

Article

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

Andreas Madlung,* Martina Bremer,[†] Edward Himmelblau,[‡] 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.

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

(Iossi, 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 International 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

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

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

^a 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.

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

A

EXAMPLE QUESTIONS FROM THE BIOLOGY-ONLY MODULE.

Question 5: Assuming you use a p-value of 0.05 as your "cut-off", would you reject or accept the null hypothesis if your calculated p-value is 0.04?

Question 6: The t-test method testing whether the log-ratio of a gene is significantly different from zero depends on computing the standard deviation s of several M-values for each gene. Suppose you have a microarray slide which has 20,000 genes on it but each is spotted only once. In this case, is it possible to compute standard deviation of the M-value for a gene? Why or why not.

B

#3: Conduct a t-test for a microarray experiment.

We need to make a decision for each gene (represented by several spots on the array). Is the gene expressed differently in the treatment and control group? Therefore, for each gene, we will carry out the hypothesis test separately.

Before the Test: Pick a gene. Find all the red and green intensity values for this gene in your data file. Compute all the M-values for this gene.

Step 1: We need to set up the null hypothesis and alternative hypothesis. Remember that the null hypothesis means that the treatment had no effect (on this gene) and the alternative hypothesis is what the researcher is really expecting to support. A good experiment will always provide valuable information, regardless of the outcome of the hypothesis test.

In your answer sheet write down the two hypotheses for your experiment:

Null Hypothesis:

Alternative Hypothesis:

Step 2: In this example we have already collected the data. We will use the microarray measurements as estimates for the gene expression levels in the two groups.

Step 3: We want to find a p-value for the gene. That means that we have to compute a test-statistic value and then decide how usual or unusual it is.

Example: Suppose the collected (and normalized) microarray data of gene At1g01000 looks like this:

We have six M-values for the gene At1g01000 in this table. Compute the average of the six observations: \bar{x} = _____ (fill in value) and the standard deviation s = _____ (fill in value) (see formula on p 3 or calculate in Excel or with your calculator).

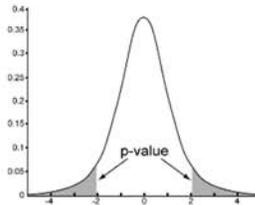
Since we have six observations, $n = 6$. Now we can compute the value of the test statistic as:

$$t = \frac{\bar{x}}{\frac{s}{\sqrt{n}}} = \frac{-1.15}{\frac{1.28}{\sqrt{6}}} = -2.08$$

| Block | Column | Row | Name | M |
|-------|--------|-----|------------|-------|
| 1 | 1 | 3 | At1g001000 | -0.16 |
| 1 | 2 | 2 | At1g001000 | -1.89 |
| 1 | 4 | 2 | At1g001000 | -1.05 |
| 2 | 1 | 3 | At1g001000 | 0.58 |
| 2 | 2 | 2 | At1g001000 | -3.05 |
| 2 | 4 | 2 | At1g001000 | -1.35 |

Calculate t and fill in value here: _____

The degree of freedom that describes the behavior of the test statistic in this case is _____. To find the p-value, we have to find the percentage of cases, in which the t-test statistic with $df = 5$ would take on more extreme values than the t-value we observed. Extreme values are the ones far away from zero (see graph to the right).



In the past, these values had to be looked up in tables. Today, Excel and other software programs have them stored in their statistics package. The p-value can be found with the Excel command "=TDIST(2.08, 5, 2)". In the example above, the exact p-value (red shaded tail area of the distribution) is 0.0921 or 9.21%.

To find a p-value using Excel, open an Excel spreadsheet, click on any empty cell and enter "=TDIST(absolute value of your test statistic, df, 2)". The "2" stands for two-sided, which means that you want the red area in both tail ends. In this example the absolute value (no minus sign) of the test statistic is 2.08 and the degree of freedom is $df = 6 - 1 = 5$.

Now calculate the p-value using the t-value you calculated above.

Step 4: What conclusion can we draw? The p-value is the probability to observe data as extreme/unusual as the one we saw if the gene expression in the two groups were the same. Our p-value 9.21% is quite large (bigger than 5%). That means that we would get observations such as these by random chance and not due to real difference in gene expression almost 10% of the time. Hence, our data is nothing unusual and we accept the null hypothesis (equal expression in both groups) for gene At1g01000.

Results of Experiment: To determine the p-values for the other genes spotted on the microarray you would repeat steps 1 - 4 above. This would provide us with a p-value for each gene on the array.

| Gene name | p-value | Differentially expressed at level 5%? |
|-----------|---------|---------------------------------------|
| At1g01000 | 0.0921 | |

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.polyplody.org/index.php/Microarray_analysis. (Continued)

(data not shown), suggesting that there was no systematic difference in the students comprising each section, enabling us to pool the data for a final sample size of 27 students. Of these 27 students, 93% had taken a course in ecology (BIOL 211), which provided them with previous instruction in using statistics to analyze biological data.

We made a few minor adjustments to the module developed for first-year students described previously to make it

appropriate for a smaller class of advanced biology students. The first modification concerned the time when the module was administered during the course. The module was administered to advanced students following several weeks of molecular biology instruction, and after they had gained some wet-lab experience in microarray work. The second modification was necessitated by the smaller number of total students in the advanced course as compared with the

C

QUIZ QUESTIONS PERTAINING TO THE MATERIAL IN 1A AND B ADMINISTERED TO FIRST-YEAR STUDENTS.

3. The "null hypothesis" in microarray data ... (Circle all that apply)
- states that a gene statistically has the same expression in the treatment as in the control group.
 - states that all genes have different expression levels in the treatment and the control group.
 - assumes that green and red dyes in a dye swap analysis incorporate into the same cDNAs of the probe at different rates.
 - states that the expression value for the same gene in treatment versus control groups does not equal 0.
 - states that the mean of all repeated measurements of the same gene as it appears on an M/A plot lies either significantly above or below the 0 line.
4. Assume that the statistical analysis of several repeated data points for the same gene from several arrays results in a p-value of >0.05. This means that... (Circle all that apply)
- the alternative hypothesis can be rejected, which means that the control and treatment group are indeed different in their gene expression.
 - according to the data, either the treatment or the control group does not contain the gene in question in its DNA.
 - the null hypothesis can be rejected, which means that the control and treatment group are indeed different in their gene expression.
 - the null hypothesis cannot be rejected, which means the control and treatment group are not different from each other in their gene expression.
 - statistically, the question whether or not the gene from the control and treatment group are different in their expression cannot be answered.

Figure 1. Continued.

A

EXCERPT FROM THE LEARNING MODULE FOR THE ADVANCED BIOLOGY COURSE.

Exercise: Conduct the dye-swap normalization. For each spot, we need to compute the average M-value for both arrays. To do this, include another column and label it "corrected M". Click into the first cell and enter "=AVERAGE(G3,K3)".

| | A | B | C | Array 1 | | | Array 2 | | | | L | |
|---|-------|--------|-----|------------|------------------|------------------|---------|------------|------------------|------------------|-------|-----------------|
| | Block | Column | Row | Name | F635 Median-B635 | F532 Median-B532 | M | Name | F635 Median-B635 | F532 Median-B532 | M | corrected M |
| 3 | 1 | 2 | 3 | At1g001010 | 1 | 2 | -1.000 | At1g001010 | 4 | 3 | 0.415 | =AVERAGE(G3,K3) |
| 4 | 2 | 2 | 3 | At1g001000 | 3 | 5 | -0.737 | At1g001000 | 2 | 2 | 0.000 | |

The data file is getting quite big at this point. We want to decide for each gene, whether it is differentially expressed in the treatment and control group. All the information we need to do this are the gene name and all the corrected M-values for this gene.

Exercise: Using Excel compute the average, Standard Deviation, test-statistic and p-value for the first gene AT1g00100. (The Excel commands needed are AVERAGE, and STDEV). Do these calculations in your Excel sheet in the columns next to "Corrected M". First compute the average, then the SD, then type in a formula for t as you find it below. (Ask for help if needed). The (rounded) value of the test statistic then becomes:

$$t = \frac{\bar{x}}{\sqrt{\frac{s^2}{n}}} = -8.684$$

df = n-1 = 2-1 = 1. Click on any empty cell in Excel and type "=TDIST(your t-value,1,2)" to obtain the p-value (rounded to three digits). Show the p-value and your computations to me before proceeding.

B

QUIZ QUESTIONS PERTAINING TO THE MATERIAL IN 2A ADMINISTERED TO ADVANCED STUDENTS

6. In microarray analysis thousands of data points are analyzed resulting in thousands of p-values after statistical analysis has been performed. This results in "multiple testing" problems because... (Circle all that apply)
- ... due to the large number of genes analyzed, the p-value for each gene is less likely to reflect true statistical (in-)significance for the gene that it is assigned to.
 - ... due to the large number of genes analyzed, the likelihood that among the p-value assigned several are false positives is close to 100%.
 - ... due to the large number of genes analyzed, the likelihood that among those p-value are false positives is equal to or less than 5% (0.05).
 - ...only one p-value is calculated to reflect the significance for all represented genes on the array.
 - ... each p-value has to be re-calculated as many times as there are genes on the array.
7. The Bonferroni correction method... (Circle all that apply)
- ... calculates p-values not by the use of t-statistics but by assigning ranks to each gene, and dividing the rank by the total sample size.
 - ... is basically a method that due to the constraints of multiple testing issues of thousands of genes considers p-values as significant even when they are greater than 0.05.
 - ... is basically a method that due to the constraints of multiple testing issues of thousands of genes considers p-values as significant only when they are much smaller than 0.05.
 - ... does not rely on a p-value but compares each gene to each other on the array in pairwise comparisons to eliminate the problems of multiple testing.
 - ... calculates a more stringent p-value because it can apply a greater degree of freedom due to the thousands of replicated data points on each array.

Figure 2. Example from the microarray learning module for the advanced students in the Plant Molecular Biology and Physiology course. The problem shown in panel (A) parallels that shown in Figure 1, A and B, for first-year biology students. The corresponding assessment questions are shown in (B). Complete classroom material can be found at www.polypldoidy.org/index.php/Microarray_analysis.

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.

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.polyploidy.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

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

| Class (N) | Stand-alone quiz questions ^a | | Embedded final examination questions ^b | |
|-----------|---|-------------|---|-------------|
| | Passive math | Active math | Passive math | Active math |
| 1 (42) | 82.8 ± 9.6 | 82.0 ± 10.4 | 72.0 ± 11.6 | 76.0 ± 10.6 |
| 2 (39) | 76.8 ± 12.0 | 76.0 ± 10.9 | 67.6 ± 11.4 | 68.4 ± 9.9 |
| 3 (36) | 70.4 ± 10.5 | 72.4 ± 8.2 | 68.4 ± 13.3 | 71.2 ± 9.1 |
| 4 (42) | 79.2 ± 6.0 | 76.4 ± 8.4 | 71.2 ± 7.8 | 67.6 ± 10.6 |

% scores are out of 25 points.

^aThere 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).

^bThere was no significant difference in the scores of the passive math and active math groups on assessment questions embedded in the final examination for any of the four introductory biology classes (two-sided *t* test, *p* = 0.26, 0.86, 0.51, and 0.20 for classes 1, 2, 3, and 4, respectively).

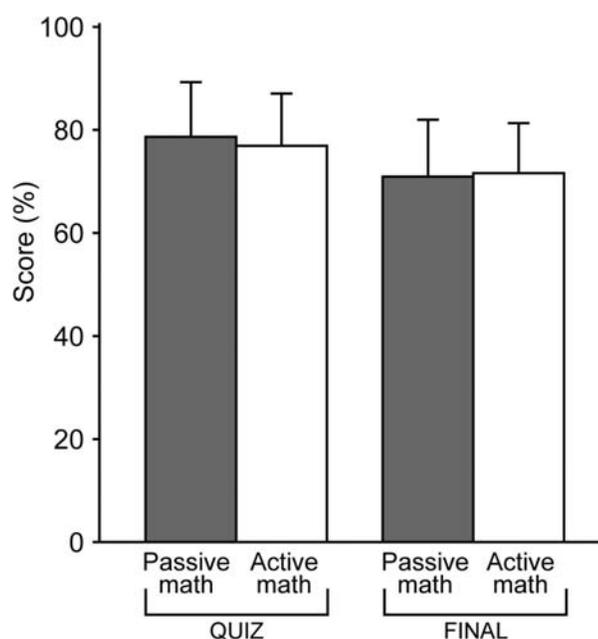


Figure 3. Adding an interdisciplinary quantitative component to the biology curriculum did not adversely affect performance of first-year students on assessment questions. During a unit on the use of microarrays in biological research, biology students were given practice problems in which they focused 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* = 0.828, *p* = 0.409) or on the final examination questions (*t* = -0.213, *p* = 0.832). 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.

groups scored significantly lower on the embedded questions compared with the stand-alone questions (paired *t* tests; passive math: *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 3. Student scores (% ± SD) on assessment questions related to a microarray learning module for advanced students

| Class (N) | Premodule | Postmodule |
|-----------|-------------|--------------------------|
| 2007 (12) | 77.2 ± 7.7 | 76.0 ± 8.9 |
| 2008 (8) | 80.0 ± 12.3 | 88.9 ± 12.5 ^a |
| 2009 (7) | 71.4 ± 6.9 | 91.4 ± 7.0 ^a |

% scores are out of 36 points.

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

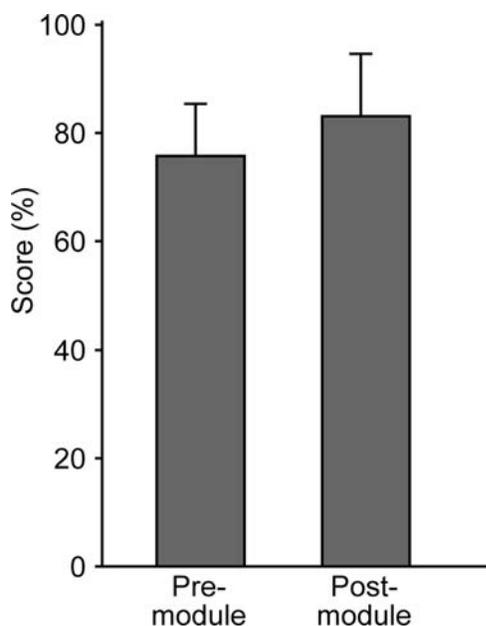


Figure 4. An interdisciplinary computer module requiring students to calculate statistical problems by hand was an effective learning tool for advanced students. Students in an upper-level Plant Molecular Biology course were asked to perform statistical computations in a lab exercise after a lecture 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.

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.

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

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

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 enroll 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 *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; Sutarso, 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

- Aiken L (1963). Personality correlates of attitude toward mathematics. *J Educ Res* 56, 476–480.
- Ares M (2004). Interdisciplinary research and the undergraduate biology student. *Nat Struct Mol Biol* 11, 1170–1172.

- Arnett A, Van Horn D (2009). Connecting mathematics and science: a learning community that helps math-phobic students. *J Coll Sci Teach* 38, 30–34.
- Ashcraft MH (2002). Math anxiety: personal, educational, and cognitive consequences. *Curr Dir Psychol Sci* 11, 181–185.
- Ashcraft MH, Krause JA (2007). Working memory, math performance, and math anxiety. *Psychon Bull Rev* 14(2), 243–248.
- Bessant KC (1995). Factors associated with types of mathematics anxiety in college students. *J Res Math Educ* 26, 327–345.
- Bialek W, Botstein D (2004). Introductory science and mathematics education for 21st-century biologists. *Science* 303, 788–790.
- Bransford JD, Brown AL, Cocking PR (eds) (2000). *How People Learn: Brain, Mind, Experience, and School*, Washington, DC: National Research Council/National Academies Press.
- Bremer M, Himelblau E, Madlung A (2010). Introduction to the statistical analysis of two-color microarray data. In: *Statistical Methods in Molecular Biology*, eds. H Bang, XK Zhou, HLV Epps, and M Mazumdar, New York: Humana Press, 285–311.
- Brent R (2004). Points of view: the interface of mathematics and biology: intuition and innumeracy. *Cell Biol Educ* 3, 88–89.
- Brown TS, Brown JT (1995). Prerequisite course grades and attitudes towards statistics. *College Student J*, 29, 502–507.
- Campbell AM (2001). DNA Microarray Methodology—Flash Animation. www.bio.davidson.edu/Courses/genomics/chip/chip.html (accessed 20 October 2010).
- Campbell AM, Zanta CA, Heyer LJ, Kittinger B, Gabric KM, Adler L (2006). DNA microarray wet lab simulation brings genomics into the high school curriculum. *Cell Biol Educ* 5, 332–339.
- Darwin F (1995). *The Life of Charles Darwin*, West Molesey, Surrey, UK: Merchant Book Company, Ltd.
- Evans BR (2007). Student attitudes, conceptions, and achievement in introductory undergraduate college statistics. *The Mathematics Educator* 17(2), 24–30.
- Gross LJ (2004). Points of view: the interface of mathematics and biology: interdisciplinarity and the undergraduate biology curriculum: finding a balance. *Cell Biol Educ* 3, 85–87.
- Hodgson T, Keck R, Patterson R, Maki D (2005). Mathematics in the biology classroom. A model of interdisciplinary education. *J Coll Sci Teach* 34, 46–50.
- Iossi L (2007). Strategies for reducing math anxiety in post-secondary students. In: *Proceedings of the Sixth Annual College of Education Research Conference: Urban and International Education Section*, eds. SM Nielsen and MS Plakhotnik, Florida International University, Miami, 30–35.
- Jungck J (2005). Challenges, connection, complexities: education for collaboration. In: *Math and Bio 2010: Linking Undergraduate Disciplines*, ed. LA Steen, Washington, DC: The Mathematical Association of America, 1–12.
- Katz V (2003). Meeting the Challenges: Education across the Biological, Mathematical, and Computer Sciences. The Mathematical Association of America. www.maa.org/features/mtc.html (accessed 28 September 2010).
- Kilpatrick J, Swafford J, Findell B (2001). *Adding It Up: Helping Children Learn Mathematics*, Washington, DC: National Academies Press.
- Lenning O, Ebberts L (1999). The powerful potential of learning communities: improving education for the future. *ASHE-ERIC Higher Education Report* 26(6).
- McEwen LA, Harris D, Schmid RF, Vogel J, Western T, Harrison P (2009). Evaluation of the redesign of an undergraduate cell biology course. *CBE Life Sci Educ* 8, 72–78.

Metz AM (2008). Teaching statistics in biology: using inquiry-based learning to strengthen understanding of statistical analysis in biology laboratory courses. *CBE Life Sci Educ* 7, 317–326.

Mills JD (2004). Students' attitudes toward statistics: implications for the future. *College Student J* 38(3), 349–361.

National Research Council (2003). *BIO 2010: Transforming Undergraduate Education for Future Research Biologists*, Washington, DC: National Academies Press.

National Science Foundation (2009). *Interdisciplinary Training for Undergraduates in Biological and Mathematical Sciences*, Program Solicitation NSF08–510.

Perney J, Ravid R (1990). The relationship between attitudes towards statistics, math self-concept, test anxiety and graduate students' achievement in an introductory statistics course. Annual Meetings of the American Educational Research Association, Boston, MA.

Perry AB (2004). Decreasing math anxiety in college students. *College Student J* 38, 321–324.

Pursell DP (2009). Enhancing interdisciplinary, mathematics, and physical science in an undergraduate life science program through physical chemistry. *CBE Life Sci Educ* 8, 15–28.

Richland LE, Zur O, Holyoak KJ (2007). Cognitive supports for analogies in the mathematics classroom. *Science* 316, 1128–1129.

Robeva R, Laubenbacher R (2009). Mathematical biology education: beyond calculus. *Science* 325, 542–543.

Suinn RM, Winston EH (2003). The Mathematics Anxiety Rating Scale, a brief version: psychometric data. *Psychol Rep* 92, 167–173.

Sutarso T (1992). Some variables in relation to students' anxiety in learning statistics. In: Annual Meeting of the Mid-South Educational Research Association, Knoxville, TN: ERIC Document Reproduction Service No. ED353334.

Tobias S (1987). Math anxiety. *Science* 237, 1556.

Zhao C, Kuh GD (2004). Adding value: learning communities and student engagement. *Res Higher Educ* 45(2), 115–138.