students’ class standing, with beginning students being less able to handle extra math assignments and more advanced students being better able to handle interdisciplinary aspects of the class. To test our hypotheses, we chose two different classes as our study systems: an introductory biology class serving mostly freshmen and sophomores, and an advanced biology class taken only by juniors and seniors.

METHODS

The courses used in this study were taught at the University of Puget Sound, WA. The University of Puget Sound is a small, private, 4-yr liberal arts college located in suburban Tacoma. Average student enrollment is approximately 2600. The module discussed in this study was administered in all sections of both courses by the same professor (A. M.).

Learning Module for an Introductory Biology Course

Students in an introductory biology course (BIOL 111, Unity of Life) served as the study population for addressing our research question on first-year biology students. BIOL 111 is designed for students majoring in biology and is taken by both majors and some nonmajors, either as a service course for other majors (e.g., Psychology) or to fulfill the university’s Natural World core requirement. This course is commonly taken in a student’s first semester at Puget Sound and familiarizes students with biomolecules, cells, major metabolic pathways, and genetics. It is also the biology class that students with Advanced Placement Biology credit or Interna-
tional Baccalaureate credit may omit. Importantly for the present study, this course traditionally has not introduced students to any statistical or mathematical concepts. Five sections of this course are taught annually in both fall and spring semesters with approximately 40 students in each section. In the present study, the module was administered to students in four separate sections of this course over three consecutive years from 2007 to 2009. Sample sizes for the four sections were 42, 39, 36, and 42, yielding a combined sample size of 159. All sections of this course used the same laboratories, had roughly the same lecture schedule, used the same textbook, and were generally homogeneous in student preparation and performance, as evidenced by comparable overall course grade averages in BIOL 111 (data not shown). An informal poll revealed that one upper-level student had previously taken statistics, so this student was excluded from the analysis. We believe that any additional differences in prior exposure to statistics that may have existed in our study population should have little bearing on our results because our assessment of outcomes (later in this article) focused on a student’s ability to use specific statistical results to interpret biological data and not on a general ability to perform statistical tests.

The module started with two consecutive 50-min lectures designed to impart a suite of concepts ranging from biology to statistics (summarized in Table 1). The first lecture focused on biological and biostatistical concepts and facts associated with microarrays. The lecture started with a slide presentation that introduced students to the biological concepts underlying how biologists use the technique of hybridization to identify genes or their transcripts. Following this, the instructor introduced the concept of experimental design and then led a discussion on the difference between technical and biological replications. The lecture concluded with a Web-based animation of the microarray technique (Campbell, 2001) and a detailed description of the procedures involved in performing RNA isolation and microarray hybridization. The second lecture focused on the mathematics and statistics involved in analyzing microarray data. The lecture opened with an introduction to relevant statistical terms, including standard deviation and the null hypothesis. Following this introduction, the instructor explained the use of t tests, including the interpretation of test statistics and p values. The lecture concluded with an explanation of the utility of the Bonferroni multiple testing correction method.

At the end of the second lecture, students were randomly divided into two groups. Students in both groups were given a packet containing a detailed handout of the lecture material and one of two versions of a take-home assignment that differed in whether or not computational tasks had to be performed. One version focused on the broader biological concepts and uses of microarrays and how statistics are used to analyze and interpret microarray data. Students receiving this version were asked a series of questions throughout the exercise (see Figure 1A for examples) but were not asked to perform any mathematical computations. Learning outcomes for this packet included the interpretation of statistical test results and knowledge of important terms for data analysis. Here, we term this version the “passive math” version because the students did not perform any of the mathematical calculations related to the statistical tests. The second version of the exercise, termed the “active math” version, was identical in content to the passive math version except that the accompanying questions included a series of hands-on computational tasks in addition to questions related to the interpretation of biostatistics in the context of microarray experimentation. All computations were related to analyzing microarray data, and included the calculation of log-ratios and standard deviations, normalization of data, and the application of a t test, which the students performed using a simple handheld calculator (see Figure 1B for examples). Although the assignments were similar in overall length and took the instructor approximately the same amount of time to complete, it is possible that the active math version took students longer and required more active engagement with the material than did the passive math version because of the need to perform statistical calculations. Take-home assignments were collected and graded for effort and completeness. We did not grade for mathematical accuracy.
because this could apply only to the assignment for the active math group. Based on the grades earned on the take-home assignments, students in both groups performed comparably on their respective homework (these data were not further evaluated for this study).

We then used two sets of questions to assess whether adding computational tasks led to differences in the students’ ability to answer questions relating to data interpretation within the biological framework of the unit. Importantly, these questions did not ask students to perform any computations, focusing instead on how statistics are used to interpret microarray data to uncover the biological meaning of the results. This approach was chosen for two reasons. First, the main objective of our study was to determine if performing computations would hamper the ability of the active math students to understand the underlying statistical concepts as they pertain to a biological question, as opposed to, for example, simply understanding the meaning of various statistical terms. Second, it would have been unreasonable to assess the passive math group on computational skills that they did not learn. The first set of assessment questions was administered as a stand-alone quiz 3–5 d after completion of the take-home assignment. The second set of questions was embedded within a comprehensive final examination, which was given 5–9 d after the quiz, depending on the final examination schedule. This second set of questions was included in the study to assess if the two groups differed in retention of information when it was presented as part of a larger body of material encompassing multiple areas of biology. Both sets of assessment questions contained five “circle all that apply” problems, each with five possible choices, effectively resulting in 25 true/false questions. Figure 1C provides an example of quiz questions related to the material presented in Figure 1, A and B. Other components on the final examination varied in each class from year to year, and performance on those aspects was not analyzed with respect to performance on the questions that were part of the module. In all cases, however, final examination grades followed a normal distribution as they did for years before the statistics module was introduced.

### Learning Module for an Advanced Biology Course

Students in a Plant Molecular Biology and Physiology course (BIOL 332) served as our advanced student population. This course was taught annually in spring semesters by the same faculty member (A.M.) and each had an enrollment of 7–12 students per year. No student enrolled in the advanced course in which the module was administered had participated in the module in the introductory course. All students in the advanced course were Biology majors and had completed at least 1 yr of course work in chemistry and 2 yr in biology. Additionally, all students were likely to have taken at least one semester of mathematics, and most had taken a statistics-intensive ecology class (BIOL 211; see later in this article).

The module for advanced students was implemented as part of a 4-wk-long unit that included hands-on microarray work in the laboratory portion of the course. The module was introduced at the end of the 4-wk-long microarray lab, approximately 10 wk after the start of the semester. Data included in this study came from three offerings of the course, taught over three consecutive years from 2007 to 2009. The final course grades for the three offerings were comparable.

---

### Table 1. Examples of key learning objectives for the microarray modules assessed in this study

<table>
<thead>
<tr>
<th>Broad learning objectives</th>
<th>Specific concepts</th>
<th>Emphasized in passive math version&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Emphasized in active math version&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Biological concepts</td>
<td>Experimental design</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Understanding the concept of gene expression changes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>versus complete gene silencing in response to experimental treatment</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Understanding the biology underlying hybridization techniques</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Understanding the biological basis of nonspecific background noise in experimental data</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Biostatistical concepts and data interpretation skills</td>
<td>Interpreting the value of normalized versus absolute data</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Interpreting the biological significance of p values derived from Student t tests</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Understanding the importance of replication and distinguishing between biological and technical replications</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Understanding the importance of nonspecific background noise in analyzing experimental data</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Mathematical and statistical computation skills</td>
<td>Calculating standard deviations from small data sets using a basic calculator only</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Calculating t values in Student t tests using a basic calculator only</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Calculating p values using tables and Excel software</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Stating null and alternative hypotheses in a statistics problem</td>
<td>No</td>
<td>Yes</td>
</tr>
</tbody>
</table>

<sup>a</sup> These versions pertain to the treatment groups in the first-year biology course. All concepts were emphasized in the module administered to students in the advanced class. See Methods section for further explanation.
Figure 1. Examples of exercise and assessment material from the introductory biology microarray learning module. (A) Exercise questions related to statistical decision making from the passive math version, and (B) the corresponding active math version of the module. Assessment questions related to this material are provided in (C). All students received the same set of assessment questions. Complete classroom material and the assessment tools can be found at www.polyploidy.org/index.php/Microarray_analysis. (Continued)
Effects in Interdisciplinary Instruction

Figure 1. Continued.

Vol. 10, Highlights of 2011 109

introductory course. Whereas we were able to make side-by-side comparisons between control and experimental groups in the introductory course, the small sample size in the advanced course restricted us to a pre- and posttreatment experimental design. Although this experimental design enabled us to gauge the effectiveness of the assignment in enhancing understanding, it precluded us from assessing differences between a treatment and control group. Finally, to ensure the largest possible sample size, the statistics assignment was administered to advanced students as a computer exercise during the regularly scheduled lab section of the course, instead of as a take-home assignment as was the case for the introductory students. Students downloaded the worksheets and instructions on individual laptops and worked independently in class. The professor intervened only as needed to help with the downloading, to clarify the written instructions, and to collect the completed exercise at the end of the class. Importantly, even though the exercise was administered in class rather than as a take-home assignment, it was not an instructor-led exercise; students worked as independently as they would have if it had been a take-home assignment. Figure 2A shows a portion of this exercise.

![EXCEPT FROM THE LEARNING MODULE FOR THE ADVANCED BIOLOGY COURSE.](image)

![QUIZ QUESTIONS PERTAINING TO THE MATERIAL IN 2A ADMINISTERED TO ADVANCED STUDENTS.](image)

![QUIZ QUESTIONS PERTAINING TO THE MATERIAL IN 1A AND B ADMINISTERED TO FIRST-YEAR STUDENTS.](image)
Implementation of the learning module began after the microarray wet-lab with two lectures, the first of which focused on biological concepts and facts whereas the second focused on mathematics and statistics. The general learning objectives were comparable to those for first-year students (Table 1). Although the presentations were initially the same for both student populations, we were able to present the concepts of p values and standard deviations as a review to the advanced group. Building on previous knowledge allowed extra time at the end of the presentation to discuss issues of multiple testing, including tests, such as the Bonferroni correction and False Discovery Rate adjustment. Within 1–2 d following the second lecture, students were given a set of questions in the form of a stand-alone quiz designed to assess their understanding of the lecture material. The quiz contained seven “circle all that apply” problems, each with five possible choices, effectively resulting in 35 true/false questions (see Figure 2B for examples).

Following the preassignment quiz, students obtained experience with the quantitative aspect of the lecture material during the computer laboratory session. The laboratory assignment consisted of two Excel files containing small, simulated microarray data sets as well as a set of instructions guiding the students through the computational and statistical analysis of the data. We chose to use Excel as a platform for these exercises because of its relative ease and because by the time students enroll in advanced biology courses they will be proficient in using it for data analysis. The ubiquity of Excel also makes distribution of these data sets via our website convenient for colleagues who might want to use the module in their own courses. All students finished the exercise and turned in their completed answer sheets within the 4-h lab period.

To assess if the laboratory portion of the module led to significant gains in students’ understanding of how statistical tools are applied to fully realize the biological meaning and implications of microarray data that were previously presented in the lecture portion of the course, we administered a set of postlaboratory assessment questions. The number and format of these questions were identical to those of the prelaboratory assessment questions, but they were embedded in a comprehensive final examination rather than given as a stand-alone quiz. The final examination was held 2–3 wk after the computer lab.

Complete classroom materials for both courses along with the assessment tools can be found at www.polyplody.org/index.php/Microarray_analysis (last updated January 2010).

Experimental Design and Statistical Analysis of Assessment Questions

We hypothesized that adding mathematical computational tasks to an exercise in a biology class would negatively impact the ability of students to understand and interpret data in a biological context. The experimental design for students enrolled in a first-year biology course consisted of a side-by-side comparison of two groups: The active math group was assigned additional computations within the framework of an assignment given to the passive math (control) group, which did not perform computations. We administered two sets of assessment questions (a stand-alone quiz and a set of questions embedded within the final examination) to all students. Importantly, these questions were based on material common to both assignments. We used independent sample t tests to compare scores of the passive math and active math groups on our assessment questions. These between-group comparisons were made individually for each of the four sections of this course and then again after combining the scores for all sections. We then used paired t tests to compare the scores on the quiz and the final examination questions, first within individual sections of the course and then for all sections combined. The data of the only upper-level student taking the introductory course for university core credit and who had, to our knowledge, taken statistics previously were excluded from the analysis.

Analysis of learning outcomes in the advanced class necessarily differed from that of the first-year class because of the differences in experimental design due to the much smaller class sizes described previously. To determine if the computational and statistical assignment had any significant effect on the students’ understanding of the biological material, we compared scores of the pre- and postlaboratory quizzes using a paired t test. As with the first-year course, we compared the scores within each of the three individual sections, and then again after combining the scores of all sections.

For all analyses, significance was determined at the 0.05 level using SPSS 13.0 (SPSS Inc., Chicago, IL) or Excel (Microsoft, Redmond, WA) software. For both courses, only scores from students who had participated in all aspects of the module were included in the analysis. Scores from any student who missed any lecture, quiz, or homework assignment were omitted.

RESULTS

First-Year Student Performance

Results for first-year students enrolled in an introductory biology course showed no significant difference between how well the passive math group and the active math group performed on the first set of assessment questions that was presented as a stand-alone quiz. This result held when each of the four sections was examined individually (Table 2) and when the data were pooled (Figure 3; unpaired t tests, p > 0.05 for all). We also compared the scores of the two groups on the second set of questions that were embedded in their comprehensive final examination. Again, the passive math group and the active math group performed equally well on these questions (Table 2, Figure 3; unpaired t tests, p > 0.05). Taken together, these results indicate that both groups understood equally well the meaning of the statistical results in the context of the biological concepts presented in the unit, even though the take-home assignment of the active math group emphasized quantitative tasks as opposed to passive interpretation of the results.

To test whether the addition of a computational component to the assignment had any effect on overall retention when the material was presented as part of a larger body of information, we pooled data for all sections and made a within-group comparison between scores on the stand-alone quiz and scores on the second set of assessment questions embedded within the final examination. Interestingly, both
and the final as measured by paired t tests was significant for both classes 1, 2, 3, and 4; 2 of 3 yr with N < 81, values on the pooled data shown in Table 2. Using a two-sided t test, df = 78, p < 0.001; active math: df = 81, p < 0.001). Specifically, the passive math group scored 9.7% lower on the embedded questions as compared to the quiz, and the active math group scored 7.7% lower. Finally, although pooled data revealed that the active math group had a slightly higher score on the embedded questions than did the passive math group, this difference was not significant (unpaired t test, p > 0.05).

**Advanced Student Performance**

We used a pre- and postassignment testing strategy as our experimental design to test if adding substantial mathematical and statistical computations negatively impacted the ability of advanced students to interpret biostatistical test results and understand the underlying biological concepts. This design was chosen because the small class sizes precluded us from splitting the sections into treatment and control groups. Thus, unlike our experiment with first-year students, knowledge of the advanced students was assessed before and after the assignment. As with the first-year students, assessment questions focused on biological concepts underlying microarray technology and the biological meaning of the statistical results (Figure 2). When classes were analyzed separately, two of the three classes showed significantly higher scores following the assignment (Table 3; paired t tests, df = 11, 7, and 6, respectively; 2 of 3 yr with p < 0.05). When the results of the 3 yr were pooled to generate a larger sample

<table>
<thead>
<tr>
<th>Class (N)</th>
<th>Stand-alone quiz questionsa</th>
<th>Embedded final examination questionsb</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Passive math</td>
<td>Active math</td>
</tr>
<tr>
<td>1 (42)</td>
<td>82.8 ± 9.6</td>
<td>82.0 ± 10.4</td>
</tr>
<tr>
<td>2 (39)</td>
<td>76.8 ± 12.0</td>
<td>76.0 ± 10.9</td>
</tr>
<tr>
<td>3 (36)</td>
<td>70.4 ± 10.5</td>
<td>72.4 ± 8.2</td>
</tr>
<tr>
<td>4 (42)</td>
<td>79.2 ± 6.0</td>
<td>76.4 ± 8.4</td>
</tr>
</tbody>
</table>

% scores are out of 25 points.

*aSignificantly different, paired t test, p = 0.026, 0.86, 0.51, and 0.20 for classes 1, 2, 3, and 4, respectively.

*bThere was no significant difference in the scores of the passive math and active math groups on assessment questions delivered as a stand-alone quiz for any of the four introductory biology classes (two-sided t test, p = 0.72, 0.85, 0.54, and 0.15 for classes 1, 2, 3, and 4, respectively).

**Figure 3.** Adding an interdisciplinary quantitative component to the biology curriculum did not adversely affect performance of first-year students on assessment questions. During a unit on the use of microarrays in biological research, biology students were given practice problems in which they focused on the broader concept of microarrays and their interpretation (passive math) or analyzed microarray data themselves by performing statistical computations (active math). Student performance on biological concepts and data interpretation was assessed twice within 2 wk. The first set of assessment questions was a stand-alone quiz, whereas the second set was integrated into a comprehensive final examination. Data presented in this figure represent the pooled data shown in Table 2. Using a two-sided t test there was no significant difference between how well the two groups performed on the quiz questions (t = 0.829, p = 0.409) or on the final examination questions (t = 0.321, p = 0.562). The results suggest that adding an intensive quantitative component did not negatively impact the students’ ability to interpret data in a biological context. Decreased retention of the material between the quiz and the final as measured by paired t tests was significant for both groups (p < 0.001). N: passive math = 78, active math = 81. Values on the Y axis represent percentages out of 25 points. Error bars indicate SD.

**Table 2.** Student scores (% ± SD) on assessment questions related to a microarray learning module for first-year students

<table>
<thead>
<tr>
<th>Class (N)</th>
<th>Stand-alone quiz questionsa</th>
<th>Embedded final examination questionsb</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Passive math</td>
<td>Active math</td>
</tr>
<tr>
<td>1 (42)</td>
<td>2008 (8)</td>
<td>2007 (12)</td>
</tr>
<tr>
<td>2 (39)</td>
<td>80.0 ± 12.3</td>
<td>87.2 ± 7.7</td>
</tr>
<tr>
<td>3 (36)</td>
<td>74.8 ± 6.9</td>
<td>71.4 ± 6.9</td>
</tr>
<tr>
<td>4 (42)</td>
<td>76.0 ± 10.6</td>
<td>76.0 ± 10.6</td>
</tr>
</tbody>
</table>

% scores are out of 36 points.

*aSignificantly different, paired t test p = 0.044 for 2008 and p = 0.002 for 2009.

---

**Table 3.** Student scores (% ± SD) on assessment questions related to a microarray learning module for advanced students

<table>
<thead>
<tr>
<th>Class (N)</th>
<th>Premodule</th>
<th>Postmodule</th>
</tr>
</thead>
<tbody>
<tr>
<td>2007 (12)</td>
<td>77.2 ± 7.7</td>
<td>76.0 ± 8.9</td>
</tr>
<tr>
<td>2008 (8)</td>
<td>80.0 ± 12.3</td>
<td>88.9 ± 12.5</td>
</tr>
<tr>
<td>2009 (7)</td>
<td>71.4 ± 6.9</td>
<td>91.4 ± 7.0</td>
</tr>
</tbody>
</table>

% scores are out of 36 points.

---

**Effects in Interdisciplinary Instruction**

Vol. 10, Highlights of 2011 111
and concepts. To help answer this question in the context of a first-year biology course and an advanced biology course for majors.

One potential danger of including extensive statistical computations in first-year biology courses is that students might focus more on “getting the math right” than on understanding the biological concepts that the statistics are intended to inform, a phenomenon that could be magnified by anxiety about mathematics in general. Such a shift in focus could lead to students knowing what various statistical terms mean and how to perform basic tests, but lacking the ability to apply this knowledge to biological questions and concepts. Results from the present study suggest that supplementing biology instruction with mathematical and statistical computations in a first-year biology course does not negatively impact the ability of students to use statistics as a means to understand biological concepts. When faced with the question of how to study possible negative influences of math exposure on the ability of first-year students to learn biological concepts and apply statistical results to gain a deeper understanding of the biology, we did a side-by-side comparison of students’ performance on an assignment related to microarray technique in which half of the students focused only on the application of statistical tools to specific biological questions and concepts and the other half performed extensive computations as part of the assignment (Table 1). Results from this experiment showed that both groups performed equally well on assessment questions related to interpreting biological microarray data with statistical tests (Table 2; Figure 3). Stated differently, these results suggest that the students who spent a considerable amount of time learning how to perform statistical tests and thinking through the mathematical concepts underlying the analysis of microarrays were as well equipped to understand the biological concepts and data interpretation as were the students who learned the material in an assignment free of computations. These results are qualitatively similar to those of a recent study that tested the effect of “learning communities” on student performance in math and biology (Arnett and Van Horn, 2009). This study (Arnett and Van Horn, 2009) compared students who took an integrative two-course module in introductory biology and algebra, analogous to our active math group, with students who took separate nonintegrative courses in the two subjects, analogous to our passive math group. Consistent with the results of our work, the authors reported that overall student performance, measured largely by subject tests in biology, was not different between the students in the two groups (Arnett and Van Horn, 2009).

When we tested first-year students on the content of the microarray unit for the second time approximately a week after the first test, the scores of both groups dropped significantly by approximately the same amount in each experimental group (Figure 3). There are at least three possible explanations for this decline: 1) Students may have simply forgotten facts and skills between the two sets of assessment questions, 2) comprehensive finals could affect recall of specific information from the microarray unit, or 3) end-of-the-semester units may be generally less effective due to time constraints. Regardless of the reason for the drop in scores, the fact that both groups showed a similar decrease in their performance on the assessment questions suggests that adding a mathematical component to the exercise had no bearing on retention of the material. Determining whether the timing of the

**DISCUSSION**

Interdisciplinary teaching in general, and mathematical biology in particular, has been heralded as one of the best ways to improve student interest and learning, and has been advocated by educators and administrators alike (NRC, 2003; Bialek and Botstein, 2004; Robeva and Laubenbacher, 2009). Despite the great interest in integrating more mathematics and/or statistics into the biological sciences, few studies have addressed the question of whether this practice could negatively impact student acquisition of biological facts and concepts. Such a negative impact could result, for example, from a general overload of material if additional mathematical units are simply added without removing other aspects of the course. Another negative impact could be alienating or “turning off” students who have a tendency toward math anxiety. The present study aimed to provide empirical data size, the results of our analysis suggested that the assignment significantly improved performance on the assessment questions (Figure 4; paired t test; df = 26, p < 0.05). Specifically, the combined scores of all three classes increased by 7.9% (Figure 4). These results suggest that experience and practice with computations increased the ability of advanced students to use statistics as a way to better understand biological results and concepts.

**Figure 4.** An interdisciplinary computer module requiring students to calculate statistical problems by hand was an effective learning tool for advanced students. Students in an upper-level Plant Molecular Biology course were asked to perform statistical computations in a lab exercise after a lecture introduction to microarray analysis. They were tested on their overall understanding of the biological implications of microarray work before and after the exercise. Data represent the pooled data shown in Table 3. Results suggest that the assignment helped students effectively apply statistical data to interpret biological results and concepts. N = 27, p < 0.05 for a two-sided, paired t test. Values on the Y axis represent percentages out of 35 points. Error bars indicate SD.
module within a course has any influence on overall retention of the material on the final examination would be valuable in a broader sense when deciding when to present any type of challenging and unfamiliar material to first-year students.

Our results for first-year students suggest that integrating a computational component into biological instruction did not affect, either positively or negatively, student performance on follow-up assessment questions. On the surface, neither positive nor negative effect could lead to the conclusion that the only benefit of integrating quantitative work into this unit was that it enabled students to practice mathematical skills in the context of a biological question. It is possible, though, that an additional benefit was realized in the form of increased receptivity to mathematics and statistics in future biology courses. Indeed, in a different study, students taking an integrated course in biology and mathematics were more positive toward mathematics, both as a separate discipline and as part of biology instruction, than were students taking separate courses in mathematics and biology (Arnett and Van Horn, 2009). Similar responses have been found among students who participate in “learning communities,” where student cohorts coenroll in two or more courses (Lenning and Ebbers, 1999). Applying what is learned in one context to a different situation in a different context has also been shown to deepen a student’s understanding (Bransford, 1999). Applying what is learned in one context to a different situation in a different context has also been shown to deepen a student’s understanding (Bransford et al., 2000) and engagement (Zhao and Kuh, 2004) with the subject. It is also possible that merely the integration of mathematics into a multidisciplinary context alleviated some math anxiety in students who might otherwise have shown signs of anxiety. Finally, studies have demonstrated that both self-paced learning and distance education contribute to reducing math anxiety (Iossi, 2007). Thus, the fact that we implemented our module as a self-paced take-home assignment to first-year students may positively influence the receptivity of these students to mathematics in future courses.

Incorporating attitudinal assessments in future studies would help determine if there are similar, less tangible benefits of integrating math and statistics into biology course work. Increased receptivity to mathematics might increase retention in students majoring in biology, and ultimately help shape their career path. To assess such less tangible benefits, one could use questionnaires asking about interest in (or fear of) math before and after the study. Such tests have been developed and used in the past (Suinn and Winston, 2003). Analysis of attitudinal changes after the use of an interdisciplinary learning tool could inform the future development of interventional teaching modules with respect to specific target groups for whom such studies would be most or least beneficial. Finally, comparing the attitudes and reactions of students who were exposed to the module as a self-paced take-home assignment versus one delivered in class would provide valuable information about how best to administer the module to achieve the desired pedagogical goals.

In contrast to the results for first-year students, results for advanced students in our study showed that the addition of a quantitative component tended to improve student performance on questions related to biological concepts and the statistical interpretation of microarray data (Table 3; Figure 4). At least three factors could help explain the different results between first-year and advanced students. First, it is possible that advanced students have an overall lower level of anxiety toward learning mathematics and statistics than do first-year students, a factor that could lead to the boost in learning seen in advanced students (Sutarso, 1992). Because of the structure of the curriculum, advanced students in this study had all completed a year of chemistry and 2 yr of biology, and are likely to have completed a year of math (or the equivalent through Advanced Placement credit). Exposure to mathematics and statistics in these courses means that they have taken more mathematics by the time they enrolled in advanced biology courses, a factor which is likely to reduce math anxiety. Second, advanced students, to a large extent, also have been self-selected for their ability and propensity to learn biology. Finally, advanced students are further along in their academic careers, suggesting that they are less intimidated by new, complex material and are more skilled at studying relevant material. Collectively, these factors could contribute to making advanced students better able to absorb, assimilate, and apply new biostatistical facts and concepts, even when presented in the context of relatively complex statistics. Future studies that systematically test how prior exposure to mathematics contributes to reducing math anxiety would provide important information for curricular design and decision making.

Differences observed between first-year and advanced students could also partly result from the differences in experimental design. Because of the much smaller class size in the advanced course, we could not perform a side-by-side comparison of control versus experimental groups. Instead, we opted for a pre- and posttreatment design, in which the treatment was the computer-based statistics exercise. Even though the advanced students performed the exercise during a scheduled lab, the independent nature of the exercise (see Methods section) makes it unlikely that the improvement was due to the professor’s presence. It is, of course, possible that the improved posttreatment performance would have resulted from any additional exposure to the material, regardless of whether it was presented as dictated by the learning module. Even if this were the case, however, we can still conclude that the addition of the statistics-intensive computer exercise did not have a negative impact on the students’ learning of the biological and biostatistical concepts and data interpretation related to microarrays.

SUMMARY

In summary, our results suggest that our module integrating more mathematics and statistics into biology courses did not have a negative effect on the performance of first-year students and can help more advanced students gain a better understanding of underlying biological principles and concepts. Although our results for first-year students do not support the findings of many studies showing that interdisciplinary learning leads to an overall gain in aptitude, it has to be stressed that we assessed performance only on material that was common to the assignments in both groups (i.e., material related to biostatistical concepts and data interpretation in the context of biology). It is likely that students in the active math group learned additional quantitative skills, but learning or retention of those skills was not assessed in our study because such a test would clearly have put the students in the passive math group at a disadvantage. Given
that our study did suggest differences between first-year and advanced students, future investigations into the impact of how integrating quantitative work into biology curricula affects the learning of biological material should continue to study students at different academic levels. The present study also serves to highlight at least three important questions in need of further investigation.

1. How does integrating quantitative work into traditional biology curricula influence a student’s subsequent attitudes toward mathematics and statistics? Attitudinal surveys administered at various time points following exposure would help address this question. Several math attitude tests exist that could be customized for this purpose (e.g., those presented or mentioned in Aiken, 1963; Sutarto, 1992; Mills, 2004; Evans, 2007).

2. Does implementing a particular quantitative exercise as an in-class exercise versus a self-paced take-home assignment lead to differences in performance on the biological elements and/or attitudes toward the quantitative elements? These questions could be approached by assessing outcomes and attitudes in a given course taught by the same professor over a number of years, with the exercise given as an in-class exercise one year and as a take-home assignment the next year.

3. Does prior curricular exposure to mathematics and/or statistics lead to differences in learning outcomes resulting from interdisciplinary exercises? Addressing this question would require a thorough analysis of each student’s high school and college transcript. In this context, it would also be valuable to assess not only the impact of prior exposure (merely taking a course) but also a student’s aptitude (what grade was earned?). In this way, it would be possible to distinguish between the impact of exposure and the impact of aptitude on learning outcomes and overall attitudes resulting from embedding more quantitative learning tools into biology course work.

If the results of our study can be generalized, class work integrating more data interpretation and underlying mathematical concepts into biology courses may thus offer students additional insights and possibly provide a path to the “extra sense” Darwin wished he had developed.

ACKNOWLEDGMENTS

This study was approved for exempt status by the Institutional Review Board of the University of Puget Sound (approval #0708-018). Funding for the study was provided by a grant from the National Science Foundation (Plant Genome Program, DBI-0501712) to E.H. and A.M. We thank Patricia Marsteller and two anonymous reviewers for their helpful comments to improve the manuscript.

REFERENCES


Effects in Interdisciplinary Instruction


Teaching Students How to Study: A Workshop on Information Processing and Self-Testing Helps Students Learn

Kathrin F. Stanger-Hall,* Floyd W. Shockley,†‡ and Rachel E. Wilson§

*Department of Plant Biology, †Department of Entomology, and §Mathematics and Science Education, University of Georgia, Athens, GA 30602

Submitted November 22, 2010; Revised March 3, 2011; Accepted March 4, 2011

Monitoring Editor: Robert L. DeHaan

We implemented a “how to study” workshop for small groups of students (6–12) for N = 93 consenting students, randomly assigned from a large introductory biology class. The goal of this workshop was to teach students self-regulating techniques with visualization-based exercises as a foundation for learning and critical thinking in two areas: information processing and self-testing. During the workshop, students worked individually or in groups and received immediate feedback on their progress. Here, we describe two individual workshop exercises, report their immediate results, describe students’ reactions (based on the workshop instructors’ experience and student feedback), and report student performance on workshop-related questions on the final exam. Students rated the workshop activities highly and performed significantly better on workshop-related final exam questions than the control groups. This was the case for both lower- and higher-order thinking questions. Student achievement (i.e., grade point average) was significantly correlated with overall final exam performance but not with workshop outcomes. This long-term (10 wk) retention of a self-testing effect across question levels and student achievement is a promising endorsement for future large-scale implementation and further evaluation of this “how to study” workshop as a study support for introductory biology (and other science) students.

INTRODUCTION

In recent years, many studies have documented a lack of proficient critical thinking skills in college students (Holschuh, 2000; Weimer, 2002; Lord and Baviskar, 2007; Crowe et al., 2008; Lord, 2008; Stanger-Hall et al., 2010). Critical thinking, defined here as higher-order thinking skills or levels 3–6 of Bloom’s taxonomy of cognition (Bloom, 1956), is generally viewed as an essential part of college training (Boyer Commission, 1998; National Research Council, 2003). Unfortunately, many students fail to understand the importance of higher-order thinking skills (Stanger-Hall, unpublished data) and consequently struggle to develop these skills.

For instructors of large introductory college lecture classes, it is often difficult to assess whether students practice critical thinking when they study for class, and failure to practice these skills is recognized by the student and the instructor only when students fail to do well on exams that assess critical thinking (application, analysis, evaluation, and synthesis) skills. This is an issue especially during the first year of college when students discover that study techniques that have been successful in high school may not necessarily be successful in college (Matt et al., 1991; Yip and Chung, 2005). More motivated students then find their way to the office of the instructor, where their study routine can be assessed and suggestions for modification of their study habits can be made and practiced. Teaching these skills to individual
students during office hours, however, is not a very efficient means by which to teach students new study skills, especially in large classes. As a result, we decided to develop (K.S.-H.) and implement (F.W.S.) a “how to study” workshop for students in a large introductory biology lecture class.

There is a growing body of literature on study techniques for college students (e.g. Paulk, 2000; Nist-Olejnik and Holschuh, 2008; van Blerkom, 2008), and online study advice is now available from many colleges and universities. Topics range from more mechanistic advice (note-taking, study scheduling) to processing skills (reading strategy, organizing information, creating exam questions, flowcharts, question design) to metacognitive skills (self-reflection on learning goals and adjustment of study time and techniques based on learning outcomes). What exactly constitutes “good” learning goals and adjustment of study time and techniques question design) to metacognitive skills (self-reflection on learning goals and adjustment of study time and techniques based on learning outcomes). What exactly constitutes “good” study behaviors and which behaviors may be detrimental because they use up study time at the expense of more effective study techniques (Gurung et al., 2010) likely depend on both the specific learning goals for an individual class and the knowledge and skill levels taught and assessed by the instructor.

In the introductory biology class in this study, critical thinking was emphasized and Bloom’s taxonomy of cognition (Bloom, 1956) was taught to students as a communication and study tool during the first week of class. At this time, the instructor emphasized to students that 25–30% of the questions on each exam would be asked at Bloom levels 3–5, assessing application, analysis, and evaluation skills. Because the exam format is limited to multiple choice, level 6 (creation/synthesis) is generally not assessed. As a result, students who desire to earn a grade of “C” or higher must master these critical thinking skills. It was apparent to the instructor during office hours, however, that students found it difficult to distinguish what they did know and what they did not know from the class material (Tobias and Everson, 2002), and generally most underestimated their knowledge and thinking skills (Isaacs and Fujita, 2006). In other words, struggling students tended to lack metacognitive skills and study strategies that would have allowed them to self-monitor their knowledge and to adjust their studying and learning outcomes to better achieve their learning goals for the class, a skill set that tends to be a trademark of higher-achieving students (Isaacs and Fujita, 2006). Furthermore, students tended to study facts in isolation rather than putting them in the context of their current knowledge: They were not using contextual thinking to improve their recall or as a basis for critical thinking (reasoning through connections to identify mistakes or misconceptions). As a result, we decided that a “how to study” workshop would demonstrate the benefits of these skills to our introductory biology students.

Visualization (in the form of mental images or as external representations) is important for all learning, but especially for learning in the sciences (Mathewson, 1999; Gilbert, 2005, 2008; Schönborn and Anderson, 2006). Furthermore, visualization relies on context and thus promotes contextual learning. Therefore, we specifically selected two different exercises as workshop activities demonstrating to students how visualization techniques can be used for contextual thinking and recall, for encouraging feedback on existing and missing knowledge and understanding, and as an opportunity for developing cognitively active learning and for practicing critical thinking skills.

EXERCISE 1. INFORMATION-PROCESSING EXERCISE: VISUAL VERSUS AUDITORY PROCESSING

Many college instructors have found themselves at some point teaching a large class and facing a sea of students looking down at their notes, trying to write down word-for-word what the instructor says, rather than paying attention to why it is being said (why it is important, what the context is, etc.). Most students taking notes during lecture tend to process information in an auditory manner by listening and simply recording what they hear. Students who process information visually (imagine what they hear), however, usually outperform students who process information in an auditory manner in terms of their recall ability (Revak and Porter, 2001). Visualization in the form of creating mental images is a fundamental cognitive process that has been shown to improve student learning (Pressley, 1976), particularly in the sciences (Wu and Shah, 2004). We decided to demonstrate this to our students through personal experience.

EXERCISE 2. SELF-TESTING EXERCISE: VISUALIZATION OF EXISTING AND MISSING KNOWLEDGE AND UNDERSTANDING

Many introductory biology students tend to limit their study activities to reviews of their lecture notes and textbook, aimed at memorization of facts and explanations (Karpicke et al., 2009; Stanger-Hall, unpublished data). They are generally unaware of self-testing as a learning strategy other than using old exams or index cards as a memorization tool. Instead, students tend to believe that, once they can recall an item, they have learned it (Karpicke, 2009). Students tend to neglect practicing information retrieval (self-testing) when studying on their own (Karpicke, 2009), but research has shown that testing inserted into the learning phase enhances long-term retention (Agarwal et al., 2008). To illustrate the importance of self-testing for student learning, we used a self-testing exercise as the second workshop activity. This self-testing exercise used drawing as a visualization tool (Gobert and Clement, 1999) to demonstrate to students what they remembered from class and what they didn’t.

Previous studies have shown that visualization of scientific principles through the use of imagery and external representations (diagrams, models) improves student conceptual understanding and ultimately student performance on assessments. We predicted that we would see such effects in our workshop students as well.

GENERAL METHODS

The prescribed “how-to-study” workshop was implemented in Fall 2009 for 99 students, as part of a larger study on the effects of various study supports on the use of critical thinking skills in introductory biology students. This study was conducted in a large introductory biology class (N = 300 students).
students). It was the goal of this larger study to test three different study supports as candidates for future large-scale implementation(s) in large introductory biology classes. The “how to study” workshops constituted one of these three study supports. Students were assigned to the three different study support groups based on their Exam 1 performance: exam scores were sorted in descending order and students were assigned to the three groups in rotating order: 1–2–3, etc. (top score: Group 1, second score: Group 2, third score: Group 3, fourth score: Group 1, etc.). Each of these three groups received a specific study support and group-specific assignments, which were part of the class grade (40 points of 1100). We generated the control groups for the three treatment groups by applying the same methodology to the same class from the year before (Fall 2008), which was taught by the same instructor (K.S.-H.) and received no treatments. This approach had the advantage of generating control groups of similar size to the treatment groups and of compensating for any possible bias generated by the 1–2–3 sequence of assigning students to groups.

We report here on the learning outcomes for the workshop group, whose study support consisted of two 90-min “how to study” workshops that each were offered during multiple (up to 20) time-slots over a period of 8 d. Workshop I was offered during the week following Exam 1, whereas Workshop II was offered during the week following Exam 2. Students signed up for their most desired time slot (maximum of 12 students) but were asked to choose another time if fewer than six students signed up. This study reports the outcomes of the first workshop in the series, which focused on the use of visualization for information processing and self-testing. There was no overlap between the workshop activities and the activities of another treatment group in Fall 2009; therefore, this treatment group served as an additional (same-semester) control group for this analysis.

The first workshop consisted of two visualization-based exercises, which demonstrated 1) different approaches to information processing (as it applies to class and reading) and their effects on remembering information, and 2) how to use self-testing as a routine study tool for remembering and critical thinking, defined as the upper four cognitive levels of Bloom’s taxonomy: application, analysis, evaluation, and synthesis (Bloom, 1956). The desired benefits of these exercises were 1) to help students learn by reviewing the class material actively: This is expected to inform them on what they know and do not know as well as improve both their lower (remembering facts and explanations) and their higher (critical) thinking skills, and 2) to help students identify misconceptions by placing what they remember into the context of other knowledge from previous classes or lectures and using their thinking skills to detect possible discrepancies.

We used assessment of student learning on the cumulative final exam to answer the following questions: 1) Does self-testing lead to learning gains for the topic used for self-testing? If so: 2) Can learning gains be documented for both lower- and higher-order thinking skills? And 3) Do students across achievement levels (as measured by grade point average [GPA]) benefit? To qualify for future large-scale implementation as a study support in introductory biology (and other science) classes, affirmative responses to all questions were required.

**Workshop Implementation**

After the first exam in Fall 2009, 99 students were assigned to the workshop group; however, only 93 students consented to participate in this study. All 93 students attended the workshop and earned 10 points (of 1100) of class credit. As a result, the evaluation of workshop activities is based on a sample size of 93 students. Students could earn additional points by filling out an online feedback survey on their workshop experience, and 79 of the 93 students elected to do so (sample size for feedback survey: \( N = 79 \)). Eleven of the 93 consenting workshop students withdrew from the class by the midpoint of the semester, leaving \( N = 82 \) students for analysis of their final exam performance (sample size for final exam analysis: \( N = 82 \)). In comparison, the same-semester control group (control 2009) had 90 consenting members, and the previous-year control group (control 2008) had 87 consenting members at the end of the semester (control samples for the final exam analysis).

The workshop students met in small groups (a maximum of 12 students per workshop) outside of the regularly scheduled class time in a small conference room to better enhance instructor–student and student–student interactions. All workshop sessions were taught by the same workshop instructor (F.W.S.). Using a workshop format rather than incorporating these activities into existing lecture classes (where the total number of students can easily exceed 300 per class at many large public universities) is not only more practical but also more personal, and better promotes follow-up discussions on how implementation of the workshop activities can directly impact student learning.

Each workshop session was scheduled for 75 min (the two workshop activities in the first workshop collectively took between 60 and 75 min to complete). We found it useful to break the workshop up into two discrete blocks, with each activity separated by short instructor-led group discussions of Bloom’s taxonomy: reminding students why this was introduced to them in the first week of lecture, how it can be applied to learning in science, and how self-testing helps improve understanding of complex material and can lead to critical thinking skills.

At the end of the workshop, students answered three basic questions about each of the two exercises in an online survey on the class website:

- How useful did you find the exercise?
- How useful was it for you to actually do the exercise (during the workshop), rather than just hearing about it?
- How likely are you to implement the workshop activity into your own learning after the workshop?

Students were asked to respond to each question by rating their opinion on a Likert scale from 1 (least: not useful at all, not at all likely to implement) to 5 (most: extremely useful, highly likely to implement).

**DATA COLLECTION AND ANALYSIS**

The immediate effect of the workshop activities on student recall and understanding was assessed during the workshop, mainly to demonstrate those effects to the students, but also...
to allow the instructor to immediately assess the impact of the workshop on student learning. The long-term effect (10 wk) of the workshop was assessed via workshop-related questions on the cumulative final exam. Final exams were not returned to the students so students in subsequent semesters could not benefit from memorizing old exams. If the workshop was effective, we predicted that workshop participants would perform better on final exam questions related to the workshop topic than students who did not participate in the workshop (F2009 and F2008 controls). We compared the distributions of correct–incorrect answers for the individual multiple-choice exam questions, the total correct answers, the total correct lower-level answers (Bloom levels 1 and 2), and the total correct higher-level answers (Bloom levels 3 and 4) between the workshop and control groups. To control for other potential influences, such as pre-existing student achievement, we compared self-reported student GPA (at the beginning of the semester), as well as overall final exam performance (all questions) between workshop and control groups. Under the Null hypothesis (that overall student achievement did not differ between control and treatment groups), we predicted no significant differences between the workshop and control groups in GPA and total final exam scores.

For each group, we tested all variables for normality (Goodness of Fit: Shapiro Wilkes Test) using JMP 8 software (SAS Institute, Cary, NC). We used SPSS 18.0 for Mac software (2010; SPSS, Chicago, IL) for quantitative statistical analyses. Only the final exam scores were normally distributed with homogeneous variances between groups. As a result, we report the results of nonparametric tests for all analyses. For the student performance data (e.g., overall performance on exam questions relating to the workshop), the total final exam scores, as well as start-of-semester GPA [pre-existing student achievement], we used nonparametric Mann-Whitney U-tests for independent samples. This is a test for both location and shape to test for differences between distributions of ranked variables. To test whether the performance of workshop and control groups on individual exam questions was the same (null hypothesis) or different (alternative hypothesis: the workshop helped students learn), we used a Pearson χ² test. The data from the respective control groups were used to calculate the expected values for the workshop group. To correct for multiple comparisons (inflated Type I error) we applied a false discovery rate (FDR) correction (Benjamini and Hochberg, 1995) and report the adjusted P values. For paired samples (pre–post comparisons) of workshop exercises, we used the Wilcoxon-signed rank test, and for correlation tests (e.g., GPA and performance on exam questions) we used Spearman correlations. With the exception of the χ² tests (alternative hypothesis: workshop students perform better: one-tailed test), all reported results are based on two-tailed tests and significance levels of P < 0.05.

EXERCISE 1: INFORMATION-PROCESSING EXERCISE—VISUAL VERSUS AUDITORY PROCESSING

Visualization in the form of creating mental images is a fundamental cognitive process that helps student learning. We decided to demonstrate this to our students through a personal experience. For this purpose, we chose the Slippery Snakes exercise, which was developed in 1993 by Don Irwin and Janet Simons from the Development Educational Learning Institute (Des Moines, IA) to illustrate the differences between visual- and auditory-encoded memory (Irwin and Simons, 1993; Bolt, 1996). We used this exercise to demonstrate different approaches to information processing while listening and the potential impacts that the different approaches can have on student recall ability. We hoped that, by learning how to engage their visual memory instead of their auditory memory alone when taking notes in class, students would not only improve their ability to remember facts, but also to contextualize information, a prerequisite for critical thinking.

For this exercise, we divided the students at the beginning of the workshop randomly and evenly into two groups. Each group was provided with a different set of written instructions on what they were supposed to do as each phrase was read aloud. One group (the visual-processing group) was given the task of trying to form a vivid mental picture or image of the action in each phrase, and rate each phrase (on a scale of 1–10) on how simple or difficult it was to visualize. The other group (the auditory-processing group) was given the task of listening to each phrase with an emphasis on pronunciation and to rate each phrase on how simple or difficult it would be to pronounce. The students were not aware that there was a difference in instructions. Each sheet of instructions had the same numbered blanks for the students to write down their ranks as each phrase was read.

The workshop instructor read the phrases aloud slowly and deliberately, one after the other. Once all the phrases had been read, the instructor gave the students a quick unannounced quiz on the content of the phrases they had just heard. The students wrote down their answers—the subject of the phrase and an associated adjective—on the back of their instruction sheets, as the questions were read aloud at the same pace as the original phrases. After the quiz, the students were asked to score their own answers as the instructor read the correct answers. The original exercise consists of 20 phrases (Irwin and Simons, 1993), as well as questions (and answers) about those phrases. For time management reasons, we only used 12 of these.

After reassuring students that memory does not equal intelligence, following the instructions by Irwin and Simons (1993), students reported their scores individually to the instructor, who recorded them on the board in separate columns for the two groups. The differences between the two groups were immediately obvious (Figure 1), with the students in the visual-processing group (mean ± SD = 10.23 ± 1.32) scoring significantly higher than those in the auditory-processing group (mean ± SD = 5.4 ± 2.12; Mann-Whitney U-test for independent samples: U = 47, P < 0.001). After looking at the reported scores on the board and averaging for each group, the instructor asked one student from each group to reveal their set of instructions by reading aloud their respective group’s instructions. The students were generally surprised by the revelation of the difference in instructions and the improved performance of students using the visualization strategy for the processing of complex information.

Owing to these striking results, we decided to control for the possibility of a bias in the higher achieving visualization group by analyzing start-of-semester GPA (pre-existing student achievement) and student preferences for the
information exchange medium (e.g. visual, acoustic). There was no difference in GPA between the visual- and auditory-processing groups (Mann-Whitney \( U = 656.5, P = 0.494 \)). To quantify student preferences for information exchange media, we used the online VARK survey (version 7.0, 2006) with permission from Neil D. Fleming, Christchurch, New Zealand, and Charles C. Bonwell, Springfield, MO (Fleming and Mills, 1992). This survey consists of 16 questions asking students about their preferred information exchange medium in everyday situations (multiple answers possible). The output scores are visual (V), auditory (A), reading (R), and kinesthetic (K) scores (total instances that medium was chosen). Please note that the VARK survey is advertised as a “learning style” assessment (for a review of the vast amount of literature on “learning styles,” see Coffield et al., 2009); however, we used the VARK as a tool to determine a student’s current preferred medium for information transfer (regardless of which influences may have contributed to this preference). After calculating the relative contribution of each medium to the preferred information processing of each student, we tested whether the students in the auditory- and the visual-processing groups of the listening exercise differed in their preferred information-processing medium. We found no significant differences between the two groups (Visual \( U = 541.5, P = 0.659 \); Auditory \( U = 603.5, P = 0.749 \); Reading \( U = 512, P = 0.421 \); Kinesthetic \( U = 690, P = 0.167 \)). These findings emphasize that the significant differences in retention between the two groups were not due to differences in pre-existing student achievement (GPA) or prior information-processing preferences, but most likely due to the different processing instructions given to the two groups of students.

Although this exercise in itself was a valuable learning experience for our students, we found it useful to relate the value of this experience to student learning in class. For example, by visualizing the information they hear in class, students can incorporate context into the processing of this information and into their lecture notes. This strategy also works well for reading assigned textbook material. Whereas textbooks tend to emphasize isolated terms (in boldface type) over context, by visualizing the textbook material students can focus on context and higher-level processing. At the end of the workshop, most students rated this exercise as very useful (Table 1), especially doing the exercise rather than just hearing about it. Most students planned to implement what they had learned into their own information processing and note-taking during class (mean ± SD: 4.11 ± 0.9 on a scale from 1 to 5), and to a lesser extent, during their textbook reading (mean ± SD: 3.75 ± 1.1).

EXERCISE 2: SELF-TESTING EXERCISE—VISUALIZING EXISTING AND MISSING KNOWLEDGE AND UNDERSTANDING

Many students in introductory biology classes are generally unaware of self-testing as a learning strategy. As a continuation of the visualization theme, we used a visual representation of the generalized plant life cycle to illustrate the importance of self-testing for student learning. Life cycles were taught in lecture the week before the workshop sessions were conducted. During class, the course instructor (K.S.-H.) emphasized that, during all sexual life cycles, specific
structures—as defined by two characteristics: their cellularity and their ploidy—are transformed into one another via three basic processes: mitosis (cell division that maintains ploidy), meiosis (cell division that reduces ploidy), and fertilization (cell fusion that increases ploidy). After this introduction, all three generalized life cycles, including the generalized plant life cycle (Figure 2), were developed and drawn on the overhead camera in a collaborative effort between students and the instructor. Visual representations of complex information during class are extremely useful because instructors and students alike can use them to incorporate a large amount of information into a simple schematic diagram that provides context for conceptual understanding and recall. Another advantage of this approach is its utility for practicing reasoning skills, checking logical connections and relationships between different pieces of information, thereby helping students construct a more comprehensive understanding. To demonstrate how to self-test in a productive way during studying, we conducted the following exercise with our workshop students.

At the beginning of the self-testing exercise, the workshop instructor handed out one blank note card (4” × 6”) to each student. Students were asked to draw as much of the generalized plant life cycle as they could remember on one side of the note card. Students were given as much time as they needed and were encouraged to recall as many structures and processes as possible. Once all students finished the self-test, the workshop instructor had the students turn the note cards over so they could not see their initial drawings and led a group discussion of the generalized plant life cycle that served to remind students of what they knew. Students volunteered the information, which structures (names) are part of the generalized plant life cycle, which structural characteristics are used to define them, and which processes are involved in transforming one structure into the next. The instructor generated a table from student responses on the whiteboard (Table 2). Students were not allowed to take notes during this review, rather they had to remain engaged in the discussion. If needed, the instructor assisted only by revealing the number of items in each category, asking students to recall what they had learned in class until the list was completed. Please note that this list contained only the names of structures and the structural characteristics and processes to be considered. There was no discussion about the details of the plant life cycle, specifically, how the individual structures were defined or what the processes did to those structures. At this point, the instructor commended the students on their brainstorming and reminded them that they themselves had generated

---

**Table 1. Student feedback**

<table>
<thead>
<tr>
<th>Exercise</th>
<th>Learning Mean</th>
<th>SD</th>
<th>Doing Mean</th>
<th>SD</th>
<th>Implementing Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Information</td>
<td>4.20</td>
<td>0.85</td>
<td>4.38</td>
<td>0.87</td>
<td>4.11</td>
<td>0.91</td>
</tr>
<tr>
<td>Self-testing</td>
<td>3.90</td>
<td>1.03</td>
<td>4.04</td>
<td>1.04</td>
<td>3.93</td>
<td>0.98</td>
</tr>
</tbody>
</table>

Student responses to three survey questions regarding the two workshop exercises, based on a Likert scale from 1 (least) to 5 (most). On average, students found the workshop, including its practical aspects, very useful and reported that they were very likely to implement what they had learned in their learning.

---

**Table 2. Self-testing group review**

<table>
<thead>
<tr>
<th>Structure names</th>
<th>Structure characteristics</th>
<th>Processes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sporophyte</td>
<td>Cellularity (unicellular, multicellular)</td>
<td>Mitosis</td>
</tr>
<tr>
<td>Gametophyte</td>
<td>Ploidy (haploid, diploid)</td>
<td>Meiosis</td>
</tr>
<tr>
<td>Spore</td>
<td></td>
<td>Fertilization</td>
</tr>
<tr>
<td>Gamete</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zygotte</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In all workshop groups, students generated this overall list of structure names, structure characteristics, and processes at work in the general plant life cycle. This list was compiled on the white board by the workshop instructor as students generated this information. Only the terms, not their relationships, were listed during the review (students had to identify the specific characteristics of each structure and the processes that transformed one structure into another on their own during their postreview self-test).
Figure 3. Pre- and postreview self-testing scores. Student performance ($N = 93$) is shown as mean (%) possible score (possible process score: 5 points; possible structure score: 15 points). Students performed significantly better in the postreview self-test than in the prereview self-test for both processes and structures (Wilcoxon signed rank test for paired samples (combined): $Z < -2.58; P < 0.01$; for processes only: $Z = -2.585; P = 0.01$; for structures only: $Z = -7.152, P < 0.001$).

For this table, so as a group they clearly knew more than they initially thought they did, and that they should practice this approach while studying.

For the final step in this exercise, the list was removed from the whiteboard, and the students were asked to use the remaining blank side of the note card to diagram the plant life cycle again, using as many of the structures and processes as they could. At the end of the workshop, students were asked to identify which was their first (prereview) and second (postreview) attempt and to hand in their note cards.

For the purpose of this study, we scored how well students did in their two attempts (this did not affect student grades). Their note cards were scored by assigning three points for each structure (one point each for the name and the two characteristics that define that structure) and one point for each correctly placed process for a total of 20 possible points (structures = 15 points; processes = 5 points). Both the prereview and the postreview attempts were graded in the same manner for each student. To allow direct comparison of student performance on structures and processes, we transformed the data to percent of total points possible for each category (structures and processes).

**Student Performance on Workshop Exercise**

The life cycles produced during the first (prereview) self-test were largely incomplete, and students were better able to recall processes (mean ± SD = $1.978 ± 1.56$ of 5) than structures and their characteristics ($3.66 ± 3.63$ of 15). In contrast, during their second (postreview) attempt, students were able to produce a much more complete and accurate diagram of the plant life cycle ($2.43 ± 1.57$ of 5 processes and $7.47 ± 4.21$ of 15 characteristics; Wilcoxon signed rank test for paired samples: $Z < -2.58; P < 0.01$). Between the first and second attempt, the correct placement of structure names and their characteristics improved by 56% and 47%, respectively, whereas correct placement of processes only improved by 17% (Figure 3). This result is partly due to the fact that students struggled more to name the correct structures and their characteristics (of 15, 24% correct) on their first attempt, than they did in naming the processes (of 5, 40% correct). In their second attempt, students scored on average ~50% correct for both structures (associating names with the correct cellularity and ploidy state) and processes (making sure that the process matched the change in structural characteristics). This represents a significant improvement for both categories (Figure 3) without looking up the complete life cycle.

**Student Behavior during Self-Test**

The workshop instructor observed students exhibiting “helpless” behavior during the first self-test: Students tended to give up at the first point in the cycle where they encountered a structure or process they could not recall, and did not attempt to start over from a different point and work the problem from there.

During their second self-test, the students who “gave up” during their first attempt were able to complete more of the life cycle, despite the fact that the placement of structures and processes and their logical connections were not practiced (or revealed) during the group review. When asked about the outcome of their second attempt after the exercise, the
students generally attributed their higher success to the “help” by the instructor, even though they themselves generated the list of terms during the group review and applied the list to their second attempt, underestimating their own role in the process. We have found that this response is fairly typical for students not used to self-testing, including students who go through this exercise individually (Stanger-Hall, unpublished data).

Student Feedback

Many students found the self-testing exercise very useful (Table 1). Most students expressed intentions to implement self-testing as a study strategy was positively correlated with how well they did in their first (Spearman \( \rho = 0.261, P = 0.02 \)) and second attempt (Spearman \( \rho = 0.278, P = 0.013 \)) (Figure 4) to draw and label the plant life cycle.

Student Performance on the Final Exam

There were seven life cycle-related questions on the final exam in 2009 (14 points of 270), and four of these questions were also represented in 2008 (8 points of 270). Two of the questions (Questions 2 and 3) were higher-level application questions. We assessed overall performance on life cycle questions, performance on lower-level (Bloom levels 1 and 2) and higher-level (Bloom levels 3 and 4) life cycle questions, as well as on the individual life cycle questions separately. For comparison, we also assessed overall student performance on the final exam (all questions). To address the question of how the workshop affected students of different ability levels, we used two approaches: 1) We assessed whether GPA was correlated with performance on workshop-related questions on the final exam and how this correlation (if any) compared with the correlation between GPA and final exam performance overall; in addition, 2) we determined the median GPA for the workshop group (3.4) and divided the workshop students into two groups (low GPA = GPA < 3.4, and high GPA = GPA > 3.4; students with GPA = 3.4 \[ N = 12 \] were excluded from this analysis to create two distinct groups). We used the ratio of points earned for life cycle questions and total points earned for the other questions on the final exam to assess whether the potential workshop benefits (as assessed by the life cycle questions) differed between low- and high-GPA students.

If student achievement (GPA) did not influence workshop benefits (Null hypothesis), we expected to find 1) no significant correlation between GPA and exam performance and 2) no difference between the low-GPA and the high-GPA groups in how well the students did on the life cycle questions relative to all the other exam questions on the final exam.

Control 2009

The advantage of using this control group is that students were exposed to the exact same lectures (style, delivery, and examples used) as the workshop group. The disadvantage of this control group is that cross-talk between students (in the same class) may have occurred, that is, students may have learned from each other during study sessions. Despite this possibility, the workshop group scored significantly higher than the control group (\( U = 4665.5, P = 0.001 \)) on the life cycle questions of the final exam (Figure 5A). Compared to the control group, the workshop students performed significantly better (all \( P \) values are reported after FDR correction for multiple comparisons) on Question 1 (lower-level: Pearson \( \chi^2 = 3.739, \text{borderline significant at } P(1) < 0.05 \)), Question 2 (higher-level: Pearson \( \chi^2 = 7.464, P(1) < 0.05 \)), and Question 7 (lower-level: Pearson \( \chi^2 = 12.158, P(1) < 0.01 \)) of the life cycle question series on the final exam. In fact, workshop students tended to perform better than the control group on all life cycle questions (Figure 5B), but the differences for the other questions were not significant. The overall performance on the final exam was not significantly different (\( U = 3723, P = 0.809 \)) between groups, and there was no significant difference in students’ pre-existing achievement, as measured by self-reported GPA (workshop GPA = 3.353 ± 0.377, Fall 2009 control GPA = 3.283 ± 0.5; Mann-Whitney \( U = 3041.5, P = 0.438 \)).

Control 2008

The advantage of using this control group is the absence of cross-talk between students, but the disadvantage is that the lectures (not content, but style and delivery) may have varied between the 2 yr, even though the classes were taught by the same instructor (K.S.-H). Overall, workshop students scored significantly better on the life cycle questions of the final exam (\( N = 4 \) questions: \( U = 4448, P = 0.004 \)) than the control group (Figure 5A). The workshop group performed significantly better on three of the four individual questions in the life cycle question series on the final exam (\( P \) values are reported after FDR correction for multiple comparisons): Question 1 (lower-level: Pearson \( \chi^2 = 11.91, P(1) < 0.01 \)), Question 2 (higher-level: Pearson \( \chi^2 = 16.089, P(1) < 0.01 \)), and Question 4 (lower-level: Pearson \( \chi^2 = 5.149, P(1) < 0.05 \)).

Figure 4. Student motivation to implement self-testing as a study tool. Students who reported a higher motivation to implement self-testing as a study tool after the exercise had performed better (points out of 20) during the postview self-test than students who reported a lower motivation. For example, students with the highest level of motivation to implement self-testing (5: highly likely to implement, \( N = 29 \)) scored 11 points (55%) in their postview self-test; the student with the lowest level (not at all likely to implement, \( N = 1 \)) scored 4 points (20%).
Workshop students tended to perform better than the control group on all life cycle questions (Figure 5B), but the difference for Question 3 was not significant. The overall performance on the final exam was significantly different ($U = 2223, P < 0.001$) between the workshop and the control group, but the Fall 2008 control group performed significantly better than the Fall 2009 workshop group. This makes the significantly better performance of the workshop students on the life cycle questions even more relevant. There was no significant difference in students’ pre-existing achievement (workshop GPA = 3.353 ± 0.37, Fall 2008 control GPA = 3.415 ± 0.34; Mann-Whitney $U = 3041.5, P = 0.438$).

**Achievement Effects within the Workshop Group**

There was no significant correlation between GPA and overall performance on the life cycle questions on the final exam within the workshop group (Spearman’s $\rho = 0.219, P = 0.055$), but GPA was positively correlated with overall final exam performance (Spearman’s $\rho = 0.552, P < 0.001$). In addition, low-GPA students did not gain significantly more or less from the self-testing exercise during the workshop than high-GPA students (relative to their performance on the other final exam questions; Mann-Whitney $U = 577, P = 0.383$). As a result, we can conclude that the self-testing exercise benefited students across achievement levels.

**SAMPLES OF STUDENT FEEDBACK**

Both workshop exercises were designed to help students learn, specifically to illustrate how visualization techniques can be used to improve information processing during class and reading, as well as for self-testing as a means to review and study. When students were asked, “What was the best or most useful part of the workshop?” the majority (52%) of respondents listed the information-processing exercise, 10% listed the self-testing exercise, and 36% had no preference (most of these liked both). The following are sample student responses to this question, demonstrating the wide variability in responses:

“The self-testing part. I had never thought about how I would go about making up my own questions but now I know how.”

“Learning to visualize the concepts that we talk about in class, not just listen to them.”
“The self test really made me see how much I did, and didn’t know.”

“It showed me the difference between the effectiveness of visual and auditory learning styles. Also, it showed me how to focus my learning on ways that will help me remember it in the long-run.”

“I though [sic] participating in the demonstration was very helpful and allowed us to actually see the effectiveness of the process.”

“Realizing that if I better visualize my notes as I am taking them/studying them, I can better retain and understand the information.”

**DISCUSSION**

There is growing recognition that visualization is an essential thinking skill in science and science education (Matthewson, 1999; Gilbert, 2005, 2008; Schönborn and Anderson, 2006), and we have shown in this study that visualization in the form of mental images (during information processing) or external representations of complex information (during self-testing) can help college students learn. The next logical step is the large-scale implementation and further evaluation of this workshop for large introductory biology (or other science) classes, including development of more learning activities that allow students to practice their visualization skills during class and at home.

In the sciences, instructors and students alike can use visual representations to incorporate complex information into a simple schematic diagram that provides context for conceptual understanding and recall. Another advantage of creating visual representations is their utility for practicing reasoning skills, for example by checking logical connections and relationships between different pieces of information. This helps students construct a more holistic and comprehensive understanding, as compared with simply describing the relationships in writing (Gobert and Clement, 1999). Agarwal et al. (2008) showed that additional testing enhances long-term retention, and the self-testing approach described here would serve that purpose.

In general, during any self-testing exercise, students should write down all the information on a given topic that they can remember (notes and text closed) and organize it on paper (e.g., in the form of a diagram). After identifying possible gaps and missing pieces, students should then brainstorm what else they might know about this or closely related topics, making a list as they go. This process is the first step in identifying possible candidates for the missing pieces to fill gaps in their diagram, the goal of this exercise. The next step is to define everything on that list, and possibly compare—contrast similar terms, structures, or processes. This process further helps the students learn the reviewed material, recognizing possible connections to their diagram, and deducing some of the missing pieces and relationships—a strategy many students already use when solving Sudoku puzzles in the school newspaper before class but fail to employ during studying. When students engage in self-testing during studying, it is important that they do so without immediately going to their notes and textbook for help. Practiced as described, self-testing can be a very effective and cognitively active form of learning (Karpicke, 2009; Karpicke et al., 2009), a prerequisite for critical thinking that helps students review information while providing a context for this information at the same time. In contrast, looking up information or asking for answers without reflection only leads to passive memorization and helpless behavior in a testing situation.

It is noteworthy that the plant life cycle self-test in the workshop group took place during the week after students drew the life cycle (guided by the instructor) in class. There was a wide range in recall (0–20 of 20 points) between students during the first (prereview) self-test, but on average students struggled to remember (5.46 of 20 points), and many students gave up rather than trying other approaches. This helpless behavior seems to be a characteristic response to a challenge by students without self-regulation skills, who tend to be low-achieving students (Isaacson and Fujita, 2006). Isaacson and Fujita (2006) suggested that such learned helplessness might be the consequence of high performance expectations that remain unadjusted by actual performance. By overestimating their abilities and not adjusting their study strategies, these students will decrease their efforts over time and ultimately give up and fail (Isaacson and Fujita, 2006).

The motivation of students to implement self-testing as a study strategy was positively correlated with how well they did in their first (prereview) and second (postreview) attempt to draw and label the plant LC. Difficult tests are better at differentiating between low and high achievers, and self-tests that also assess higher-level skills are generally more challenging than self-tests that only require lower-level skills (Isaacson and Fujita, 2006). This suggests that, although it is important to challenge students during a self-testing exercise, some success should be built in to convince students of the value of the exercise, and to move from helpless behavior to filling in as many pieces and connections as possible, even if skipping of steps is required. This can be best achieved by using recent material from class (as in this exercise) or by asking students to review their notes (or a textbook passage) on previously untested material. We recommend asking students whether the material “makes sense” to them before the start of a self-testing exercise. Students routinely will state that the notes (text) make sense but then struggle when asked to replicate the information on paper during a (closed-book) self-testing exercise. This apparent conflict between self-assessment and performance leads to the insight that “making sense” is not the same as understanding the material, which is a key step in convincing students to be receptive to making changes in their study routine.

The reviewing and self-testing skills are not only important for students during studying (Karpicke, 2009) but also during assessments (quizzes and exams). Rather than giving up immediately when not remembering an answer to a question, or when having to apply, analyze, synthesize, or evaluate the learned material to answer a higher-level question, these self-testing skills give students the opportunity during an exam (or in their everyday life) to figure out what they do not remember, or how to answer a critical thinking question. Because critical thinking is a life-long learning skill, active self-testing should be an integral part of studying and practicing for life-long learning, rather than immediately giving up and looking to others (notes, text, teacher) for answers when faced with a “hard” or “unfair” question. Unfortunately, the latter approach is far more common but promotes only
short-term memory, not long-term retention and higher-level thinking.

The benefit of this self-testing exercise for students was at least threefold: 1) it helped students review the class material actively, thereby helping them distinguish what they knew and what they did not; 2) it helped them place what they remembered in context, a prerequisite for identifying mistakes or misconceptions; and 3) it provided them with the tools to deduce missing pieces and to practice their critical thinking skills.

The written (documented) aspect of reviewing and self-testing is in our opinion crucial to achieve these outcomes. Although many students insist that they “do self-testing in their head” (Stanger-Hall, unpublished data), it is key for the success of this self-testing routine that a student documents her/his knowledge and knowledge gaps on paper (as an external visual representation) so they can be organized visually and their (existing and missing) connections (context) becomes apparent to the student. Students who self-test “in their head” usually just go through lists of terms and processes without recognizing their overall relationships and connections (Stanger Hall, personal observation), whereas students who document their self-testing on paper better recognize possible connections and can inspect them for flaws in logic afterwards.

STUDENT FEEDBACK

Overall, students perceived both workshop exercises positively but rated the information-processing exercise as more useful than the self-testing exercise. In addition, they reported that they were somewhat more likely to implement what they learned during the information-processing exercise in their own learning (Table 1). This preference may have been influenced by several factors. For example, the information-processing exercise may simply have been more fun and was therefore better received by the students (as suggested by the 56% preference rating). It could also in part be due to the results being immediate for the information-processing exercise during the workshop, whereas the self-testing exercise and its results were not as immediate since it required grading by the workshop instructor, despite obvious improvements between the first and second self-tests (Figure 3).

Regardless of overall preference, students generally valued the visualization techniques learned in this workshop. For example, several students contacted the workshop instructor to relay how they had tried implementing one or both of these techniques into their own note-taking and study practices. According to those students, there was an immediate positive effect in terms of their performance on the next exam.

SUMMARY AND CONCLUSIONS

We feel that the workshop format, working in multiple, smaller groups (up to 12 students), is particularly useful for demonstrating the efficacy of these techniques and for promoting discussion of how these techniques might be used to enhance student learning. Although both exercises could potentially be done in larger groups, in our experience, larger groups of students take more time to lead through the exercises, and large group sizes are less effective at promoting individual interaction and discussion than smaller groups. We also feel that using a separate workshop instructor adds credibility to the lecture instructor’s efforts to teach students critical thinking skills, by providing a second “independent” proponent of this approach. However, there is an inherent increase in implementation cost with large college classes because multiple workshop instructors would need to be hired and trained for this small-group approach. This cost needs to be justified by significant gains in student learning, the ultimate purpose of college instruction.

Our small-scale study (within a large class) convincingly showed that students who use visualization during information processing remember more during a recall test immediately following the exercise than students who do not. In addition, our assessment of student learning on the cumulative final exam demonstrated that: 1) self-testing led to long-term learning gains for the self-testing topic; 2) these learning gains apply to both lower- and higher-order thinking skills; and 3) students across achievement levels (GPA) benefit. Thus, all the previously stated criteria for large-scale implementation and further evaluation have been met.

Ideally, the self-testing exercise (taught to all students in class) would be followed by a series of weekly study groups where visualization and self-testing are practiced by students and applied to all class topics. These weekly study groups could even be facilitated by trained peers, recruited from the same class or from previous classes (Crouch and Mazur, 2001; Stanger-Hall et al., 2010). This practice will help achieve the ultimate goal of teaching students self-regulating techniques, encouraging them to use what they learned from the workshop exercises for all other class topics and to do so while studying on their own.

ACKNOWLEDGMENTS

This workshop was implemented through funding by a Faculty Research Grant (#790) to K.S.-H. from the University of Georgia Research Foundation. This study was conducted under the guidelines of IRB # 2007–10197-4. Thanks to P. Lemons, P. Brickman, N. Armstrong, and two excellent (anonymous) reviewers for constructive comments on an earlier version of the manuscript. This is a publication of the UGA Science Education Research Group.

REFERENCES


Lord T, Baviskar S (2007). Moving students from information recitation to information understanding: exploiting Bloom’s taxonomy in creating science questions. JCScT 36, 40–44.
Article

Combining Peer Discussion with Instructor Explanation Increases Student Learning from In-Class Concept Questions

M.K. Smith,*†‡ W.B. Wood,*† K. Krauter,* and J.K. Knight*†

*Department of Molecular, Cellular, and Developmental Biology and †Science Education Initiative, University of Colorado at Boulder, Boulder, CO 80309

Submitted August 18, 2010; Revised December 1, 2010; Accepted December 9, 2010
Monitoring Editor: Diane Ebert-May

Use of in-class concept questions with clickers can transform an instructor-centered “transmissionist” environment to a more learner-centered constructivist classroom. To compare the effectiveness of three different approaches using clickers, pairs of similar questions were used to monitor student understanding in majors’ and nonmajors’ genetics courses. After answering the first question individually, students participated in peer discussion only, listened to an instructor explanation only, or engaged in peer discussion followed by instructor explanation, before answering a second question individually. Our results show that the combination of peer discussion followed by instructor explanation improved average student performance substantially when compared with either alone. When gains in learning were analyzed for three ability groups of students (weak, medium, and strong, based on overall clicker performance), all groups benefited most from the combination approach, suggesting that peer discussion and instructor explanation are synergistic in helping students. However, this analysis also revealed that, for the nonmajors, the gains of weak performers using the combination approach were only slightly better than their gains using instructor explanation alone. In contrast, the strong performers in both courses were not helped by the instructor-only approach, emphasizing the importance of peer discussion, even among top-performing students.

INTRODUCTION

Active-learning activities can significantly increase student learning in biology courses (Udovic et al., 2002; Kitchen et al., 2003; Knight and Wood, 2005; Freeman et al., 2007; Walker et al., 2008). Among the many kinds of such activities that are practical in large lecture classrooms, in-class concept questions using personal response systems or “clickers” have received recent attention (e.g. Wood, 2004; Caldwell, 2007). Typically, instructors pose multiple-choice questions requiring application of a recently presented concept at several time points during a class, and students record their answers using clickers. In addition to breaking up lectures into smaller chunks, concept questions provide students with opportunities to practice solving problems and monitor their understanding during class. Recent work in cognitive psychology has shown that frequent assessment of students in this manner has a powerful impact on both learning and retention (reviewed in Roediger et al., 2010).

In connection with in-class concept questions, instructors often use an approach called peer instruction, which encourages students to verbalize their thinking and interact with their peers to arrive at an answer (Mazur, 1997). In one commonly used mode, students first answer a concept question individually, discuss the question with their peers, and then revote before the answer to the question is revealed. The instructor then explains the question and often shows a...
histogram of the student responses, which gives both instructors and students immediate feedback on how well a concept is understood.

Many instructors report that the frequency of correct answers increases after peer discussion (Mazur, 1997; Crouch and Mazur, 2001; Knight and Wood, 2005; Smith et al., 2009). Two alternative hypotheses could explain this observation: 1) active engagement of students during discussion with peers leads to increased conceptual understanding, resulting in improved performance on the revote, or 2) students do not necessarily learn from the discussion, but simply choose the answer most strongly advocated by neighbors they perceive to be knowledgeable.

In a previous study (Smith et al., 2009), strong support was obtained for the first hypothesis using matched pairs of in-class concept questions that addressed the same concept and required similar reasoning but had a different story line. Students answered the first question of a pair (Q1) individually. After a few minutes spent discussing their responses in small groups, they revoted on Q1. Students then answered a second question (Q2) individually, and only then were the answers and the histograms for both questions revealed and discussed. Subsequent tracking of student responses using the clicker software showed that students who changed their answers to Q1 from incorrect to correct after discussion performed on average much better on Q2 than students who did not change their answers. Moreover, on the more difficult questions, performance on both Q1 after discussion and Q2 increased markedly, even for groups in which no student initially answered Q1 correctly, indicating that the process of discussion itself rather than the influence of knowledgeable peers could lead students to increased understanding.

Although this study demonstrated that peer discussion had a positive effect on student learning, it did not attempt to compare peer discussion with explanation from the instructor or an alternative activity between Q1 and Q2. In informal discussions with instructors, we learned that some instructors skip peer discussion, believing that their explanation of a clicker question answer will be clearer, more efficient, and more informative than what students are likely to hear in conversations with each other. However, the constructivist viewpoint supported by the above study of Smith et al. (2009) predicts that the process of verbalization and discussion could promote understanding more effectively than even a clear instructor explanation. In addition, grappling with a question in discussions with peers could enhance the learning value of a subsequent explanation by the instructor (Schwartz and Bransford, 1998). To explore the merits of these alternative views, we applied a modification of our earlier protocol using matched pairs of questions (referred to in that study as isomorphic questions; Smith et al., 2009) to ask which of the following three presentation modes leads to the greatest improvement in student performance: having a peer discussion, listening to an instructor explanation, or engaging in peer discussion followed by an instructor explanation (the combination approach). We evaluated the effects on student learning gains in two classes, genetics for majors and nonmajors, as well as for three different ability groups of students classified as strong, medium, and weak clicker performers.

**METHODS**

**Student Demographics**

This study was conducted in an undergraduate introductory genetics course required for majors (Fall semester of 2008) and a genetics course for nonmajors (Fall semester of 2009) (student demographics shown in Table 1). These courses were taught in the Department of Molecular, Cellular, and Developmental Biology (MCDB) at the University of Colorado, Boulder, by two of the authors: K.K. (majors) and J.K.K. (nonmajors). Both courses met for three 50-min sessions per week, and student grades in both courses were based on a similar distribution of points (Table 2).

**Instructional Modes and Experimental Protocols**

In both courses, an average of four in-class concept questions were asked per class, and approximately half the class periods included matched-pair questions that were used in this study. Even though all the in-class concept questions were awarded only participation points, students had an incentive to do so.

| Table 1. Demographic information on students who participated in this study |
|-----------------|-----------------|
| Category        | Majors’ course  | Nonmajors’ course |
| Gender          | 41% female, 59% male | 66% female, 34% male |
| Year in college | 7% freshman, 31% sophomore, 29% junior, 26% senior, 7% other | 36% freshman, 37% sophomore, 11% junior, 13% senior, 3% other |
| Major           | 55% biology; 98% indicated they hoped to pursue a career related to science | 10% biology* |
| Grade distribution in genetics course | 26% A, 39% B, 22% C, 9% D, 4% F | 37% A, 36% B, 23% C, 3% D, 0% F |

*There are three biology majors at CU-Boulder: Molecular, Cellular, and Developmental Biology (MCDB); Ecology and Evolutionary Biology (EBIO); and Integrative Physiology (IPHY). These students are EBIO and IPHY majors.

---

**Table 2. Distribution of grading points in the majors’ and nonmajors’ genetics courses**

<table>
<thead>
<tr>
<th></th>
<th>Majors (%)</th>
<th>Nonmajors (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exams</td>
<td>71</td>
<td>61</td>
</tr>
<tr>
<td>Homework</td>
<td>17</td>
<td>16</td>
</tr>
<tr>
<td>Clicker participation</td>
<td>10</td>
<td>12</td>
</tr>
<tr>
<td>Additional participation (surveys, reflections)</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>Group project</td>
<td>N/A</td>
<td>11</td>
</tr>
</tbody>
</table>
well on these questions because they were told that clicker questions gave them practice for the exams.

To determine whether engaging in peer discussion, listening to an instructor explanation, or participating in a combination of peer discussion followed by instructor explanation was more effective for student learning, we followed each of three different modes of the experimental protocol outlined in Figure 1, using matched pairs of questions that test genetics concepts as described in the Introduction. Both courses have similar learning objectives and questions on similar topics, but the question pairs used were different because of the higher level of detail appropriate for the majors course (examples shown in Figure 2).

Our protocol included three types of questions: Q1, Q1ad, and Q2 (Figure 1). In all three modes, students answered Q1 individually to provide a measure of student understanding after listening to a lecture on the topic. The three modes were as follows:

- In the peer-discussion mode, students voted on Q1 after discussion (Q1ad). After recording their vote, they were told the correct answer to the question, but no additional explanation was given.
- In the instructor mode, after students had answered Q1 individually, the instructor asked the students to volunteer their reasons for selecting specific answers, explained the solution to Q1, and answered any student questions.
- For the combination mode, after students answered Q1 individually, they discussed the question with their neighbors and then voted on the same question again (Q1ad), just as in the peer-discussion mode. The instructor then asked the students to volunteer their reasons for selecting specific answers, explained the solution to Q1, and answered any student questions, as in the instructor-explanation mode.

In all three modes, students then voted individually on Q2.

After all of the Q2 votes were recorded, the instructor explained the solutions to Q1 (for the peer-discussion mode) and Q2. Histograms of student responses to Q1 and Q2 were shown only after the Q2 vote, because showing histogram results can bias student responses and influence the student discussion (Perez et al., 2010).

To compare these instructional modes in a normal classroom setting, there were no time limits placed on the instructor explanations or student voting. Both instructors generally let student voting continue until 75–80% of the students had recorded their vote, encouraged the remaining students to vote, and then stopped the voting 10–20 s later. Consequently, mean amounts of time devoted to consideration of Q1 for the different modes varied (Table 3). For both the majors’ and the nonmajors’ courses, more time on average was spent considering answers to Q1 in the combination mode than in the peer-discussion or instructor-explanation modes, as might be expected (implications of these differences are explored in the Discussion).

All three modes of the protocol were used in the majors course, but only the instructor-explanation and combination modes were used in the nonmajors’ course.

Description of Matched Pair Questions

To minimize any bias toward writing an easier Q2 question, the questions in each pair were randomly assigned to be Q1/Q1ad or Q2 after they were written (Smith et al., 2009). Which of the three modes of presentation to use was also randomly determined for each question pair. Q1/Q1ad and Q2 were assigned to a mode of presentation and inserted into the slide presentations shortly before class to minimize the possibility of altering the lecture to favor one mode of presentation over another.

Both instructors agreed that questions where the individual Q1 vote was greater than 80% correct were insufficiently challenging and left little opportunity for gains in learning; these questions were not included in this study. Although we intended all questions to be challenging, one peer-discussion and two combination questions had Q1 scores of >80% correct in the majors’ course. In the nonmajors’ course, three instructor-explanation questions had Q1 scores of >80% correct.

After the course was completed, all the question pairs used were judged for similarity by two independent reviewers, who did not have access to the student performance results.
Combining Discussion and Explanation

Figure 2. Examples of the Q1/Q1ad and Q2 question pairs used in this study. The correct answers are underlined.

The reviewers were familiar with the content of the genetics courses and had participated in an earlier study that used matched pairs of questions to measure the benefits of peer discussion (Smith et al., 2009). These reviewers were asked to judge whether they thought the question pairs in this study were testing the same concept. Data from five question pairs, three from the majors’ course and two from the nonmajors’ course, were removed from the data set, because two independent reviewers judged them as not testing identical concepts. Individual responses were also removed from the data.
Table 3. Average elapsed times between the end of individual Q1 and the start of individual Q2 for the three modes of discussing Q1

<table>
<thead>
<tr>
<th>Mode of discussing Q1</th>
<th>Majorsa</th>
<th>Nonmajorsb</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peer discussion</td>
<td>2 min 54 s</td>
<td>N/A b</td>
</tr>
<tr>
<td></td>
<td>(17 s)</td>
<td></td>
</tr>
<tr>
<td>Instructor explanation</td>
<td>1 min 54 s</td>
<td>3 min 19 s</td>
</tr>
<tr>
<td></td>
<td>(25 s)</td>
<td>(45 s)</td>
</tr>
<tr>
<td>Combination</td>
<td>4 min 42 s</td>
<td>5 min 5 s</td>
</tr>
<tr>
<td></td>
<td>(53 s)</td>
<td>(60 s)</td>
</tr>
</tbody>
</table>

aThe SEM is in parentheses.
bThe peer discussion alone mode was not used in the nonmajors’ course (see text).

set if a student did not answer all questions in a question set (e.g., answered Q1 and Q1ad but not Q2).

The remaining 32 question pairs were also rated for cognitive level according to Bloom’s taxonomy (Bloom and Krathwohl, 1956) by two independent reviewers not associated with the course who are experts at these rankings (Crowe et al., 2008). The raters were given all 64 questions in a random order and were not told which questions were matched-pair sets. Both raters independently determined that the average level of the questions was 3 (application level). The two raters concluded that 87% and 81%, respectively, of the Q1–Q2 pairs were at the same Bloom’s level. For the question pairs that did not match, 60% of the time the raters concluded that Q2 was at a higher level than Q1.

The data set included responses from 150 students in the majors class and 62 students in the nonmajors class. These students answered at least one complete set of questions in each of the different modes. Table 4 shows the mean number of questions answered for each protocol mode.

Data Analysis

The change in learning between question pairs was computed for each individual student using a modified version of the Hake normalized gain formula (Hake, 1998) known as normalized change \(<c>\) (Marx and Cummings, 2007). Normalized change values provide a measure of how much a student’s performance increases compared with that individual’s maximum possible increase. When calculating the mean normalized change between Q1 and Q2 over all question pairs for a given student, the following formula was used when an individual’s mean Q2 score was higher than the mean Q1 score (most cases): \(<c> = 100(\text{mean } Q2 - \text{mean } Q1)/(100 - \text{mean } Q1)\). Alternatively, if an individual’s mean Q1 score was higher than the mean Q2 score, \(<c> = 100((Q2 - Q1)/Q1)\), was used. In cases where an individual’s mean Q1 score and the mean Q2 score equaled either 100 or 0, the response for that student was removed from the data set, because otherwise \(<c>\) would be recorded as 0. Significant differences between mean \(<c>\) values between two populations cannot be determined because they are nonlinear computed quantities that are not normally distributed. Instead, the standard error measurements on reported \(<c>\) values are used to provide a coarse depiction of the spread of values (Marx and Cummings, 2007).

All statistical analyses were performed with SPSS (SPSS, Chicago, IL) or Excel (Microsoft, Redmond, WA). Item discrimination values (D) were calculated by rank, ordering students by their overall Q1 percent correct score. The top 27% and the bottom 27% of students in the majors’ and nonmajors’ courses were compared for this analysis. For each Q1 question, the following formula was used: D = \((RU - RL)/(1/2T)\). RU is the number of students in the upper group who answered correctly, RL is the number of students in the lower group who answered correctly, and T is the total number of students included in the item analysis (Gronlund, 1976). The average item discrimination values for Q1 questions were then calculated for each protocol mode in the majors’ and nonmajors’ courses.

Institutional Review Board Protocols

Approval to evaluate student clicker responses (exempt status, Protocol No. 0108.9) and end-of-year survey responses (expedited status, Protocol No. 0603.08) was granted by the Institutional Review Board, University of Colorado, Boulder.

RESULTS

Q1 Questions Have Equivalent Difficulty and Adequate Item Discrimination for Question Pairs Administered in Each of the Three Modes

The mean percentages of correct individual Q1 answers were not statistically different between the three different protocol modes for the majors (Figure 3A, repeated measures analysis of variance, p > 0.05). Similarly, for the nonmajors, the percentages of correct individual Q1 answers were not statistically different between the instructor-explanation and combination modes (Figure 3B, paired t-test, p > 0.05). Also, the average item discrimination values (D) for the Q1 questions were greater than 0.3 for all protocol modes in both the majors’ and nonmajors’ courses (Table 5). Questions with D values above 0.3 are generally considered good discriminators of the top and bottom students (Ebel, 1965).

The Learning Gains between Q1 and Q1ad Were Similar in Both Modes That Involved Peer Discussion

An initial measure of learning through peer discussion was calculated by recording mean student performance on individual Q1s and the same questions after discussion (Q1ad) (Figure 3). In all cases, students’ mean performance on Q1ad was significantly higher than on Q1 (dependent t-test, p < 0.05). In addition, we calculated the mean normalized change \(<c>\) between Q1 and Q1ad for each individual student. In
Combining Discussion and Explanation

The Effect of Different Discussion Modes on Percentage of Correct Answers

Figure 3. Effects of three different modes of discussing Q1 on percentage of Q1ad and Q2 correct answers in the majors’ (A) and nonmajors’ (B) genetics courses. Performance results were averaged for each individual before computing the means shown. The SEMs are shown with error bars.

The majors' course, this value was 41.5% (±3.4) for the peer-discussion mode and 37.2% (±3.4) for the combination mode. The similarity of these values suggests that peer discussion resulted in similar performance improvement for both these modes. For the nonmajors, the mean <c> between Q1 and Q1ad for the combination mode was somewhat higher at 56.9% (±5.5%).

The Combination Mode Led to Larger Learning Gains between Q1 and Q2 Than Either Peer Discussion or Instructor Explanation Alone

In the majors’ course, when all three modes of the protocol were compared for student performance on Q1 and Q2, the mean percentage of correct answers was higher for Q2, indicating that performance improved in all three modes (Figure 3, dependent t-test, p < 0.05 in all cases). Similarly in the nonmajors’ course, the mean percentage of correct answers was significantly higher for Q2 than for Q1, indicating that performance improved in both the instructor-explanation and combination modes (Figure 3, dependent t-test, p < 0.05 in all cases).

Two principal findings from comparisons of learning gains are presented in Figure 4, which shows the Q1-to-Q2 mean <c> values for each mode in both the majors’ and nonmajors’ genetics courses. First, in the majors’ course, the peer discussion and instructor-explanation modes resulted in similar mean <c> values, suggesting that each of these modes alone is equally effective. Second, in both courses, the combination mode resulted in strikingly higher <c> values than either of the other modes alone.

Table 5. Mean item discrimination (D) values for Q1 questions in the different protocol modes

<table>
<thead>
<tr>
<th>Method of discussing Q1</th>
<th>Majors</th>
<th>Nonmajors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peer discussion</td>
<td>0.42</td>
<td>N/A*</td>
</tr>
<tr>
<td>Instructor explanation</td>
<td>0.39</td>
<td>0.39</td>
</tr>
<tr>
<td>Combination</td>
<td>0.37</td>
<td>0.55</td>
</tr>
</tbody>
</table>

*The peer discussion alone mode was not used in the nonmajors’ course (see text).

The Combination Mode Results in the Largest Gain in Learning for All Ability Levels of Students

To determine whether a certain instructional mode is better for students who tend to do well or poorly on in-class concept questions, mean Q1 percent correct scores for all instructional modes were calculated for each student. Then the students in each course were divided into three groups based on these scores, in which weak, medium, and strong clicker performers were designated as having mean Q1 scores of <33.3%, 33.3–66.6%, and >66.6%, respectively. Table 6 shows the percentage of students who fell into each category. The majority of students in both courses fell into the medium clicker performer category.

Figure 5 shows the learning gains represented by average Q1-to-Q2 <c> scores for the weak, medium, and strong clicker performers for all three instructional modes in both courses. In the majors’ course (Figure 5A), the combination mode was clearly most effective for all three groups of students. For the weak and medium groups, the peer-discussion and instructor-explanation modes appeared equally effective, whereas for the strong group, the instructor mode appeared least effective.

Similar trends were seen in the nonmajors’ course (Figure 5B). Namely, the combination mode was most effective for all three groups of students, except for the weak performers, for whom the gains for the instructor-explanation and combination modes were similar. As was true with the majors, the instructor-explanation mode was least effective for the strong students.
DISCUSSION

Summary of Results

Our results show that genetics students, both majors and nonmajors, learn from in-class concept questions whether the mode of administration comprises peer discussion alone, instructor explanation alone, or a combination mode in which peer discussion is followed by instructor explanation (Figure 3). However, the combination mode results in substantially higher learning gains compared with either the peer-discussion or instructor-explanation modes, as measured by the normalized change \(<c>\) in scores between Q1 and Q2 (Figure 4). Analysis of the results for three ability groups of students, designated weak, medium, and strong based on mean Q1 scores, showed that the combination mode was most effective for all three groups in both the majors’ and the nonmajors’ courses (Figure 5).

Strikingly, the strong clicker performers in both classes showed the smallest learning gains when the instructor-explanation mode was used (Figure 5). We hypothesize that discussing questions with peers in either the peer-discussion mode or the combination mode keeps the strong clicker performers engaged with the material. Without this element in the instructor mode, the strong students may pay less attention to the subsequent question, Q2. These results are in agreement with a study that compared overall student learning gains in introductory physics courses taught using traditional lecturing or interactive engagement (Beichner and Saul, 2003). In this study, the stronger students learned more in the interactive courses, possibly because they were cementing their own understanding by helping their peers. Our results support the conclusions of these authors that interactive approaches such as peer discussion benefit the high-achieving students.

We see differences between students in the majors’ and nonmajors’ genetics courses with respect to the weak clicker performers. In the majors’ course, the weak students show...
substantially larger learning gains with the combination mode than with either of the other two modes (Figure 5A). However, for weak students in the nonmajors’ course, the combination mode is only slightly more effective than the instructor-explanation mode (Figure 5B). One likely reason for this difference is that nonmajors were less inclined to regard their peers as learning resources. Several lines of support for this idea come from a previous study in which behaviors and motivation levels of nonmajor genetics students were measured (Knight and Smith, 2010). Observations of these students revealed that they were more likely than majors to ask an instructor rather than peers for help when working on group activities. Nonmajors in this study also studied outside of class significantly less than did majors, consistent with lower levels of motivation. These factors may combine to generate an environment for nonmajors in which the weaker students are less inclined to participate in peer discussion, and thus do not benefit as much as other groups.

Our data from the majors’ genetics course show that the peer-discussion and instructor-explanation modes result in similar learning gains, at least for the weak and medium clicker performers (Figures 4 and 5A). Peer discussion has benefits over listening to an instructor, such as breaking up the monotony of lecture and giving students a chance to practice putting their thoughts into words (Mazur, 1997; Smith et al., 2009). However, in our experience, many students report that peer discussion without any instructor explanation or feedback can be frustrating.

**Why Is the Combination of Peer Discussion Followed by Instructor Explanation So Effective for Student Learning?**

The effectiveness of the combination mode is consistent with previous findings in cognitive psychology, showing that student engagement in a learning activity such as answering questions predisposes them to learn from a subsequent lecture (Schwartz and Bransford, 1998). During peer discussion, students engage with the material by sharing their ideas with others. In short, students are figuring out what they understand and what they have questions about. The instructor explanation immediately following peer discussion in our protocol corresponds to the subsequent lecture in the Schwartz and Bransford (1998) study. Additional studies have shown that feedback to students, which allows them to gauge their current understanding of a topic, can have a positive impact on their future performance. Feedback is especially helpful when it includes a statement of the correct answer and information on why it is correct (reviewed in Roediger et al., 2010). In both the instructor-explanation and combination modes used in this study, students received extensive feedback of this nature from their instructor, as well as explanations of why other answers were incorrect.

In the combination mode, students spent on average about 2–3 min more total time engaging with Q1 than in either of the other two modes (Table 3). Time on task was not strictly controlled in our study because we wanted to compare these modes in a normal classroom setting without imposing time limits on the instructor. However, several considerations argue that time on task alone cannot account for the superior effectiveness of the combination mode. From our experience in the classroom, in agreement with published guides to best practices with clickers (Caldwell, 2007), useful peer discussion following a question is generally limited to 2–3 min, after which most of the students turn to conversations on other matters or personal pursuits, such as email or texting. Therefore, simply adding time to peer discussion would be highly unlikely to increase the effectiveness of this mode significantly. Consistent with this view is a recent physics education study, in which students individually answered a clicker question and then for the next minute discussed the question with their peers, reflected on their answers silently, or were distracted by a cartoon (Lasry et al., 2009). When the students then voted on the question again, the percent change in performance was highest when students engaged in peer discussion, suggesting that the benefit of this activity is not simply to provide additional time considering the question. Based on these arguments, the substantially increased learning gains that result from adding instructor explanation after peer discussion are highly unlikely to be attributed merely to the increased time on task.

Could a longer and more detailed instructor explanation following administration of Q1 have increased the effectiveness of this mode alone to the level of the combination mode? Although we did not do the experiment, three considerations suggest that this possibility is unlikely as well. The first is the size of the effect; the combination mode was on average approximately twice as effective in promoting learning gains as the instructor-explanation mode, in both courses. Adding 2 or 3 min of instructor explanation to an already complete explanation is unlikely to have produced such a doubling. Second, the findings of Schwartz and Bransford (1998) suggest that, after grappling with a question, students are primed to gain more from a subsequent lecture. These results suggest an apparent synergy, which we have also observed, between peer discussion and instructor explanation in the combination mode. Third, student surveys also indicate that they perceive the combination mode to be synergistic. On end-of-course surveys in both courses (n = 122 major respondents, n = 45 nonmajor respondents), students were asked to indicate agreement or disagreement with the following statement: “Having a discussion with my neighbors prepares me to listen to the instructor’s explanation.” Sixty-four percent of the majors and 84% of the nonmajors agreed, and when asked to explain why, several students described how peer discussion helps prepare them to learn. Two such descriptions follow:

- It gets me thinking about the topic before [the instructor’s] lecture, rather than just passively listening to what he has to say – I am already engaged.
- Discussion helps get the ideas and thoughts flowing, which makes what [the instructor] says more concrete.

The class time required for administration and discussion of concept questions, especially using the combination mode, may seem daunting to some instructors. However, several studies have shown the value of modifying course structure so as to place more responsibility on students for learning factual material outside of class, thereby freeing class time for active-learning activities, such as clicker questions and discussion (reviewed in Wood, 2009). In the courses described here, the instructors focused primarily on conceptual understanding in class rather than transmission of detailed factual
knowledge. In addition, students practiced general skills and application of concepts through regular online homework assignments outside of class. These modifications allowed the instructors to require mastery of basic content while still leaving time for in-class active-learning activities.

CONCLUSION AND FUTURE DIRECTIONS

Our research further defines best practices for using in-class concept questions and clickers. From previous work, we know that active engagement of students during peer discussion leads to improved performance (Smith et al., 2009). The results presented here show that, in two different courses, the largest gains in student performance occur when peer discussion is immediately followed by instructor explanation. This combination mode is probably so effective because it combines student engagement through peer discussion with instructor feedback. Qualitative studies on the content of student discussions during peer interaction, currently in progress, should help to better understand the benefits of this mode of clicker use.

Although our results indicate that the combination mode is better for in-class performance than the single modes tested, we still do not know which modes of instruction best promote retention of material. Following the evidence from cognitive psychology studies (Schwartz and Bransford, 1998; Roediger et al., 2010), we would predict that peer discussion immediately followed by instructor explanation should enhance not only short-term learning, but also retention as well. Longitudinal studies are needed to explore this prediction.

ACKNOWLEDGMENTS

We are grateful to the Science Education Initiative at the University of Colorado, Boulder, for full support of M.K.S and partial support of J.K.K during this project. Thanks to Alison Crowe and Mary Pat Wenderoth for assessing the Bloom’s levels of the question pairs. We also thank Carl Wieman and Wendy Adams for intellectual support and comments on the manuscript and Tin Tin Su for help with the experimental design.

REFERENCES


Using Digital Images of the Zebra Finch Song System as a Tool to Teach Organizational Effects of Steroid Hormones: A Free Downloadable Module

William Grisham,* Natalie A. Schottler,* Lisa M. Beck McCauley,† Anh P. Pham,‡ Maureen L. Ruiz,§ Michelle C. Fong,§ and Xinran Cui‡

Zebra finch song behavior is sexually dimorphic: males sing and females do not. The neural system underlying this behavior is sexually dimorphic, and this sex difference is easy to quantify. During development, the zebra finch song system can be altered by steroid hormones, specifically estradiol, which actually masculinizes it. Because of the ease of quantification and experimental manipulation, the zebra finch song system has great potential for use in undergraduate labs. Unfortunately, the underlying costs prohibit use of this system in undergraduate labs. Further, the time required to perform a developmental study renders such undertakings unrealistic within a single academic term. We have overcome these barriers by creating digital tools, including an image library of song nuclei from zebra finch brains. Students using this library replicate and extend a published experiment examining the dose of estradiol required to masculinize the female zebra finch brain. We have used this library for several terms, and students not only obtain significant experimental results but also make gains in understanding content, experimental controls, and inferential statistics (analysis of variance and post hoc tests). We have provided free access to these digital tools at the following website: http://mdcune.psych.ucla.edu/modules/birdsong.

INTRODUCTION

The bird song system has had a long record of yielding cutting-edge findings in neurobiology. Adult neurogenesis (Alvarez-Buylla et al., 1988) and dramatic sex differences in the brain (Nottebohm and Arnold, 1976) were first described in this system. This system is still the focus of ongoing investigations on neurosteroids (Schlinger et al., 2001), learning and memory (Troyer and Doupe, 2000; Nordeen and Nordeen, 2004), and genes and proteins involved in vocal learning (White et al., 2006; Xie et al., 2010). The song system is also a valuable model of both the genetic and hormonal bases of sexual differentiation of nervous systems (Grisham and Arnold, 1995; Agate et al., 2003; Duncan et al., 2009).

The song system consists of a set of interconnected nuclei whose only known function is the learning and production of song (Figure 1). This system can be divided into song acquisition and song production pathways. The song acquisition pathway consists of nuclei in the anterior forebrain pathway: lateral magnocellular nucleus of anterior nidopallium (IMAN), nucleus dorsolateralis anterior thalami, pars medialis (DLM), and Area X, which is its proper name (nomenclature from Reiner et al., 2004; Nixdorf-Bergweiler and Bischof, 2007). Lesions in this pathway disrupt the acquisition (Bottjer et al., 1984; Scharff and Nottebohm, 1991; Johnson and Bottjer,
Figure 1. Semi-schematic of a sagittal view of a songbird brain highlighting the relationships among the song nuclei. RA, IMAN, Area X, and HVC are all telencephalic nuclei. RA, IMAN, and HVC can be considered as roughly homologous to aspects of the mammalian cortex. Area X is a part of the basal ganglia and is also part of the forebrain. DLM and nXII are in the brainstem.

1992; Nottebohm, 2005) and maintenance of song (Williams and Mehta 1999; Roy and Mooney, 2007). The production pathway consists of HVC (here used as a proper name), the robust nucleus of arcopallium (RA), and hypoglossal nucleus (nXII) (nomenclature from Reiner et al., 1999). Lesions of this pathway result in a loss of song (Nottebohm et al., 1976). Electrophysiological evidence suggests that song syllable timing is organized in HVC and the syllables’ actual acoustic properties are organized in RA (McCasland, 1987; Vu et al., 1994; Suthers and Margoliash, 2002).

In most (but not all, cf. Brenowitz and Arnold, 1986) songbird species, the song behavior is sexually dimorphic with males singing more than females. Accordingly, in most species, the telencephalic song nuclei (Area X, HVC, and RA) are dramatically larger in males than in females (Nottebohm and Arnold, 1976; Gurney and Konishi, 1980; Grisham and Arnold, 1995; see Figure 2). This sex difference can be altered in development via the organizational effect of steroid hormones, which determine aspects of the sexual phenotype irrespective of the genetic code carried on the sex chromosomes (cf. Breedlove and Hampson, 2002; Phoenix et al., 1959). Exogenous steroid hormones, particularly estradiol, can masculinize the brain phenotype in developing female zebra finches (Gurney and Konishi, 1980; Grisham and Arnold, 1995, Grisham et al., 2008) and is also known to masculinize mammalian brains in development (cf. Breedlove and Hampson, 2002).

The rich literature on the song system, the robust difference between the sexes in the size of song nuclei, and the dramatic masculinizing effect of hormones on this system’s development (Gurney and Konishi, 1980; Grisham and Arnold, 1995) make the zebra finch song system valuable for demonstrating the effects of hormones on development and sex differences in brain morphology. In particular, undergraduate neuroscience, biological psychology, and animal behavior laboratories could profit from using this system. Nevertheless, studying this system requires significant investments: supporting a bird colony and purchasing good-quality microscopes, microtomes, histological supplies for processing brains, and digital microscope cameras. Such extensive facilities and equipment requirements make studying the bird song system out of reach for many institutions. We have overcome these barriers by creating a digital image library of the bird song system and distributing it at no cost to users at the following website: http://mdcune.psych.ucla.edu.

This library consists of images of female zebra finch brains used in a previously published study (Grisham et al., 2008), along with added images from control males. The females were administered different doses of estradiol at hatching, and their brains were examined and photographed in adulthood. Using this library, students not only repeat an experiment examining the relationship between the dose of estradiol and the degree of masculinization in the song system, but also extend the original experiment because they quantify images from control males that were not part of the original experiment (Grisham et al., 2008). Because of the robust sex differences and large treatment effects, even inexperienced students will obtain data yielding significant differences.

METHODS

Materials
All that is required for this laboratory is a computer. The image library (available for download at the above address) consists of the relevant images of Area X, HVC, and RA in control male and control female zebra finches, as well as females that received 5, 15, or 50 μg estradiol (E2) at hatching. These images were made from tissue prepared for Grisham et al., 2008, so students can refer to this article for procedural details (Pubmed Central ID #PMC2605609). The males came from other studies in which they were treated at hatching with either blank implants (Grisham and Arnold, 1995) or implants of indomethacin, a prostaglandin inhibitor like aspirin, that had no discernable effect on the song system (Borowskit et al., 2010). The male brains were prepared in an identical way to that reported in Grisham et al., 2008. Students quantify the size of song regions using ImageJ, which is a free software package that can run on any platform (National Institutes of Health, 1997).

Procedure
Students are given the following five specific objectives to focus their thinking and to help structure their journal-style laboratory report: 1) Although this module isn’t an independent replication, it is still useful to see whether students’ measurement and reanalysis yield the sex differences in the song system established in the literature (Nottebohm and Arnold, 1976; Grisham and Arnold, 1995). We use this as an opportunity to underscore why this step is necessary to validate their techniques and findings. We point out that there is no need to try to explore the underlying basis of a phenomenon if one cannot repeat it with the same materials. 2) To probe the sensitivity of the female zebra finch song system to E2-induced masculinization. This study was originally motivated by another study that substantially reduced E2 synthesis and found no impact on the masculine development of the song system (Wade et al., 1996). Nonetheless, if masculinization processes were exquisitely sensitive to E2, the residual E2 could have been
sufficient to masculinize the system. 3) To discern whether masculinization is an all-or-none event once a threshold was reached, or whether it is proportional to the dose of E₂. 4) To ascertain whether different parts of the song system show differential sensitivity to E₂, which could provide some clues about song system development. 5) To determine whether our largest dose of E₂ will fully masculinize the song system in females to the same extent as males. We provide students with background readings (Grisham and Arnold, 1995; Breedlove and Hampson, 2002; Nottebohm, 2005; Grisham et al., 2008) as well as background lectures (see http://mdcune.psych.ucla.edu/modules/birdsong) and ask them to formulate predictions based on the readings.

We use this module to teach about experimental procedures to avoid possible confounds. First, each student is assigned a total of five birds, one from each different treatment condition (control male, control female, and 5, 15, and 50 μg E₂ females; see Figure 3). This balancing procedure prevents students' differential quantification style from potentially confounding the results (discussed further in the Complete Users’ Manual: Thorough Guide to Obtaining and Analyzing Data for Students and Instructors, which can be found on the website). We use this design consideration to illustrate that, when extraneous variables are present and cannot be eliminated, an experiment can be designed so that they do not operate differentially across treatments and potentially confound the results. Second, students are kept blind to the sex and treatment of the birds while they are making measurements of the song system, and we explain to students that this controls for bias.

Using NIH ImageJ, students make cross-sectional measurements to the nearest 0.01 mm² of all the images of Area X, HVC, and RA. PDFs on the website, including the Complete Users’ Manual: Thorough Guide to Obtaining and Analyzing Data for Students and Instructors, and FAQs provide details on quantification, including pitfalls. Students can
convert these measurements to volumes because they know the sampling interval between sections (see the Complete Users’ Manual on our website).

Students combine their data, and ultimately several birds’ data are represented in each treatment condition. A master spreadsheet containing all the data along with information about the sex and treatment of each bird is distributed to each student. (Instructors may access a spreadsheet by going to http://mdcune.psych.ucla.edu/faculty and setting up a free faculty account.) Students are guided in analyzing the data using OpenStat (Miller, 2010) or VassarStats (Lowry, 2010), which are free statistical analysis packages available online. Specifically, students need to decide which type of experimental design is used and which analysis would be appropriate (a one-way analysis of variance [ANOVA] with five levels). We use this data analysis to introduce ANOVA and post hoc statistical tests and why they are appropriate to analyze our data. As a post hoc test, we typically use Fisher’s Least Significant Difference test, which OpenStat does automatically at $p < 0.05$ if a significant overall F ratio is obtained. (Details on our lesson on ANOVA and factors that can influence statistical power are available as part of the PowerPoint presentation, Lecture 3, available on our website: http://mdcune.psych.ucla.edu/modules/birdsong.)

Students characteristically obtain remarkably good data and invariably find significant differences. Data from one section of our students are displayed in Figure 4. Students are assigned to write a report in the form of a journal article and are guided through this process via discussion of the five objectives outlined above (also see PowerPoint slides in Lectures 3 and 4 on our website: http://mdcune.psych.ucla.edu/modules/birdsong).

Initially, we taught this module by interpreting the data for students. Both formal and informal feedback suggested

Figure 3. RA in a single section from each of our groups. Differences are quite evident and easy to quantify. (Again, students quantify multiple sections for each nucleus of each bird.)
that this “spoon feeding” wasn’t desirable. Rather, students indicated that they would appreciate the intellectual challenge of interpreting the data themselves, especially since they already had the five-question framework. Teaching the module without interpreting the data for students predictably led to significantly lower grades \( t(141) = 2.088, p < 0.05 \), with higher variance in scores, but the absolute difference in mean scores was only 3.13%. The intellectual exercise seemed well justified.

Implementing the Bird Song System Module

This module works well with a class of 24 students, but it is adaptable for smaller or larger enrollments. We devote 3 wk of 3-h labs and three 1-h lectures to this module. Students are easily able to complete their measurements and data analyses in this time. When we assign five birds (one from each condition) to each student, it takes about 3–5 h, maximum, for students to finish quantifying Area X, HVC, and RA in all of their birds. (This time includes orienting students to the ImageJ quantification tool and collecting all of their data in a master spreadsheet.) Data analyses are accomplished fairly quickly, but it takes much longer to explain the rationale and meaning of the analyses. Analyzing data, explaining their purpose, interpreting the results, and demonstrating how to present data in a meaningful manner takes a couple hours, which we usually distribute between lab and lecture. (See PowerPoint slides from Lectures 3 and 4.) Finally, as we are going through this project, we train students to write a journal-style lab report, and this effort usually takes 1–2 h, which can either be in lecture or lab (again, see PowerPoint slides from Lectures 3 and 4).

Our students are either Psychobiology or Neuroscience juniors and seniors. Most students have had a course in statistics, and some have had exposure to neuroanatomy. Some have had some exposure to discussions on hormones and behavior, but most have not. Most have had a course in which they learn how to draft a journal article report. In teaching this module, we have found it useful to review relevant neuroendocrinology, neuroanatomy, statistics, and discussions on formulating a write-up. Notably, we discovered that, even when students have a background in statistics and writing journal-style reports, they still need a refresher tutorial. To assist faculty with these tutorials, we have provided PowerPoint presentations focusing on neuroendocrine background (Lectures 1 and 2), statistics (Lecture 3), and writing (Lectures 3 and 4). Thus, prerequisite courses are probably not necessary if the instructor provides relevant background on these topics. We use the exercise to reinforce statistical and writing skills that the students acquire in other courses, but this exercise also could be used to introduce these skills.

From the instructor’s point of view, this teaching module is easy to implement. In teaching this unit, we identified and remedied many obstacles to make it a better learning experience for both students and instructors. Many common problems are addressed in the Complete Users’ Manual and the FAQ list on our website.

Possible Shorter Version of This Module

Instructors could construct a module that would take less time but would still be informative by using fewer groups. For example, instructors could use the control females and control males along with the females treated with 50 μg in order to demonstrate the sex difference and show that masculinization can be achieved by administering E2. Alternatively, instructors could use just the control males and control females to demonstrate the sex differences.

Assessing the Effectiveness of This Module

All assessment measures had IRB approval (UCLA IRB Exemption #07–211). Efficacy of this module was assessed in two consecutive terms in lab classes composed of UCLA Neuroscience and Psychobiology majors. Both groups have very similar demographics and career ambitions. We administered a pretest and a posttest including a scale that assessed understanding of the content, statistics, and experimental design with a subscale of items that assessed “inappropriate thinking biases” based on those presented by Stanovich (2009). Students also completed a Materials Evaluation questionnaire that included Likert-scale items and an open-ended item to assess their experience with the module and their opinions about the materials provided. The open-ended question asked students to describe the purpose of the Bird Song System module from a learning standpoint. The pre/posttest and Student Materials Evaluation can be found in the Supplemental Material.

Because repeated testing itself can sometimes raise scores (Campbell and Stanley, 1963; Trochim, 1986, 2006), we administered the posttest alone in a second sample of students in a subsequent term. All students received the Materials Evaluation questionnaire, and their responses on this questionnaire were combined across both samples.

RESULTS

We compared the posttest scores to the pretest scores separately on our two scales: the module-specific and
“inappropriate thinking biases” scales. One item on the module-specific scale was discarded because of poor psychometric properties (item #15), which seemed to be poorly worded; the pattern of results was not appreciably altered by discarding this item. The posttest means on the module-specific scale, addressing content, statistics, and experimental design, were significantly higher than the pretest of the first sample of students, $t(111) = 17.78, p < 0.001$, and $t(148) = 10.74, p < 0.001$, respectively (Figure 5). The posttest means on the module-specific scale were not significantly different from each other, $t(148) = 0.118, p = 0.906$, suggesting that the gains between the posttest and pretest reflected genuine increases in learning and reasoning skills and not just an artifact of repeated testing (Figure 5). Also, grades on the module were not significantly correlated with pretest scores, $r(110) = 0.06, p > 0.50$, suggesting that better performance was not correlated with better preparation before the module. Posttest scores did correlate with grades on the module in both terms, $r(36) = 0.37, p < 0.05$ and $r(110) = 0.33, p < 0.05$, providing convergent validity for our module-specific pre/posttest. Not surprisingly, scores on the four-item “inappropriate thinking biases” subscale (items 19–22 on the posttest) were uninfluenced by discarding this item. The posttest means on the module-specific scale were not significantly different from each other, $t(148) = 0.118, p = 0.906$, suggesting that the gains between the posttest and pretest reflected genuine increases in learning and reasoning skills and not just an artifact of repeated testing (Figure 5). Also, grades on the module were not significantly correlated with pretest scores, $r(110) = 0.06, p > 0.50$, suggesting that better performance was not correlated with better preparation before the module. Posttest scores did correlate with grades on the module in both terms, $r(36) = 0.37, p < 0.05$ and $r(110) = 0.33, p < 0.05$, providing convergent validity for our module-specific pre/posttest. Not surprisingly, scores on the four-item “inappropriate thinking biases” subscale (items 19–22 on the pre/posttest; see Supplemental Material) were uninfluenced by the module; students performed roughly at chance both before and after the module, and comparisons on this subscale were nonsignificant, all $p > 0.50$ (data not shown).

Responses to the Likert-scale items reflected a general enthusiasm for the module and showed that students believed that they were gaining in knowledge and skills. Students generally agreed that they had learned a lot about sexual differentiation of the brain and data analysis (Figure 6). Students also liked getting significant results and repeating and extending a published experiment. Most students also reported that using the digitized images was easy, and they felt that they learned as much from the digitized images as they would have from tissue on slides. (Notably, they had a side project that dealt with using actual tissue on slides, so they had an apt comparison.)

Students responded to the following open-ended prompt: “Please describe the purpose of using the Bird Song System module from a learning standpoint.” Responses were coded as content-related (“sexual differentiation of the zebra finch,” “role that estradiol plays on the female zebra finch system”), comprehension (“to learn,” “to see,” “give students a better grasp”), experimental method (“how a scientific study is produced from formulation of hypothesis to the data collection,” “learn how to conduct a thorough research experiment”), data analysis (“students learned about ANOVA tests, how to justify and draw conclusions from data,” “applying the statistics that I’ve learned….to ‘real’ data, and build critical thinking skills by analyzing the data”), writing skills (“wrapping up all the ideas in a formal research article format,” “formatting of a proper paper”), hands-on experience, extended relevance (“applicable to that of the human brain,” “generalize to human beings,” “connected to actual behaviors,” “further studies in mammals”), positive affect (“excellent,” “module did a great job,” “Overall, I liked this module!” “It was very fun and enjoyable….”), and negative affect, which included all negative comments, whether regarding clarity or effectiveness of the materials. Table 1 presents the frequency of each response category.

**Table 1.** Categorization of responses to the open-ended question “Please describe the purpose of using the Bird Song System module from a learning standpoint.”

<table>
<thead>
<tr>
<th>Category</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comprehension</td>
<td>100</td>
</tr>
<tr>
<td>Content</td>
<td>89</td>
</tr>
<tr>
<td>Experimental method</td>
<td>57</td>
</tr>
<tr>
<td>Data analysis</td>
<td>40</td>
</tr>
<tr>
<td>Writing skills</td>
<td>36</td>
</tr>
<tr>
<td>Hands-on experience</td>
<td>24</td>
</tr>
<tr>
<td>Extended relevance</td>
<td>12</td>
</tr>
<tr>
<td>Positive effect</td>
<td>25</td>
</tr>
<tr>
<td>Negative effect</td>
<td>4</td>
</tr>
</tbody>
</table>

A given student’s responses can be coded in more than one category ($n = 136$ students providing data on this item).

**DISCUSSION**

This lab experience illustrates the organizational effect of gonadal steroids on brain development (Breedlove and Hampson, 2002), which occurs in early life during a critical/sensitive period and persists for the lifetime of the individual (Phoenix et al., 1959; Breedlove and Hampson, 2002). $E_2$ was administered when the birds hatched, yet the effects of this early exposure were still manifest in adulthood. Although the masculinization of the genetic females’ song system is only partial, it is still impressive (see Figure 5).

In the course of this module, students learn how to do a careful image analysis, how to work with a large data...
Figure 6. Responses to selected questions on our Likert scale; percent of total students in both samples combined \((n = 153)\) as a function of scale point (full set of responses and questions in the Supplemental Material). Question 11: I learned a lot about sexual differentiation of the brain through the Bird Song System module. Question 12: I learned a lot about analyzing data through the Bird Song System module. Question 17: I felt like I learned as much using the digitized images as I would have using tissue on slides. Question 20: I appreciated being able to jump right in and collect data.

The open-ended item, which tapped students’ impressions of the purpose of the module (Table 1), indicated that the experience enhanced students’ understanding of experimental methodology and design, data analysis, and writing skills as well as content. Although affect wasn’t explicitly queried in the open-ended item, positive comments outweighed negative by a ratio of >5:1 (Table 1).

The Likert-scale questionnaire indicated that students agreed that they had learned a lot about content and data analysis (Figure 6). Further, students appreciated being able to obtain significant differences in their data. One advantage of this lab is that effect sizes and sex differences are so large that even inexperienced students can obtain data that will yield significant differences. In teaching this module across several terms, students have never failed to obtain data that yielded significant differences: students’ data usually come out as they did in Figure 4. Thus, students obtain the sex differences described in the literature and can relate their results to published studies. Students find that, although the female song system is not exquisitely sensitive to estradiol (5 μg never masculinizes), usually 15 μg is sufficient to masculinize the song system. Generally, students find that masculinization maximizes with a dose of 15 μg and doesn’t increase with a dose of 50 μg. Further, 50 μg, which is a massive dose, does not completely sex-reverse the song system. Usually, different parts of the song system don’t show differential sensitivity to estradiol, but because this is an inquiry-based lab, students’ data don’t always come out the same way (see the Complete
Users’ Manual on the website for a discussion of anomalous results).

Students reported that using the digitized images was easy, and they felt that they learned as much from the digitized images as they would have from tissue on slides. Notably, we had students perform a side project in which they sectioned a zebra finch brain and then tried to determine the sex of their “mystery” bird by examining the song system. Although students seem to enjoy slicing and staining the actual tissue, we have found that using the image library is far superior to using the actual tissue on slides. When using the actual tissue, students and even graduate teaching assistants were observed wandering around the lower brainstem trying to find the song nuclei that are in the forebrain. Clearly, the image library focuses neophytes on the task better than using the actual tissue.

Our data suggest that this learning experience was an enjoyable and valuable one from the students’ perspective. We have created the materials so that instructors can easily adopt the module. Further, no specialized equipment is required to provide this experience to students, only computers connected to the Internet. Due to their digital format, these materials can be incorporated into a standard classroom, a laboratory setting, or even distance learning.

All materials required for teaching this module, including the image library, the Complete Users’ Manual, PDF handouts, PowerPoint slides with and without voice-over commentary, links to relevant websites, as well as lectures on the material, are available for free at http://mdcune.psych.ucla.edu/modules/birdsong. Grading keys/rubrics, keys to the treatment group for each bird, and instructor-specific materials also can be obtained for free by accessing http://mdcune.psych.ucla.edu/faculty/create-a-faculty-account and setting up a faculty account.

ACKNOWLEDGMENTS

We thank A.-M. Schaaf for her help in preparing the manuscript and Jamie Crisostomo for help in formatting the manuscript. This study was supported by NSF CCLI DUE-0717306 (to W.G.) and a grant from the UCLA office of Instructional Development (to W.G.).

REFERENCES


Article

Redesigning a Large-Enrollment Introductory Biology Course

Catherine Ueckert,* Alison Adams,* and Judith Lock†

*Department of Biological Sciences, Northern Arizona University, Flagstaff, AZ 86011; †School of Biological Sciences, University of Southampton, Southampton S017 1BJ, United Kingdom

Submitted October 7, 2010; Revised January 13, 2011; Accepted January 20, 2011

Using an action research model, biology faculty examined, implemented, and evaluated learner-centered instructional strategies to reach the goal of increasing the level of student achievement in the introductory biology course BIO 181: Unity of Life I, which was characterized by both high enrollments and a high DFW rate. Outcomes included the creation and implementation of an assessment tool for biology content knowledge and attitudes, development and implementation of a common syllabus, modification of the course to include learner-centered instructional strategies, and the collection and analysis of data to evaluate the success of the modifications. The redesigned course resulted in greater student success, as measured by grades (reduced %DFW and increased %AB) as well as by achievement in the course assessment tool. In addition, the redesigned course led to increased student satisfaction and greater consistency among different sections. These findings have important implications for both students and institutions, as the significantly lower DFW rate means that fewer students have to retake the course.

INTRODUCTION

Research in biology has changed dramatically over the past three decades due to the development of new technologies, such as recombinant DNA and polymerase chain reaction, which have led to new questions and approaches and the emergence of new fields of study such as proteomics and bioinformatics. However, according to the National Academies (2002, p. 1), “In contrast, undergraduate biology education is still geared to the biology of the past…. [M]ost courses, especially those for first-year students, are still primarily lecture-based, and do not convey the exciting reality of biology.”

Since 1990, reform in American science education has been led by documents such as Science for All Americans: Project 2061, Benchmarks for Science Literacy, and the National Science Education Standards (American Association for the Advancement of Science [AAAS], 1989, 1993; National Research Council [NRC], 1996). These documents call for a major shift in not only what is taught but also how it is taught in kindergarten through 12th grade. Federal dollars have been appropriated through several programs, including the U.S. Department of Education (Eisenhower Secondary Mathematics and Science Education Program and the Bush No Child Left Behind Act), to provide professional development opportunities for states to instigate the suggested changes and monitor student performance. Colleges and universities have been involved as partner institutions that developed and offered various workshops and courses for in-service and preservice science teachers.

Efforts to research biology, chemistry, and science education in general have provided valuable insights (Herron and Nurrenbern, 1999; Klymkowsky et al., 2003; Stokstad, 2003; Yehudit et al., 2003; Handelsman et al., 2003; Handelsman et al., 2004; Mazur, 2009) on the effectiveness of various pedagogical interventions. However, the physics education research community has led the effort to implement and research the effectiveness of interactive methodologies (McDermott, 1991, 1996; Hestenes et al., 1992; Hake, 1998; Redish, 1999; Meltzer and Manivannan, 2002; Rebello and Zollman, 2004; Heron and Meltzer, 2005; Wieman and Perkins, 2005; Mason and Singh, 2010). The research from the physics community stresses the importance...
for students to be engaged in problem solving, modeling, discussions, and other modes of involving students in their own learning. Many of the pedagogies included in the revision described in this article are based on these efforts and the research on how people learn (Bransford et al., 2000).

Engaging students in their own learning whether by interaction with the content, peers, or the course instructor increases learning (Slater et al., 2006). Classrooms that utilize active-learning strategies are noisy. Students are debating ideas, asking questions and comparing answers to what is known, using evidence to develop explanations, considering alternatives, communicating their ideas, and, upon reflection, often changing their ideas. Sometimes students work in groups to collect data in real-life contexts or apply knowledge gained in the classroom to societal problems. These increased interactions with content and peers make thinking explicit. Active-learning strategies simulate the processes involved in scientific inquiry and appeal to many students because they accommodate different interest and learning preferences (Ueckert and Gess-Newsome, 2007).

Science is an active process. It involves asking questions, making observations, and collecting and analyzing data, which are then used to justify explanations. Active learning is similar to the process of science as it requires students to engage with the content and with others in order to answer questions and solve problems, unveil prior ideas, and make connections between ideas. As recognized in the National Science Education Standards, “Student understanding is actively constructed through individual and social processes” (NRC, 1996, p. 29). Learning involves connecting what we already know with new knowledge. This is an active process and requires students to take responsibility for their own learning. Learners must recognize what they know and what they still have questions about (metacognition). Understanding a concept is often manifested when students are asked a question in which they must transfer prior knowledge and understandings to a new context. This transfer occurs only if learners are aware of the principles underlying their thinking (Bransford et al., 2000; NRC, 2000). The research on learning implies that students need multiple opportunities to think deeply and purposefully about the content and to gain feedback on their learning (Ueckert and Gess-Newsome, 2008).

We encountered several challenges in teaching BIO 181: Unity of Life I that are documented in the literature on large lecture-based classes, including 1) ill-prepared students, 2) poor attendance, 3) diverse learning abilities, 4) passivity, and 5) lack of student feedback (Slater et al., 2006; Freeman et al., 2007; Walker et al., 2008). The academic problem addressed in this study was the low level of student success in the course as reflected by one of the highest rates of students who received a D, an F, or a W at our university. For the previous seven semesters, the DFW rate for students taking BIO 181 varied from 28 to 51% (depending on the section). This DFW rate not only impacted the potential of students to enter the major of their choice but also contributed to overenrollment in BIO 181 and the problem of limited lab space. Our goal was to use learner-centered education strategies to increase the success of students in BIO 181 and to increase consistency among sections.

Changing how things have been done traditionally is often met with resistance from both faculty and students, and the impact is often limited to a single professor making the changes (Brainard, 2007). To make a more widespread impact on BIO 181, a systematic approach was used. Principles of instructional design suggested that the process should begin with an analysis of the student population and identification of learning outcomes. Following the implementation of changes, data should then be collected and analyzed to guide continual curricular improvements. These guidelines were used to redesign BIO 181.

The reform project discussed in this article was developed using these central elements of science education reform (AAAS, 1989, 1993; NRC, 1996): collaboration on course goals and content, establishment of a collegial climate in which ideas are openly discussed and refined, and deliberate design of an assessment plan that aligned with course goals and would guide future course development. The project was partially funded by an Arizona Board of Regents Learner-Centered Education grant.

The course redesign had five components: 1) development of a common syllabus with common learning outcomes and pre/posttests to increase consistency among instructors; 2) introduction of online tools, including tutorials, simulations, and PowerPoints, to increase student engagement and learning; 3) use of online quizzes to help students keep abreast of material and to provide students with evidence of successful learning; 4) implementation of classroom response systems, together with small-group work, to increase student participation in the learning process during class time; and 5) alignment of laboratory investigations and course content to help reinforce lecture material. The course was also enhanced by a concerted effort to include real-life applications of concepts wherever possible. A timeline of our course redesign efforts is found in Table 1.

NOTE: Both the Eisenhower Secondary Mathematics and Science Education Program and the Bush No Child Left Behind Act are readily known and probably do not need to be referenced; however, the following link describes the Eisenhower Professional Development Program: www2.ed.gov/programs/eisenhower/index.html. This program was replaced by the Improving Teaching Quality State Grants program that was established under Title II-A of the No Child Left Behind (NCLB) Act of 2001. “The primary goal of each grant is to improve teachers’ pedagogical and academic content knowledge through a program of rigorous professional development.” www2.ed.gov/programs/teacherqual/performance.html.

METHODS
The present study began in 2006 when funding was received to incorporate learner-centered instructional strategies to reach the goal of increasing the level of student achievement in a gateway introductory biology course (BIO 181) with high enrollment and a high rate of failure as measured by the DFW rate. BIO 181 is the first course required for all biology majors and 22 other majors at our public, high research institution with an enrollment of 23,600 students statewide and 16,000 on campus. This introductory course for majors serves 900 students in the fall semester and 500 students each spring. There are 34 lab sections offered in the fall and about 19 lab sections offered in the spring. There is a high demand for the course; however, the number of students taking BIO 181 is limited due to the availability of lab and lecture space. There
are three to five lecture sections of BIO 181, each containing 120–240 students and taught each semester by a variety of instructors (including one part-time lecturer, a visiting assistant professor, adjunct faculty, and tenured/tenure-track faculty).

BIO 181 is the first course of a two-semester introductory biology sequence. BIO 181 focuses on cell and molecular biology, and BIO 182 (the second course) focuses on organismal biology, ecology, and evolution. The topics for BIO 181 are shown later in Table 3. Success in learning these concepts is critical for students when taking more specialized courses of microbiology, genetics, anatomy and physiology, evolution, and so forth.

Many students in BIO 181 are freshmen majoring in biology (19%), although 81% are either pursuing other majors in the scientific field (64%), undeclared (15%), or simply taking the course because they want to (1%). A typical demographic makeup is 59% female and 41% male, with 73% white, 11% Hispanic, 8% American Indian, and 8% other.

In the summer of 2006, six faculty members met to determine the essential learning expectations for BIO 181. A small stipend was paid. The process began by discussing the following questions: What did the instructors like about BIO 181? What did they dislike about BIO 181? Which of the “dislikes” could change? What were the goals of the course in terms of the instructors’ expectations for students?

This discussion led to an organizational concept of the content for BIO 181. Three key organizing concepts to the study of BIO 181 emerged: self-organization, self-maintenance, and self-replication of cells (Figure 1). Crosscutting these three learning goals were the overarching goals of “science as a way of knowing” and the interrelated concepts of relevance, application, and problem solving. It was upon this model that the table of specifications for BIO 181 (Table 2) was developed.

### Table of Specifications: Obtaining Consensus on Amount of Time Spent on Various Topics

One of the issues that emerged in discussion with the various instructors was that there was a huge difference in focus by individual instructors. For example, some topics that were considered essential to one instructor were not covered at all by another instructor, and the amount of time that was devoted to various topics varied widely. We therefore developed a table of specifications (Table 2) that reflected common beliefs as to what should be covered and what proportion of the course should be devoted to the various topics.

A table of specifications is a two-dimensional table that relates key concepts, the levels of skills desired, and the amount of time or emphasis for each concept. A table of specifications should be developed before the assessment is written and before teaching begins (Kubiszyn and Borich, 2003). A table of specifications requires considerable time to construct due to the need to reach consensus as to which concepts will be taught in the course and the desired level of skill. In our opinion the benefits greatly outweigh the time invested because it ensures that the instructor teaches the required content at a designated level of comprehension (Bloom et al., 1971). The table of specifications (Table 2) ensured that a fair and representative sample of questions appeared on the pre/posttest measure and that thinking skills beyond concept comprehension—namely, application, analysis,
synthesis, and evaluation—were included on the assessment (Notar et al., 2004).

**Generation of a Course Assessment Tool**

The table of specifications (Table 2) was used to guide the construction of a course assessment tool (pre/posttest). Instructors began the process by choosing questions from the question bank from the textbook publisher that were both content appropriate and representative of the required skill level, specifically, critical thinking, concept attainment, relevance/connections, and application (see Table 2). Some of the distracters were modified to include common misunderstandings observed during the cumulative experiences of the instructors. Such distracters allow assessment of higher-order thinking skills (Treagust, 1988; Rebello and Zollman, 2004). The distribution of questions reflected the fraction of time spent on each topic. A multiple-choice format was used to administer the test to hundreds of students, provide rapid feedback, and probe for various levels of understanding of the fundamental concepts taught in this course. This test initially consisted of 100 multiple-choice questions, each one categorized into four areas (see Table 2). The principal investigator of the grant created a bank of questions that was reviewed by the instructors and the department chair.

Students were given the pretest during the first week of the semester and a posttest that could be taken during the last week of the semester. The tests were administered online as a low-stakes assignment. Course points were awarded to students who completed the tests. No rewards were given for the correct answer.

The pre- and posttests were administered through Blackboard VISTA, beginning in Fall 2006. Examining the pretest results from Fall 2006 and Spring 2007 showed that, prior to taking the course, there were no significant differences among the sections of BIO 181. The similar pretest mean scores indicate that the various sections of BIO 181 contained students with similar background knowledge of biology. Comparison of pretest and posttest scores indicated that student learning occurred in all sections. However, data analysis of this initial trial revealed significant differences between sections taught by different instructors (unpublished data). There are various possible reasons for this difference: 1) Emphasis on the posttest exam may have varied from one instructor to another. For example, the effort made by students varied widely, as indicated by the amount of time on task (sometimes less than 10 min). 2) Difficulties with Blackboard VISTA were expressed by some students. 3) There may have been significant variation in course content between sections. This finding was used to foster further discussions to identify strategies that could be used to minimize the large discrepancies in student learning that was occurring between sections.

**Developing a Common Syllabus**

The table of specifications was used to draft a common syllabus. BIO 181 instructors from the university and local community colleges met nearly every week during the Fall 2007 semester to 1) develop common learning outcomes, 2) determine the essential content for BIO 181, 3) develop a common syllabus with matched reading assignments from a common text, and 4) build collaboration between local community colleges and the university.

In this process the following questions were discussed: What content could be eliminated? What content was absolutely necessary? What was the prerequisite knowledge that students needed? What knowledge could students obtain from text/tutorial, and what needed to be taught during class time? What were the applications of the information to other majors? The group agreed that the depth and breadth of the content taught should be consistent across all sections. A way to establish this consistency was to create a common syllabus and a set of course outline notes that everyone would use. These notes would consist of the minimum content that must be taught for a topic. Additional content could be added based on the preference of the instructor. Beyond the core course notes, it was agreed that the group would create a BIO 181 Toolbox, which would be an online resource for course instructors. This toolbox would include small-group activities, teaching strategies for specific concepts, PowerPoint presentations, relevant examples, websites, animations, videos, and so forth. The basic content of the course would be the same for everyone, but the details and method of course delivery could vary.

During the weekly meeting time, the instructors divided into working groups to develop a common course outline. One member of each group then rotated to another group to ask questions and provide input. This process continued until a common set and order of topics emerged (Table 3). Using this outline, a common syllabus, lecture outlines, and PowerPoint presentations were developed. These were posted on the instructors’ website for all to use as a basic resource along with animations and videos that instructors have found useful. Strategies for effectively teaching BIO 181 were also shared. These included active-learning strategies such as classroom response systems (clickers), think-pair-share, small-group work, Web-based interactive tutorials, and shorter lectures with more interaction with the students.

The content was organized so that the course began with a discussion of what science is and how science is done. The next topic was cells—something the students were already familiar with and at the heart of what the course was about. This led into a discussion of membranes and their lipid composition. The biochemistry of lipids was therefore taught in context. The structure of lipids was connected to their

<table>
<thead>
<tr>
<th>Concept and time dedicated to each intersection</th>
<th>Critical thinking 32.5%</th>
<th>Concept attainment 32.5%</th>
<th>Relevance/connections 22.5%</th>
<th>Application of concept 12.5%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Science as a way of knowing 10%</td>
<td>3.25%</td>
<td>3.25%</td>
<td>2.25%</td>
<td>1.25%</td>
</tr>
<tr>
<td>Self-organization 30%</td>
<td>9.75%</td>
<td>9.75%</td>
<td>6.75%</td>
<td>3.75%</td>
</tr>
<tr>
<td>Self-maintenance 30%</td>
<td>9.75%</td>
<td>9.75%</td>
<td>6.75%</td>
<td>3.75%</td>
</tr>
<tr>
<td>Self-replication 30%</td>
<td>9.75%</td>
<td>9.75%</td>
<td>6.75%</td>
<td>3.75%</td>
</tr>
</tbody>
</table>

Table 2. Table of specifications

Redesigning an Introductory Biology Course
function and how this function was dependent on their chemical composition and interactions. The next topic was structure and replication of DNA, which led easily into the concept of the cell cycle and topics such as cell-cycle control and mitosis. Following this, students explored some of the issues of gene expression and a discussion of proteins as machines in the cells. As with lipids, the biochemistry of proteins was taught in context so that students could connect the structure of the various amino acids to the function of protein machines. At this point, they could also understand how variations arise through mutations and how these mutations are expressed as changed amino acids with altered properties. This then led into ideas of how mutations are inherited with discussions of meiosis and genetic crosses. Finally, students were exposed to the idea that cells need energy for everything they do. The structure and metabolism of carbohydrates was introduced in this context. Throughout the course an effort was made to connect new ideas to previous learning—a strategy that is known to increase student understanding and retention (Bransford et al., 2000).

It is important to note that despite using a common syllabus, instructors still have flexibility in how they teach the material. For instance, one instructor uses MasteringBiology (Pearson Publishing; http://masteringbiology.com/site/index.html; a widely used platform for tutorials, homework, and assessment) and extensive animations. Another instructor incorporates simulations that involve students and various props. Some instructors use PowerPoints in class, and others use them simply as an online summary of lecture material that is taught on the board in class.

**Coordinating the Lectures with the Labs**

The lab for BIO 181 has always been taken separately from the lecture; however, students must be enrolled in both simultaneously. Over time, the lecture and lab in BIO 181 had become disjointed, with many lab topics unrelated to lecture topics. The common syllabus enabled the lab and lecture material to be coordinated as all instructors were now teaching the material in the same order (Table 3).

The lab that accompanied BIO 181 was therefore redesigned to align with the lecture. Revisions included new topics with an emphasis on applying concepts in novel situations that require students to make interdisciplinary connections; alignment of topics with lecture; a new lab format that uses a quantitative approach, that is, a focus on equations, data sets, and graphical analysis; and an emphasis on the nature of science (Ueckert, unpublished results).

**Incorporating Concepts of Chemistry in the Syllabus**

One of the issues that emerged from our discussions was how the chemistry component of BIO 181 should be taught. Biology courses, especially introductory courses, are laden with chemistry—so much so that students often remark during the first weeks of the semester that it seems as if they are enrolled in a chemistry rather than a biology course. Facts and terminology are often taught in rapid succession in the belief that students cannot learn the biological concepts without first understanding such basic topics as atomic structure, redox reactions, polarity, bonding, pH, chemical reactions, properties of water, and energy. To address this issue we decided to 1) reduce the in-class time spent teaching basic chemistry by developing an online component of the course specifically devoted to chemistry and 2) teach more advanced concepts as they came up (i.e., on a “need-to-know” basis). Development of an online test and tutorial occurred in collaboration with the chemistry department; these were intended to teach basic chemistry concepts and to ensure students had a working knowledge of these concepts. Students were required to pass the chemistry test by the end of the first week in order to continue with the class. According to the National Science Education Standards, by the time students graduate from high school they should know about atomic structure, chemical bonding, and chemical reactions (NRC, 1996). The chemistry test thus served as a review of previously learned concepts.

More advanced topics in chemistry were taught in context, that is, when those chemistry topics were most crucial for the understanding of a specific biological concept. For instance, pH was introduced as enzymes were discussed; types of bonds and their significance in organisms were explored when students were learning about lipids and structure and function, and again when students were learning about protein structure and function; and redox reactions were taught during discussions of photosynthesis and respiration.

**Increasing Relevance of Course Material**

Another goal of the course redesign was to make the curriculum more effective by increasing relevance and including discussion of applications to real-world issues and societal challenges. The students were surveyed to determine the composition of the students enrolled in BIO 181. The student population was quite diverse. Only ~20% of our students were biology majors. Far more (~80%) were majors in health sciences, exercise science/athletic training, or forestry/environmental engineering or were undeclared.

Relevance was added to the course by making a concerted effort to include pertinent examples and case studies in lectures to add interest and appeal to the various audiences and designing in-class group tasks that required students to apply concepts to real-life applications. For example, Nabhan and Tewsbury’s study (Freeman, 2005) on directed deterrence appealed to environmentalists/forestry majors and the molecular basis for genetic diseases appealed to students in health-related majors. We strove to make connections between topics and current events that students might be interested in such as the relationship of the cell cycle to cancer, the connection between mutations and phenylketonuria, and
the role of DNA analysis in forensics, paternity tests, and so forth. Placing concepts in meaningful contexts is well known to increase students’ motivation to learn (Bransford et al., 2000; Hulleman and Harackiewicz, 2009).

**Helping Students Understand Science as a Process**

A concerted effort was made to teach biology not as a set of facts but rather as a process of discovery that has evolved over time, especially in terms of increased sophistication of tools and technology. We wanted students to realize that the findings of others become the foundation for future research, science has a cultural and societal component, and science may be controversial. For instance, at the beginning of the semester, students are given written descriptions of five major scientists and their experiments that played a role in the 200-yr debate on spontaneous generation. The scientists were from different countries and professions and held contradictory beliefs about spontaneous generations. Students form informal small groups to discuss their assigned scientists and experiments. A person from each group reports the main points to the class. The information is used to create a summary table of the contributions and lessons learned from each scientist. The goal is for students to discover that scientists come from various cultures and backgrounds, scientific experimentation has evolved to its present state, science is tentative, and people are reluctant to accept ideas different from their own. These understandings provide the foundation for studying the development of other theories and hypotheses during the semester.

**Increasing Student Engagement through the Use of Class Response Systems**

Most of the instructors used classroom response systems, or “clickers,” to involve students in their own learning. Clickers can be used in a number of ways, for example, to see what the students already know, to determine whether they have read the assigned reading, to check for student understanding, and to challenge students to connect ideas. Answer choices of common misconceptions and “trouble spots” observed in previous students are included. The students’ responses help to reveal gaps in their understanding. This information is used to alter lectures and teach what is needed to help students understand. Some multifaceted questions that require students to use what they know to build new understandings are also included. In this situation, the students answer the question (the instructor does not reveal the correct answer), then discuss their reasoning for choosing their answer in informal groups of two or three, and then reanswer the question. A discussion follows with students sharing their reasoning for their answer choices. These think-pair-share exercises result in beautiful teachable moments, and the students are eager to learn.

Clicker questions give students rapid feedback on their level of understanding and a preview of the types of questions that will be asked on quizzes and exams. Both of these are needed for students to not only learn but also feel good about their accomplishments and motivate them to learn more. Classroom response systems were implemented to check student understanding and foster small-group conversations about key concepts in large lecture settings. Students are asked to predict an answer to various problems or situations and try to convince their neighbor that their answer is correct. The ensuing discussions not only are of high energy but also involve connection of various concepts and the opportunity to clear up any misunderstandings.

Use of the clickers also allows students’ attendance to be tracked and compared with their performance in the course. We found a direct correlation between class attendance and academic success (similar to other findings, e.g., Van Blerkom, 1992; Gunn, 1993; Moore et al., 2003; Newman-Ford et al., 2008). Students who attended class earned higher course grades than those who did not. The scatterplot (Figure 2) suggests a strong linear trend between attendance score and course score in BIO 181. The correlation coefficient between attendance score and course score is .86 (based on 132 students).

**Helping Students Engage with the Material**

Most sections of BIO 181 have 120–240 students. To help students with the material, the following were used:

- Web links that teach students to study for biology or assist in the enhancement/review of background knowledge and skills. The Web tutorials help with the transition from high school to college.
- Scaffolded reading assignments that include learning objectives, smaller sections of text to read and answer questions about, and in-class questions on the reading assignment to increase student accountability.
- Small-group work that engages students in the preparation and presentations of content applications.
- Weekly Web-based quizzes. In introductory classes, students have often not learned how to study on a regular basis, and they often get behind on the material. To address this issue, weekly Web-based quizzes were implemented.
Figure 3. Average grade distributions (%) in all sections of BIO 181 over a 6-yr interval. The redesign process occurred during academic year (AY) 2007 and was implemented across all sections in AY 2008. Average grade distributions for each semester were obtained by adding numbers of students in all sections obtaining an AB, C, or DFW that semester and dividing by the total number of BIO 181 students that semester. Blue, %AB; red, %C; green, %DFW.

These provided students with frequent feedback on their learning and helped them keep up with the material.

Implementing the Course Redesign
Students were informed of the innovations in BIO 181, the reason for the changes, and the anticipated improved learning outcomes using the department’s website. In addition, a description of the changes and reasons for the changes were posted on the Blackboard VISTA course page and included in the syllabus. These changes and justification of the changes were also discussed by the instructors with the students early in the course. Students were given specific instructions on how to use Web-based technologies. It should be noted that Web-based quizzes, tutorials, and so forth were already in use across campus. Students therefore had many opportunities to learn and use the technology. There are numerous computer labs across campus and wireless capabilities in dorm rooms, classrooms, and labs. Instructional Technology Services provides a 24-h hotline to help students and faculty with technical problems.

RESULTS
We hypothesized that the redesigned course would have positive effects on student learning as evidenced by 1) lowering the DFW rate and increasing the number of As and Bs, 2) increasing consistency between sections, and 3) increasing student satisfaction.

Impact of the Redesign on Grades
Grade distribution data were collected over a 6-yr interval (3 yr prior to redesign, the year of the redesign, and 2 yr since the redesign was implemented). Figure 3 shows that over the 6-yr interval, the average DFW rate fell across sections. These changes are statistically significant, as linear regression analysis shows there is a significant negative correlation over time ($F_{1,9} = 9.659, R^2 = 0.518, p = 0.013$). Moreover, using the $t$-test, there is a significant difference between %DFW before (mean = 34.46) and after (mean = 26.25) the change ($t = 4.491, df = 7, p = 0.003$).

The percentage of Cs remained fairly constant over the 6-yr interval (Figure 3). Regression analysis shows no significant change over time ($F_{1,9} = 1.357, p = 0.274$). Furthermore, a $t$-test comparing the mean %C before (26.14) and after (22.925) the redesign showed no significant change ($t = 1.060, df = 7, p = 0.325$).

The percentage of As and Bs increased both during the redesign process and after the redesign was implemented (Figure 3). This finding is statistically significant, as linear regression analysis reveals that there is a significant positive correlation over time ($F_{1,4} = 15.329, R^2 = 0.589, p = 0.004$). Moreover, there is a significant difference between the mean %AB before (39.52) and after (50.88) the change ($t = −5.603, df = 7, p = 0.001$).

The data in Figure 3 and associated text show a significant improvement in student grades after the course redesign, and the variation between sections is addressed later in Figures 6 and 7 and associated text. However, it is formally possible that these changes merely reflect grade inflation caused by more lenient grading standards. To address this possibility, we examined pre/posttest scores. These scores were less subject to the possibility of grade inflation as the questions used were the same as those used prior to the redesign, and there was no change in the way these tests were administered before versus after the redesign. Figure 4 shows the difference between posttest and pretest scores in various sections.
Redesigning an Introductory Biology Course

sections over a 4-yr span (beginning in Fall 2006). As seen in the figure, there is a general trend toward an increase in posttest minus pretest gain after the redesign.

Analysis using a box-and-whisker plot (Figure 5) shows that these changes are substantial. In all such plots, the top of the box represents the 75th percentile, the bottom of the box represents the 25th percentile, and the line in the middle represents the 50th percentile (median). The whiskers (the lines that extend out from the top and bottom of the box) represent the highest and lowest values that are not outliers or extreme values. Outliers (values that are between 1.5 and 3 times the interquartile range) and extreme values (values that are more than three times the interquartile range) are represented by circles beyond the whiskers.

The finding that after implementation of the course redesign there is a significant improvement in posttest minus pretest scores confirms the observations from the grade data and suggests there was not a general trend toward grade inflation over this time.

Impact of the Redesign on Consistency between Sections

One of the goals of the redesign was to reduce the variability between sections of BIO 181. DFW rates were determined for each section before, during, and after the redesign. Figure 6 shows the DFW rate for various sections during this 5-yr interval. From the graph, it is clear that the variability in DFW rates during and after the redesign process is much reduced compared with the variability before the redesign. The decrease in variability between instructors is also clear from the box-and-whisker plot shown in Figure 7.

A reduction in variability between sections was also suggested by the data comparing the difference between pretest and posttest scores (see Figure 4). For example, prior to the redesign the posttest minus pretest score varied during one semester (Spring 0707) from 4.3 to 13.5 points (a 9.2-point difference). After the redesign, the maximum difference in any one semester was just three points. However, when these apparent differences were examined with a box-and-whisker plot (Figure 5), it became clear that although the median point difference increased after the redesign, the variability between sections was unchanged. Future work with this course will use these data to address issues concerning the pre- and posttest. For example, we will examine whether the pre- and posttest questions are well matched to the information that is actually taught across the various sections.

Impact of the Redesign on Student Satisfaction

A 22-item survey was administered to all four sections of BIO 181 at the beginning of the Fall 2006 semester to find out more about the students enrolled in BIO 181 and to determine whether the different sections were similar in composition. There were no significant differences in demographics among the four sections in terms of gender, ethnic make-up, or major. The Fall 2006 results established the baseline that was used to compare student satisfaction after full implementation of the course redesign. The percentage of agreement to each of the four questions was compared with the baseline data to determine whether a change in attitude had occurred from the beginning of the redesign in Fall 2006 to Spring 2010 (see Figure 8). Instructor impact on student responses was minimized by the researcher administering the anonymous survey in the lab portion of the course. All students enrolled in BIO 181 lab in Fall 2009 and Spring 2010 were surveyed.

Figure 5. Box-and-whisker plot of difference in posttest minus pretest score before (semester type 1) and after (semester type 2) the course redesign. There is a single outlier (small circle) above the box representing data from before the redesign (also apparent in Figure 4).

Figure 6. DFW rates over a 6-yr interval (11 semesters) in various sections of BIO 181. The redesign process occurred during AY 2007 and was implemented across all sections in AY 2008.
Students are required to enroll in both the lecture and lab portions of the course.

Students were asked to rate their level of agreement on 14 items related to interest, confidence, and self-efficacy. A five-point rating scale was used, ranging from strongly disagree (value = 1) to strongly agree (value = 5). Data for the following four statements are shown in Figure 8:

1. I feel confident learning basic concepts in this course.
2. I think this course material was useful for me to learn.
3. I’m confident I understood the most complex material presented.
4. Biology is difficult to understand.

Figure 8 shows that since the course redesign, students have become more confident in their abilities to learn biology (questions 1 and 3), and they have increased their perception that the material is useful for them to learn (question 2). In addition, their perception that biology is difficult to understand has decreased (question 4). The percentage of agreement difference between Fall 2006 and Fall 2009/Spring 2010 is statistically significant, as shown by a P test comparing Fall 2006 with Fall 2009 or Spring 2010, with an α of .05%. However, comparison of the two semesters since the redesign shows that there is no statistical significance at the 5% level between Fall 2009 and Spring 2010. This results shows that the revision of BIO 181 has increased student satisfaction.

DISCUSSION
The redesign of BIO 181 was conducted with the goal to improve the student success rate in this course. In this section we discuss both the benefits and the challenges of redesigning this course.

Benefits of the Redesigned Course
Our findings described above indicate that the redesigned course had a clear impact on student success rates. There are several major benefits of this improved success rate. First and most obviously, the redesigned course allows students to proceed with their course of study more quickly. This is particularly relevant as BIO 181 is a gateway course that is required for all other courses in biology. Second, there are financial implications of the redesigned course for not only the students but also the university and the department. For example, with a class size of 1000 students per year, every 1% reduction in DFW rates corresponds to approximately 10 students fewer retaking the course. We have achieved an approximately 8.5% reduction in DFW rates (from an average of 35.75% before the redesign to an average of 27.25% after the redesign), which translates to about 100 fewer students retaking BIO 181 per year. This is the equivalent of almost an entire section and therefore an additional instructor. This also
equates to four lab sections (24 students/lab section), plus more than one teaching assistant (TA; each TA teaches three sections). However, the number of lab sections has increased due to an increase in enrollment. Institutional data on the number of repeats are not available; however, the change in grade distribution implies that fewer students are repeating the course as the students who traditionally repeat the class would have earned a D or F or withdrawn from the class. Those numbers have been significantly reduced, as seen in Figures 3, 6, and 7.

A third major benefit of the redesigned course is that we can now provide a syllabus and helpful tools to new (often part-time) faculty hired to teach BIO 181 due to sabbaticals, increased enrollment, illness, and so forth. This helps new instructors to teach this course effectively and to cover material that is consistent with the goals of the course and department. The materials to teach BIO 181 are conveniently located on Blackboard VISTA and include a toolbox containing videos, animations, teaching strategies, and a PowerPoint skeleton for the various lectures.

Finally, an additional benefit of the redesign is that it led to a more collegial atmosphere among instructors of BIO 181. Instructors of the course are now more actively engaged in talking to each other about teaching strategies and sharing newly found resources such as animations.

### Challenges Faced in the Revision of BIO 181

Revising a curriculum is a difficult challenge. One issue faced initially was maintaining the autonomy of the instructors while developing a common syllabus. In addition, although most instructors realized that course adjustments needed to be made, some believed the problem could largely be resolved if we adopted a different textbook. Some instructors were also resistant to engaging in time-consuming discussions when they already felt overworked. Moreover, the instructors had very different research interests and initially had different ideas about topics that should be included. Eventually, however, the differences became an asset as the group became more comfortable with the project and focused on the task at hand. Negotiations resulted in a table of specifications that proved to be very helpful as decisions were made on what to teach and how long to spend on each concept. We found it was crucial to have open discourse and sharing of progress and even obstacles that were encountered. We found that involving all interested parties in the designing and implementation strengthened the final product and helped to reassure faculty that they were an integral part of the redesign.

There were also technical difficulties. One of the initial challenges was developing and administrating the pre/posttest. Placing the test on Blackboard VISTA was time consuming and frustrating. After administrating the pretest, it was realized that some diagrams had not uploaded successfully, a couple of wrong answers existed in the answer key, and there was a duplicate question. In addition, the 100-question test proved to be too long. The test was therefore revised to contain just 50 questions. Another technical challenge concerned the use of clickers. The classroom response systems were new to Northern Arizona University (NAU) in 2006. Learning to use them was time consuming and sometimes exasperating, and, indeed, some of the instructors tried using them but gave up.

### CONCLUSION

Curriculum revision of BIO 181 has been a long, slow process involving a plethora of meetings and decisions. The result is a complete revamping and reordering of the concepts taught, incorporation of technology, and alignment of the laboratory exercises with lecture. A major change was the incorporation of active-learning strategies such as think-pair-share, classroom response systems (clickers), and small-group work. Research has consistently shown that all students, including college science students, learn more when actively engaged (Ueckert and Gess-Newsome, 2007). Active-learning strategies mirror many of the processes of scientific inquiry and accommodate different interests and learning references.

The impact of redesigning our introductory biology course for majors was measured by comparing pre/posttest exam scores as well as grade distributions and student satisfaction. The redesign resulted in a decrease in the DFW rate, an increase in the number of As and Bs, improved posttest minus pretest scores, increased student satisfaction, and reduced variability between sections.

BIO 181 is a course that is approved for articulation with the community colleges. While we see the benefits for the redesign of BIO 181 as impacting primarily NAU, we would also like to use this opportunity to expand the conversation with local community colleges. Two community college instructors participated in the redesign team. Both teach BIO 181 at their respective institutions and have contacted us about wanting to share ideas about teaching this course. We believe that this partnership will improve the articulation of the course across institutions and may spark interest in the modification of this course beyond NAU.

Finally, we note that changing curricula is challenging and should be a continuing venture as demanded by advances in research, technology, access to Internet resources, and knowledge about teaching and learning. The response by postsecondary institutions to reform biology education appears to be slow and isolated; that is, a course or two may be revised by a department or institution, but there is no concerted effort to change biology at the national level. Perhaps the absence of a systemic effort to reform postsecondary science is due to the lack of leadership provided by the documents developed for elementary and secondary education. However, reforming the content and teaching of biology at the university level is currently receiving renewed attention largely due to the BIO2010: Transforming Undergraduate Education for Future Research Biologists and the Scientific Foundations for Future Physicians report (NRC, 2003; Association of American Medical Colleges and Howard Hughes Medical Institute, 2009). The recommendations of BIO2010: Transforming Undergraduate Education for Future Research Biologists are motivating us to pursue additional collaborations to strengthen interdisciplinary and cross-disciplinary connections in BIO 181. The anticipated goal is the incorporation of complex problem solving into both the introductory chemistry and biology courses.

### ACKNOWLEDGMENTS

We thank the BIO 181 instructors for their participation in the course redesign and collection of data. In particular, we thank Kitty Gehring,
REFERENCES


CBE—Life Sciences Education Instructions for Authors

CBE—Life Sciences Education (CBE–LSE) is an online, quarterly journal owned and published by the American Society for Cell Biology (ASCB). The journal publishes original, previously unpublished, peer-reviewed articles on life sciences education research and evidence-based practice at the K–12, undergraduate, and graduate levels. The editorial board of CBE-LSE is particularly interested in students’ conceptions and cognition related to biology and scientific inquiry and students’ engagement in the study of life sciences, as well as the efficacy of particular approaches to teaching in the life sciences. In addition, the ASCB believes that biology learning encompasses diverse fields, including math, chemistry, physics, engineering, and computer science, as well as the interdisciplinary intersections of biology with these fields. One goal of the journal is to encourage teachers and instructors to view teaching and learning the way scientists view their research, as an intellectual undertaking that is informed by systematic collection, analysis, and interpretation of data related to student learning. Target audiences include those involved in education in K–12 schools, two-year colleges, four-year colleges, science centers and museums, universities, and professional schools, including graduate students and postdoctoral researchers.

CBE-LSE is published online four times a year: March (Spring issue), June (Summer issue), September (Fall issue), and December (Winter issue). CBE-LSE also prints a highlights issue each year in December, featuring contributions selected from the online issues. Submissions are accepted at any time. Articles are assigned to particular issues by the editors. To be included in an issue, manuscripts must be accepted in final form at least two months prior to the publication date.

All published articles are available freely online without subscription. In addition, published articles are indexed in PubMed and available through PubMed Central. A goal of CBE-LSE is to stimulate dialogue. CBE-LSE invites readers to send comments that will be considered for publication in the Letters to the Editor column.

Determining the Suitability of a Manuscript for CBE-LSE

Articles. CBE-LSE publishes articles that focus on questions of science education research and practice. Articles can either break new ground in understanding teaching and learning or describe the implementation and evaluation of educational innovations. Articles must include collection, analysis, and interpretation of data, which can be quantitative or qualitative in nature. In addition, CBE-LSE articles should: 1) address a clear educational problem or education research question, 2) demonstrate clear alignment among the problem or question being addressed, the design of the study or educational innovation, the claims being made, and the evidence used to support those claims, 3) describe how results are applicable or transferable to other settings, 4) be relevant to a defined audience of educators, and 5) make reference to related educational literature. When possible, articles should include any newly described instructional resources so that they are available for use by readers. The online nature of the journal facilitates the inclusion of instructional materials such as syllabi, assignments, laboratory protocols, or professional development guidelines.

Articles detailing educational innovations must include assessment appropriate for the innovation. CBE-LSE is interested in educational innovations that have not been described extensively elsewhere, although articles about the outcomes of adopting or adapting particular approaches to teaching and learning with diverse learners and in diverse learning contexts are welcomed. Carefully constructed descriptions of relevant literature are important for demonstrating how a particular innovation is unique compared with previously published work. Instructors interested in publishing their educational innovations in CBE-LSE should give careful thought to how they will assess student learning. Articles that lack adequate assessment will be returned to authors without review.

Because the concept of assessment may be unfamiliar to many instructors, some suggestions for assessment are presented here. Answering three questions can help guide the process of assessment: 1) What are your instructional goals? 2) What should your students know or be able to do if you met your goals? 3) How can you measure or otherwise document whether your students know or are able to do what you intend? Documentation of intended outcomes can be accomplished through systematic analysis of data collected through diverse approaches, such as pretest/posttest, interviews, focus groups, surveys, or performance on coursework, including exams, papers, or lab reports. Authors should present their innovations in the same way that life scientists present their research: claims regarding efficacy should be supported by evidence.

Answering the following questions can help you determine whether your manuscript is ready for consideration by CBE-LSE: What is the background or motivation for your project? Does it meet a previously unmet need or provide a test of some educational theory? What was your goal in developing it? How do you know if you have succeeded? Can you show success to readers? Are the outcomes and supporting evidence sufficient to convince others to adapt your materials, approach, or project for their own use? Are the outcomes sufficiently novel to warrant publication?

Essays. CBE-LSE publishes essays, which describe approaches to educational challenges, including assessment methods, student engagement, curriculum innovations, K–20 continuum, and other topics. Essays are framed by personal experience and provide specific examples, but describe a problem or approach of general interest and may be synthetic across the work of many individuals. Appropriate foci for essays include reviews of current practices, policies, or research that have implications for biology teaching and learning, or personal perspectives on issues that are provocative or would otherwise be of widespread interest. The problem or approach should be presented within a scholarly context, citing references and resources that address the topic. If claims are to be made, there should be evidence from the literature or the authors’ own work. Although it is not a requirement, essays can include ideas for assessment or future research as appropriate.

Meeting Reports. The journal publishes occasional reports of meetings related to education in the life sciences. Meeting reports received through the normal online submission process will be judged on the basis of their relevance and interest for the broad readership of the journal.

Letters to the Editor. CBE-LSE welcomes letters as a means for initiating discussion of issues of interest to readers. CBE-LSE requires that letters are directly responsive to an article published in the journal or that they bring a new or underrecognized issue to the attention of readers in a way that
is informed by the inclusion of relevant data or references to the literature. Letters received through the normal online submission process will be reviewed by the Editor-in-Chief and published at her discretion. Appropriate topics for letters include commentary on education-related articles in this or other journals as well as opinions on more general topics of concern to the readership.

Features. Articles listed under the heading of “Features” are by invitation only. Authors interested in contributing to a feature should contact the editor.

Announcements. CBE-LSE accepts noncommercial announcements of meetings, workshops, and conferences and of funding opportunities and fellowships open to all.

Peer-Review Process

All submitted manuscripts and educational materials are subject to peer review. After a manuscript has been submitted to CBE-LSE, the Editor-in-Chief selects an editorial board member to guide the paper through the review process. Editorial board members select two reviewers to submit written evaluations. The board member will assess the peer reviews and determine whether the submission will be accepted as is, accepted with suggested revisions, temporarily rejected with suggestions for improvements before resubmission, or rejected with reasons explaining this decision. The corresponding author can usually expect an initial response within four weeks.

License and Publishing Agreement

Authors are required to sign a License and Publishing Agreement when a manuscript is accepted for publication. Under the terms of that agreement, authors retain copyright but grant the ASCB a perpetual license to publish the manuscript. Authors also grant to the general public the nonexclusive right to copy, distribute, or display the manuscript subject to the terms of the Creative Commons—Noncommercial—Share Alike 3.0 Unported license (http://creativecommons.org/licenses/by-nc-sa/3.0).

ASCB Policy on Research Misconduct by Authors

By submitting a paper to CBE-LSE, an author acknowledges that he or she is subject to the ASCB Policy on Research Misconduct by Authors. The policy is posted at www.ascb.org/files/research_misconduct.pdf.

Institutional Review Board Approval

Manuscripts reporting on studies involving human subjects must include explicit assurance that the research reported was approved by a local Institutional Review Board, including an Institutional Review Board number, unless the research is exempt from such review according to U.S. Department of Education guidelines. Prospective authors are advised that permission must be obtained in advance.

Guidelines for Preparing Articles, Essays, and Features

General Instructions. Submitted papers must be original (i.e., not published or submitted for publication elsewhere). Previously published material will not be considered. Publication of a short abstract does not constitute prior publication, nor does presentation of data at a scientific meeting or in a webcast of such a meeting. The posting of a manuscript on an author’s personal website or in an online institutional repository prior to submission of the manuscript to CBE-LSE is not considered prior publication. Publication of a paper in the proceedings of a scientific meeting generally does constitute prior publication.

The text should be written in clear, concise, and grammatical English. Manuscripts ordinarily begin with an overview of how the work presented is relevant to the classroom, laboratory, or curriculum and what student outcomes are expected. Whenever possible, incorporate materials by citing relevant publications, without repeating already published works. Manuscript files must be submitted in either .doc or .rtf format.

Tables and illustrations should convey information effectively. Graphs and figures should be provided digitally as separate TIFF or EPS files. Authors are encouraged to take advantage of the online nature of CBE-LSE. Video, audio, databases, images, animations, molecular structures, and other electronic resources may be linked as supplemental material for further consideration by readers.

International authors may wish to consider using an editorial service, such as ScienceDocs (www.sciencedocs.com), The Medical Editor (www.themedicaleditor.com), American Journal Experts (www.jounealExperts.com), Editage (www.editage.com), Bioscience Writers (www.biosciencewriters.com), or Squirrel Scribe (www.squirrelscribe.com). CBE-LSE does not endorse any particular service, and cannot be held responsible for the services they provide.

Length Guidelines. The following manuscript submission lengths are intended to aid authors in preparing their manuscripts; however, submissions outside these ranges will be considered.

Articles: 30,000–60,000 characters (with spaces), or 5–10 journal “pages”; typically do not exceed 20 journal pages, or 120,000 characters.

Essays: 12,000–30,000 characters (with spaces), or 2–5 journal “pages”; typically do not exceed 10 journal pages, or 60,000 characters.

Features: 6,000–12,000 characters (with spaces), or 1–2 journal “pages.”

Letters: 3,000–6,000 characters (with spaces), or up to half a journal “page.”

Cover Letter. Authors should submit a cover letter from the corresponding author stating that the work is being submitted exclusively to CBE-LSE and indicating why it is appropriate for the journal. If there is a connection between an author and a commercial product being used or reported, full disclosure is required in the cover letter and appropriate statements should be included in the manuscript. (See “Title Page” below.) Authors are invited to suggest reviewers (please include institution and email address).

Title Page. Page 1 should include the title of the manuscript, the type of manuscript being submitted (e.g., article, essay, feature, letter to the editor, response), the number of characters in the manuscript, a shortened running title (not to exceed 42 characters and spaces), the name of the corresponding author, and the names and affiliations of all other authors. For
the corresponding author, give complete postal and email addresses and telephone and fax numbers. Keywords should also appear on page 1. Include at least five keywords selected from the text of the article. If possible, include a keyword that indicates the target student group (primary, secondary, undergraduate, graduate, or postdoc) and a second keyword that indicates which cells or organism are studied in the course or lab exercise (e.g., bacteria, fungi, plants, yeast, Drosophila, nematodes).

If one or more of the authors of a research paper that assesses the effectiveness of a product or curriculum was also involved in producing the product or curriculum, readers need to be fully aware of this potential conflict of interest. Therefore, any potential conflicts of interest should be clearly stated on the title page of the manuscript. The author and the product should be identified, and a statement included that no promotion of a particular product to the exclusion of other similar products should be construed. This will be noted under the byline if the manuscript is accepted for publication.

Abstract. Page 2 should contain the abstract, which should be no more than 200 words long and should summarize the important points in the manuscript.

Manuscript Text. The text of the paper should begin on page 3. CBE-LSE follows the style guidelines of the Council of Biology Editors Style Manual. For chemical nomenclature, follow the Subject Index of Chemical Abstracts. Capitalize trade names and give manufacturer names and addresses. Do not include figures or tables within the body of the manuscript. A format of Introduction, Methods, Results, Discussion, and References is encouraged, but other formats may be more appropriate for some topics. All pages must be numbered.

Accessing Materials. Describe how to access new educational materials if the study or use of such materials is the subject of the paper. If materials are online, provide a URL to the material. Any registration requirements or agreements inherent in the use of the materials should be described. If there are no online materials, simply state “No additional materials available online.” For other new educational materials presented in the manuscript, authors should describe how readers can access the materials, what format is available (e.g., DVD, CD-ROM, PDF files, and html pages), how to request copies, and if there are any costs. CBE-LSE encourages provision of materials on a nonprofit basis, but recognizes that this is not always feasible. Please contact CBE-LSE with any questions regarding this policy.

Acknowledgments. Identify financial sources and other sources of support for the research being reported in the manuscript.

References. Place the reference list immediately following the manuscript text (beginning on a new page). Consult the most recent issue of CBE-LSE for reference citation formats. Cite references in the text alphabetically by name and date (e.g., Jones and Smith, 1987; Smith et al., 1988). Only published articles or manuscripts accepted for publication can be listed in the Reference section. References should contain complete titles and inclusive page numbers and should be listed in alphabetical order.Abbreviate the names of journals as in PubMed. The abbreviations can be found at LinkOut Journal Lists.

Citations to online websites are acceptable reference links. In the text, include the name of author for the cited website and the date accessed (e.g., National Center for Biotechnology Information, 2005). The citation should include the author (if known), date of latest modification (if known) or copyright, title of page, title of complete work (if appropriate), URL (e.g., www.website.org), and date of access. For example: National Center for Biotechnology Information (2005). NCBI Home Page. www.ncbi.org (accessed 5 February 2005).

Unpublished results, including personal communications and submitted manuscripts, should be cited in such as in the text. Personal communications must be accompanied by permission letters unless they are from the authors’ laboratory. If you cite no references, include the results of a literature search that cites your own and others’ work that has a bearing on the manuscript you are submitting and explain why no citations are needed. Include in this section only material that has been published in the literature or on the Internet.

Footnotes. Call out footnotes at the appropriate place in the text with a superscript numeral. The footnote text should be placed on a separate page after the References.

Tables. All tables must be cited in order in the text of the manuscript. Table files need to be uploaded separately.

Figure Legends. Figure legends should appear in numerical order after the References. Figure legends should provide an overview of the figure and details that describe any component parts.

Guidelines for Preparing Book Reviews

Book reviews should include the complete book title, publisher information, author detail, ISBN, and retail price. The reviewer should suggest a title for the book review, but the title should be informative and descriptive of the book to most readers.

Guidelines for Preparing Digital Artwork

Digital artwork must accompany the manuscript submission. Figures should be uploaded as separate files with the manuscript through the online manuscript submission system. Because artwork must be of sufficient quality for print reproduction, CBE-LSE asks that all artwork be prepared using professional graphic art software. Word processing and presentation software packages (such as Word and PowerPoint) are inadequate for preparing high-quality digital artwork.

Prepare all digital artwork as RGB TIF images, at 300 dpi resolution or EPS images:

Figure Size. Prepare figures at the size they are to be published.

Up to 1 column wide: Figure width should be 4.23–8.47 cm.
1 to 1.5 columns wide: Figure width should be 10.16–11.43 cm.
2 columns wide: Figure width should be 14.39–17.57 cm.
The figure depth must be less than or equal to 23.5 cm.

Guidelines for Preparing Electronic Resources

It is possible for authors to submit for peer review electronic works including, but not limited to, animations, Chime tutorials, movies, interactive websites that may include quizzes,
images (electron micrographs, photomicrographs, etc.), Java Applets, searchable databases, etc. Articles that describe new educational uses of existing resources (e.g., Expression Connection, FlyBase, Database of Interacting Proteins, etc.) are also of interest. Manuscripts of this type should provide detailed instructions for use of the resource by the target audience.

For works using someone else’s electronic resource (such as a database), a letter from the creator or curator of the resource indicating their willingness to support free pedagogical use of their work must be included. The same general rules for evaluation will apply to all electronic submissions. All submissions will be evaluated for (1) pedagogical content, (2) clear description of goals and expected student outcomes, (3) transferability to other settings, (4) appropriateness for the target audience, and (5) references to related educational literature. After publication of the electronic work(s), the authors will be encouraged to submit the work to other databases (National Digital Library, BEN, etc.); however, it is expected that the CBE-LSE publication citation will remain associated with the work. This will allow viewers to read a more in-depth discussion of the work.

All such electronic works must be freely available, and will be hosted on the CBE-LSE server or on the ASCB server, with the exception of large databases. This will ensure stable access to the works with a nonchanging URL.

A manuscript should accompany any electronic submission. The manuscript should describe 1) the learning goals or purpose of the electronic work, 2) the target audience, 3) development of the electronic work (describe hardware and software used), 4) platform availability (see below), 5) a description of any necessary hardware or software, with links to the appropriate sites for downloading (e.g., plugins, helper applications, etc.), and 6) assessment of the work’s impact on student learning.

Ideally, submitted works should work on any platform (PC, Mac, Unix) and on all browsers. If there are known restrictions, these should be included in the manuscript. CBE-LSE can help authors test their works for such limitations if they do not have access to certain platforms or browsers. Assessment may come in the form of learning gains (e.g., pre- and posttests) and/or changes in student attitudes toward the electronic work. Length of time spent using the electronic work may be one indicator of student attitudes. If it is possible to measure student responses to online questions, these responses (number of submissions before correct answer, typical errors, etc.) may further a reader’s understanding of how the electronic work could be incorporated into his or her teaching.

The electronic work may have been hosted previously on any website, but the authors may not have previously published any description of the electronic work other than the associated Web pages. Published journal descriptions of the electronic work will preclude publication in CBE-LSE, with the exception of abstracts or presentations at professional meetings.

The manuscript should include an abstract and any figures deemed helpful when the published work appears as a PDF file.

The online publication will include hyperlinks to the work that will appear in a new browser window, if appropriate.

This capacity could be helpful to the authors since they could provide directions for readers as needed to illustrate particular aspects of the work. The layout and submission process for the manuscript accompanying an electronic work should follow the same general format as other categories. The electronic work should be submitted to CBE-LSE at the same time as the manuscript. If this presents a problem, contact CBE-LSE staff for assistance.

Sharing Materials and Data

Publication of a manuscript in CBE-LSE implies that the authors agree to make available, to the extent legally permissible, all propagative materials such as mutant organisms, cell lines, recombinant plasmids, vectors, viruses, and monoclonal antibodies that were used to obtain results presented in the article. Prior to obtaining these materials, interested scientists will provide the authors with a written statement that they will be used for noncommercial research purposes only. Instructional materials that were developed in the course of a project described in a CBE-LSE article should also be made freely available, preferably in electronic format. Authors are encouraged to share raw data with qualified researchers who wish to reproduce or further analyze the authors’ work for noncommercial purposes. Sharing of data on human subjects should be consistent with the conditions of the Human Subjects protocol approved by the authors’ Institutional Review Board (and any other agreements made with the subjects) for the work reported.

How to Submit Manuscripts

Electronic Submission. Authors must submit manuscripts online at www.cellbiologyeducation.org. Specific instructions on how to submit your manuscript are available at the submission site. Authors should submit the manuscript in either .doc or .rtf as two text files, the first containing the cover letter and the second containing the manuscript file. Figures, tables, and Supplemental Material, including Appendixes, must be submitted as individual files. Do not embed figures in the manuscript. If you are submitting a feature that does not contain an abstract or keywords, write “There is no abstract” in the required abstract field and “one,” “two,” and “three” in the keyword fields.

Questions regarding submission guidelines can be directed to: cbe@ascb.org or 301-347-9304.

General Questions

At any stage in the submission process, authors with questions should contact the CBE-LSE Editorial Office at 301-347-9304 (phone); 301-347-9350 (fax); cbe@ascb.org; or The American Society for Cell Biology, 8120 Woodmont Avenue, Suite 750, Bethesda, MD 20814-2762.

You are encouraged to contact editorial board members or the Editor-in-Chief by email to discuss submissions to CBE-LSE.

The CBE-LSE website is an additional resource to authors; see earlier issues of CBE-LSE for examples of the different types of manuscripts published.
Human blood cells (red blood cells and granulocytes, latter 0.015 mm)

Early embryos of the African clawed frog (Xenopus laevis, 1.3 mm)

Sea urchin embryos (Lytechinus pictus) in second cell division (0.1 mm)

Early zebrafish embryo (Danio rerio), 2-cell stage, shell removed (0.5 mm)

Mouse embryonic stem cells (0.025 mm)

Elodea cell (0.025 mm x 0.05 mm)

Human red blood cells (0.008 mm)

Amoeba proteus stained with pH-dependent dye (0.7 mm)

Human cheek cells stained with methylene blue (0.03 mm)

Questions for Stimulating Inquiry*

There are many types of cells, and although they have numerous similarities, they also display profound differences in shape, structure, and function. These images highlight contrasts and commonalities across a spectrum of organisms. Use them as a springboard for deeper exploration of cells and biological processes.

1. Which panels display cells that are differentiated, that is, specialized for particular functions? How many of their functions can you name? Would you argue that the amoeba is a differentiated cell, or not?

2. Which panels display cells that are not differentiated? As they continue to divide, will some of their descendants become differentiated? What kinds of differentiated cells will they be?

3. How many different kinds of organelles can you identify in these images? What are their functions?

4. How do the forms of the cells in these images relate to their functions? For example, why do the two different types of blood cells have different shapes? How about the others?

5. Using the cell diameters given in the captions, how can you determine the final magnification of each image?

*For answers and discussion, see the CBE-LSE website: www.lifescied.org

All images taken by Eliott K. Li, Life Sciences Education Media, Laboratory for Physical Sciences, University of Maryland College Park. Frog embryos stained using fluorescence microscopy techniques. Amoeba and Elodea images taken by Robert Shumway, Center for Biological Image Science and Molecular Informatics (CBIS), A Laboratory of the National Institutes of Health (NIH), and by Monty Pernis and Ruth W. Fulbright.
The HHMI Exceptional Research Opportunities Program aims to increase the diversity of the science professoriate. We match well-prepared undergraduate students with HHMI scientists in a summer-long mentored research experience. Since 2003, EXROP has sponsored 518 students from 115 colleges and universities who worked with 175 HHMI scientists. More than three-fourths of the students are underrepresented minorities. Of the 301 EXROP alumni who have completed the baccalaureate degree, 283 are in science careers including 143 who are in graduate school.

Here are the EXROP students who participated in summer of 2011.

David Albores  
Cornell University

David Baker  
University of Washington

Jawara Allen  
Duke University

Dan R. Littman  
New York University

Jorge Alvarado  
Amherst College

Thomas C. Südhof  
Stanford University School of Medicine

Visesh Appadoo  
Vassar College

Richard N. Zare  
Stanford University

Mark Araujo  
Emory University

Charles S. Zuker  
Columbia University College of Physicians & Surgeons

María Ayala Ramírez  
Mount Holyoke College

Brendan Lee  
Baylor College of Medicine

Julian Baptiste  
University of Pittsburgh

Susan Ferro-Novick  
UC San Diego School of Medicine

Jose Gabriel Barcia Duran  
Yale University

Shahin Rafii  
Cornell University Joan & Sanford I. Weill Medical College

Donald Bryant  
Emory University

Richard Lee  
Brigham and Women’s Hospital

Stephen Canton  
University of Pittsburgh

Hugo J. Bellen  
Baylor College of Medicine

Isadora Corulio  
Columbia University

Michael Rosbash  
Brandeis University

Jun Chen  
City University of New York Hunter College

Martin D. Burke  
University of Illinois at Urbana-Champaign

Jessica Coates  
Spelman College

Daniel L. Balslev  
University of Texas at Austin

Nicolas Contreras  
Gonzaga University

Arthur L. Horwich  
Yale School of Medicine

Joseph Donahue-Torres  
University of Colorado at Boulder

Arthur Weiss  
UC San Francisco

Sabrina Dumas  
University of Arizona

Tamir Gonen  
University of Washington

Ricardo Dunwoodie  
Morehouse College

Joachim Frank  
Columbia University College of Physicians & Surgeons

Jessica Franco  
University of Arizona

Joseph L. DeRisi  
UC San Francisco

Kimberly Frutoz  
UC Los Angeles

Elliot Chaikof  
Beth Israel Deaconess Medical Center

Ariel Garcia  
Purdue University

Elizabeth C. Engle  
Children’s Hospital Boston

Sarah Gordon  
Columbia University

Craig G. Mello  
University of Massachusetts Medical School

Daniel Gutierrez  
San Jose State University

Paul Alquist  
University of Wisconsin-Madison

Usman Hameedi  
Brandeis University

John Kuriyan  
UC Berkeley

Hayden Hatch  
Harvey Mudd College

Eric R. Kandel  
Columbia University College of Physicians & Surgeons

Rebecca Hernandez  
Rice University

Erol Fikrig  
Yale School of Medicine

Leslie Higueta Montoya  
Furman University

Scott Keeney  
Memorial Sloan-Kettering Cancer Center

James Hill  
Duke University

Norbert Perrimon  
Harvard Medical School

Meryl Horn  
City University of New York Hunter College

Li-Huei Tsai  
Massachusetts Institute of Technology

Tien-Phat Huynh  
UC Los Angeles

Tom A. Rapoport  
Harvard Medical School

Odadyin Irianye  
North Carolina Central University

Alex L. Kolodkin  
The Johns Hopkins University School of Medicine

Michael Ishida  
California State University Fullerton

Jennifer A. Doudna  
UC Berkeley

Sharose Jefferson  
College of William and Mary

Michael R. Green  
University of Massachusetts Medical School

Sorna Kamara  
Georgetown College

Upul Banerjee  
UC Los Angeles

Kayla Lee  
Hampton University

Ronald R. Breakey  
Yale University

Angel Leiva  
University of Miami

Michele Pagano  
New York University

Tiffany Lemon  
Louisiana State University

Bruce D. Walker  
Massachusetts General Hospital
Diana Lenis  
Stony Brook University  
Joan A. Steitz  
Yale School of Medicine  
Jian Li  
Yale University  
Andrew P. McMahon  
Harvard University  
Kristine Lyon  
Lewis and Clark College  
David D. Ginty  
The Johns Hopkins University School of Medicine  
Sesath Mack  
Howard University  
Thomas R. Cech  
University of Colorado at Boulder  
Richard Madanu  
University of Maryland Baltimore County  
Matthew P. Scott  
Stanford University School of Medicine  
Brenda Marin-Rodriguez  
UC Davis  
Catherine Dulac  
Harvard University  
Carlos Marti  
University of Puerto Rico Mayaguez Campus  
Axel T. Brunger  
Stanford University  
Rey Martin  
UC Los Angeles  
Rebecca Richards-Kortum  
Rice University  
Rosa Martinez  
University of Puerto Rico Mayaguez Campus  
Chris Q. Doe  
University of Oregon  
Daniel Mendez  
California State University Fullerton  
Stanley Fields  
University of Washington School of Medicine  
Emmanuel Mendoza  
UC Davis  
Leonard I. Zon  
Children’s Hospital Boston  
Cristina Murphy  
Louisiana State University  
Richard Axel  
Columbia University College of Physicians & Surgeons  
Martin Mwangi  
UC Los Angeles  
Graham F. Hatfull  
University of Pittsburgh  
Marla Nguyen  
Oregon State University  
Joseph G. Gleeson  
UC San Diego School of Medicine  
Kaitlyn Norman  
New Mexico State University  
James J. Collins  
Baston University  
Fiona Odu  
University of Missouri Columbia  
Luís A. Nunes Amaral  
Northwestern University  
Alejandra Olvera  
Weislyan University  
Pamela J. Bjorkman  
California Institute of Technology  
Fabian Ortega  
Yale University  
David G. Schatz  
Yale School of Medicine  
Jennifer Ortega  
Harvard College  
Louis J. Piché  
UC San Francisco  
Gorneva Parkinson  
Howard University  
Kornelia Polyak  
Dana-Farber Cancer Institute  
Carlos Pelambo  
UC San Diego  
Susan M. Koech  
Yale University  
Fernando Pimentel  
Stanford University  
Stephen P. Gaff  
Columbia University College of Physicians & Surgeons  
Karolina Puskarczyk  
Duke University  
Joes Teixeira  
Massachusetts General Hospital  
Sofia Quinodoz  
Princeton University  
Michael B. Elowitz  
California Institute of Technology  
Ryan Quiroz  
UC Los Angeles  
Stuart L. Schreiber  
Broad Institute  
Gabriel Rangel  
Purdue University  
Christopher P. Lowe  
University of Maryland Baltimore, School of Medicine  
Samuel Regalado  
UC Berkeley  
Amy Wagers  
Harvard University  
Donald Richards  
UC Riverside  
Leslie A. Leinwand  
University of Colorado at Boulder  
Sharnara Robinson  
Spelman College  
David Pellman  
Dana-Farber Cancer Institute  
Lauren Rodriguez  
UC Santa Cruz  
William R. Jacobs Jr.  
Albert Einstein College of Medicine  
Juan Ruiz  
University of Miami  
Jeffrey Karp  
Brigham and Women’s Hospital  
Talibat Salami  
Wellesley College  
Steven A. Siegelbaum  
Columbia University College of Physicians & Surgeons  
Angelica Simmons  
Louisiana State University  
Yiwan Cheung  
Children’s Hospital Philadelphia  
Charesa Smith  
Harvard University  
Stephen C. Harrison  
Harvard Medical School  
Patrice Starck  
University of Maryland Baltimore County  
Bruce D. Walker  
Massachusetts General Hospital  
Jordyn Talley  
University of North Texas  
Joseph S. Takahashi  
University of Texas Southwestern Med Center Dallas  
Tigist Tamir  
College of William and Mary  
Randall T. Moon  
University of Washington School of Medicine  
Theodor Uzumeri  
Morehouse College  
Tyler Jacks  
Massachusetts Institute of Technology  
Arielle Valdez  
Boston University  
Richard L. Hugginer  
The Johns Hopkins University School of Medicine  
Alana Van Dervort  
Florida International University  
George Q. Daley  
Children’s Hospital Boston  
Yorleny Vicioso  
Stony Brook University  
Tim Stearns  
Stanford University  
Isaiah Woodson  
Louisiana State University  
David E. Clapham  
Children’s Hospital Boston  
Danielle Wright  
Davidson College  
David R. Walt  
Tufts University  
Paul Yarabe  
Harvard University  
Huda Y. Zoghbi  
Baylor College of Medicine  
Martha Adriana Zepeda Rivera  
University of Washington  
Matthew K. Waldor  
Brigham and Women’s Hospital

http://www.hhmi.org/grants/individuals/exrop.html
ESSAYS

Improving Graduate Education to Support a Branching Career Pipeline: Recommendations Based on a Survey of Doctoral Students in the Basic Biomedical Sciences
C. N. Fuhrmann, D. G. Halme, P. S. O'Sullivan, and B. Lindstaedt.......................................................................43–53

Teaching Biology through Statistics: Application of Statistical Methods in Genetics and Zoology Courses
Migdalisel Colon-Berlingeri and Patricia A. Burrowes ..........................................................................................54–62

ARTICLES

Increased Course Structure Improves Performance in Introductory Biology
Scott Freeman, David Haak, and Mary Pat Wenderoth...........................................................................................63–74

Culturally Relevant Inquiry-Based Laboratory Module Implementations in Upper-Division Genetics and Cell Biology Teaching Laboratories
Dimuth Siritunga, María Montero-Rojas, Katherine Carrero, Gladys Toro, Ana Vélez, and Franklin A. Carrero-Martínez .................................................................75–85

Investigation of Science Faculty with Education Specialties within the Largest University System in the United States

A Study Assessing the Potential of Negative Effects in Interdisciplinary Math–Biology Instruction
Andreas Madlung, Martina Bremer, Edward Himelblau, and Alexa Tullis .....................................................104–115

Teaching Students How to Study: A Workshop on Information Processing and Self-Testing Helps Students Learn
Kathrin F. Stanger-Hall, Floyd W. Shockley, and Rachel E. Wilson ......................................................................116–127

Combining Peer Discussion with Instructor Explanation Increases Student Learning from In-Class Concept Questions

Using Digital Images of the Zebra Finch Song System as a Tool to Teach Organizational Effects of Steroid Hormones: A Free Downloadable Module
William Grisham, Natalie A. Schottler, Lisa M. Beck McCauley, Anh P. Pham, Maureen L. Ruiz, Michelle C. Fong, and Xinran Cui ...................................................................................137–145

Redesigning a Large-Enrollment Introductory Biology Course
Catherine Ueckert, Alison Adams, and Judith Lock............................................................................................146–156

The journal is supported in part by an Undergraduate Science Education Program grant from the Howard Hughes Medical Institute.

8120 Woodmont Ave., Suite 750, Bethesda, MD 20814, USA; Phone: 301-347-9300; Fax: 301-347-9310; www.ascb.org