Statistics Across the Curriculum Using an Iterative, Interactive Approach in an Inquiry-Based Lab Sequence

By Alysa J. Remsburg, Michelle A. Harris, and Janet M. Batzli

How can science instructors prepare students for the statistics needed in authentic inquiry labs? We designed and assessed four instructional modules with the goals of increasing student confidence, appreciation, and performance in both experimental design and data analysis. Using extensions from a just-in-time teaching approach, we introduced concepts throughout a two-semester biology sequence when students could apply their knowledge to the inquiry labs. The modules included readings and active lessons that used modified "jigsaw" and immediate feedback techniques. Assessment based on pre- and postmodule surveys paired by individual student indicated improved student confidence and a rise in the number of students planning to take a statistics class. Student self-reported skill and performance were not always linked; student performance was best for the learning outcomes emphasized during class. Poor performance on low-level tasks, such as formatting statistical results, demonstrated the need for classroom time in support of reading assignments. The four short statistics modules appeared to be a minimum level of instruction for preparing science students to apply data analysis tools appropriately in their own research. Using an iterative process over several months, student-driven research and student-centered activities were important strategies in preparing students to apply statistics.

The laboratory experience is a foundation for science education (Hofstein & Lunetta, 2004). Students gain a deeper understanding of scientific concepts if their laboratory experiences are based on inquiry (e.g., American Association for the Advancement of Science, 2011; Leonard, 2000). A fundamental component of inquiry labs is analyzing and interpreting real data. Teaching statistics explicitly as part of labs improves student understanding of data analysis and interpretation (Maret & Ziemba, 1997). One challenge for instructors is that most students in introductory science courses have very little or no prior training in basic statistics. Even students who have taken statistics struggle to resolve statistical hypothesis testing with scientific hypothesis testing (Goldstein & Flynn, 2011; Maret & Ziemba, 1997). Thus a large gap emerges between the analytical tools available to students and their appropriate application in inquiry-based lab experiences. Science lab instructors struggle to prepare their students with at least functional tools for experimental design, data analysis, and data presentation because of time constraints and because these faculty have little to no training in teaching statistics. To address this gap, we developed and evaluated four statistics modules to accompany independent projects interspersed throughout two semesters of an honors introductory biology laboratory curriculum.

Math anxiety likely contributes to poorer student performance in data analysis than in other aspects of science courses (Hembree, 1990). Reviews of how undergraduate students learn statistics in psychology (Conners, McCown, & Roskos-Ewoldsen, 1998), wildlife ecology (Burger & Leopold, 2001; Kendall & Gould, 2002), and biology programs (Abrook & Weyers, 1996) indicate that many instructors have trouble motivating students to learn statistics. This can, in part, be attributed to students not appreciating the application of statistics in their chosen major (Conners et al., 1998; National Research Council, 2003). Pairing inquiry-based curriculum with a statistics-across-the-curriculum approach motivates students because they need to understand the tools of statistics while testing a hypothesis of their own creation.

Incorporating statistics routinely in science courses can improve student motivation for learning statistics and improve overall quantitative literacy (American Association for the Advancement of Science, 2011; National Council of Teachers of Mathematics, 2000; National Research Council, 2003). Familiarity with statistics early in the college curriculum also enables students to gain the most from subsequent coursework and research opportunities because they approach science in an objective, authentic way (Maret & Ziemba, 1997). "Ability to use quantitative reasoning" is considered a core competency by the American Association for the Advancement of Science (2011, p. 14) for undergraduate biology students. Therefore, introductory biology instructors need to incorporate aspects of data analysis when teaching experimental design. This quantitative literacy allows students to articulate hypotheses that can be directly evaluated on the basis of their current understanding of statistical approaches (Kendall & Gould, 2002). Statistics instruction within science courses should emphasize the role of statistics for minimizing background noise, defining data-collection protocols, explaining variability, and communicating results (Burger & Leopold, 2001; Demir, Schmidt, & Abell, 2010; Higgins, 1999; Kendall & Gould, 2002). How do instructors help students appreciate these powerful applications of statistics? Using a statistics-across-the-curriculum approach (Turner, 1981), we introduced components of experimental design and data analysis during 4 weeks of a two-semester biology lab curriculum so that students could apply knowledge of statistics to group research projects.

We observed previously that if "plug and chug" statistical tests are introduced before students fully understand summary statistics and experimental design, students do not question the output generated by the statistical software. Students also seem to assume that statistics consists of only memorizing a key phrase or rule after the computer does the work. Our first learning modules address this problem by focusing on elements of experimental design including sampling and replication. We subsequently help students become familiar with data they have generated, particularly their variation around the mean value and discerning meaningful patterns, before they apply statistical hypothesis testing. Knowledge, comprehension, application, and analysis from Bloom's taxonomy (Bloom, Krathwohl, & Masia, 1956) were all included in our learning outcomes.

The statistics instructional modules we present are focused on experimental design, process of science reasoning, and data analysis within the inquiry-based labs. As background reading and reference, we provided students with a statistics primer to supplement the lab time allocated to the statistics modules. The statistics primer served as a reference with explanations, examples, and instructions for statistics used during the two semesters. The lessons we describe here include active and group learning strategies to help diverse students practice the quantitative reasoning skills necessary to achieve our inquiry-based learning goals.

We used a scientific teaching approach (Handelsman et al., 2004) to evaluate this overall research question: To what extent do student confidence and application of statistics improve following exposure to statistics across the biology curriculum? Educational research goals were that students who experienced the statistics instructional modules would have (a) increased confidence in experimental design and data analysis, (b) increased appreciation of the importance of statistics in scientific research, and (c) improved performance in application of experimental design and data analysis concepts. Formative and final assessments helped us modify the modules in response to student needs. The specific student learning outcomes (Figure 1) address common misconceptions (based on instructors' prior experiences) rather than a set of comprehensive experimental design or data analysis concepts.

Methods

Classroom setting

We introduced 103 sophomore biology undergraduates in five laboratory sections to statistics in a two-semester honors lab sequence (2006–2007) at a large Midwestern Research 1 University (Batzli, 2005). A statistics course was not required as a prerequisite or corequisite for this introductory biology sequence, although a first-semester calculus course was a prerequisite. Institutional Review Board approval was granted for this education research. The two Bio I modules focused on experimental design, describing variation, and graphical representation of data. The two Bio II modules integrated and expanded statistics concepts taught in Bio I and added hypothesis testing. We dispersed the statistics modules throughout the two lab courses (Table 1), but these teaching materials could be applied to 4 weeks of most single-semester science laboratory courses.

Sequencing of statistics modules across the curriculum

The study system for the first statistics module was water quality of a local stream. Students were assigned readings in their statistics primer (excerpt in Appendix 1; available at www.nsta.org/college/connections.aspx) about experimental design and summary statistics to help guide their laboratory assignments. Students completed a prelab assignment that

FIGURE 1

Selected learning outcomes that were the focus of statistics modules for an honors introductory biology two-semester lab series. These outcomes formed the basis for assessments of student attitudes, confidence, and performance.

I. Experimental Design
   A. Write research hypotheses that explicitly state the variables measured.
   B. Assign experimental treatments systematically when a known factor other than the independent variable influences the dependent variable.
   C. Consider how well an experiment tests the direct effects of an independent variable.
   D. Design experimental replicates that receive equivalent but independently applied experimental conditions.
   E. Understand how sample size affects the ability to detect a significant effect when there is variation in the system (signal-to-noise ratio).

II. Data Analysis
   A. Recognize when paired *t*-tests, independent *t*-tests, and ANOVAs are useful.
   B. Conduct independent *t*-tests using Excel.
   C. Interpret *p*-values to make appropriate conclusions regarding hypotheses.
   D. Make conclusions based on the probability that data support statistical hypotheses.
   E. Follow the format of a scientific paper: statistics included in Methods section; test statistic, degrees of freedom, *p*-values and appropriate graphs reported in Results section.

included practice questions where they applied concepts from the reading to a set of hypothetical data. Students gathered stream data the following week, sharing data across all five lab sections in a common Excel spreadsheet. Following field work, we began statistics instruction (Module 1) with a discussion of summary statistics relevant to the data collected. In particular, students practiced articulating hypotheses, differentiating true replicates from nonindependent pseudoreplicates, analyzing raw data and distributions in histograms, summarizing variation in data using descriptive statistics, and determining meaningful graphical representations of the data (e.g., bar graphs with error bars, scatterplots) in Microsoft Excel using data they collected in the stream.

Modules 2 and 3 consisted of group research projects in which students developed questions and tested their own hypotheses. Students brainstormed biological questions, performed primary literature searches, presented proposals, responded to oral feedback on their research design, conducted the experiments, analyzed data, and summarized their research in a poster and a paper. Table 1 outlines the sequence of activities and assessments.

The experimental design module (Module 2) began by presenting students with a number of realistic, flawed experiments for critique. We used research examples from a variety of disciplines to engage students with diverse interests. For example, students discussed whether *Daphnia* (aquatic water flea, zooplankton) living and developing in the same beaker should be treated as independent experimental replicates. Students were asked to summarize and analyze results of their *Daphnia* research projects, but data analyses of this first project were limited to comparisons of means, standard deviations, standard error, and 95% confidence intervals. In other words, we introduced the comparison of means in light of variation among replicates but saved formal hypothesis testing for the next statistics module (in Bio II). We used nongraded research proposal presentations and the ensuing questions as a formative assessment of how well students could apply experimental design concepts to their *Daphnia* projects. Students reported they planned to use statistics, but their proposals lacked statistical terminology, suggesting their awareness of specific tools was still vague.

Module 3 in Bio II enabled students to use their understanding of experimental design to develop enzyme catalysis research hypotheses (Table 1). We asked students to refer to their statistics primer to complete a prelab assignment (Appendix 2; available at www.nsta.org/college/connections.aspx) prior to beginning their experiments and as a reference to help guide the design of their experiments (e.g., defining replicates, sampling) and the analysis of their own data.

Module 4, focused on hypothesis testing, included a set of immediate feedback questions to assess informally how well students understood when independent and paired *t*-tests would be appropriate for different research scenarios (Appendix 3; available at www.nsta.org/college/connections.aspx). We used a low-tech sticker approach that helped prepare students to discuss the answers verbally, in the same way that "clicker questions" are useful for large lectures (Levesque, 2011). This activity also enabled formative assessment by both students and instructors.

TABLE 1

Teaching framework for four statistics modules (four classroom sessions) embedded in a two-semester biology laboratory series.

Time	Topic	Activity/ Assessment	Purpose
Homework before Module 1	Summary statistics using field data	Reading assignment[a] Graded reading questions	Understand summary statistics outcomes (Figure 1) Elicit prior knowledge and misconceptions
Module 1 40 min	Summary statistics and Excel	Discuss reading question answers Outline the results section of a scientific paper based on data	Extend understanding of summary statistics Explore basic purpose of statistics; generate descriptive statistics and graphs with error bars
Module 1 80 min	Introduction of experimental system	Develop research proposal	Engage in statistics by developing a hypothesis, designing an experiment, and proposing analysis of expected results
Homework before Module 2	Experimental design and basic data analysis	Online premodule survey[b]	Instructors assess prior knowledge
Module 2 30 min	Experimental design practice	Work in groups to identify experimental design flaws in research scenarios	Reveal common misconceptions in experimental design
Module 2 30 min	Experimental design application	Work in groups to apply experimental design tips from worksheet to research proposals	Apply experimental design concepts to group research projects
Module 2 50 min	Formative evaluation	Group presentations of research proposals	Instructors assess experimental design and use just-in-time teaching to improve projects
Module 3 2 hr	Introduce experimental system	Record data; brainstorm hypotheses for independent projects	Appreciate the role of statistics in a research project
Homework before Module 4	Experimental design and data analysis Formative evaluation	Reading assignment[a] Graded reading questions[c]	Extend experimental design concepts to new scenarios Elicit prior knowledge and misconceptions
Module 4 30 min	Role of statistics, independent and paired *t*-tests	Discuss reading questions[c] and extensions in groups and as a class	Explain formal hypothesis testing to peers Consider when and why to use paired two-sample hypothesis testing
Module 4 10 min	Paired versus independent *t*-tests	Immediate feedback questions posted around room[d]	Apply understanding of when paired *t*-tests can be useful
Module 4 40 min	*t*-test mechanics, introduction to ANOVAs	Chalk talk, group jigsaw questions,[e] and group presentations	Explore and explain data analysis concepts (Figure 1)
After Module 4	Conduct group research projects	Group research paper assignment	Apply data analysis Communicate results
After Module 4	Experimental design and data analysis	Online postmodule survey[b]	Instructors assess learning gains

[a] Statistics primer excerpt, Appendix 1
[b] Pre- and postmodule surveys, Appendix 4
[c] Experimental design reading questions, Appendix 2
[d] Immediate feedback questions, Appendix 3
[e] Group jigsaw activity for *t*-tests, Figure 2

Module 4 also attempted to engage students in a more thorough consideration of *t* scores by assigning specific questions to small groups (Figure 2). Because each group presented on a different topic, students learned pieces of the whole story from each other. This is a simple extension of the "jigsaw" cooperative teaching approach (Johnson, Johnson, & Smith, 1991). Students discovered and discussed the answers by analyzing either graphs or datasets in Excel and then presented their findings to the class. Use of a software package already familiar to many students enabled them to focus on statistical concepts rather than software or computations (Garfield, Hogg, Schau, & Whittinghill, 2002).

Assessment of student confidence

Self-reported confidence was assessed using 10 survey questions that asked students to rate their own skills in experimental design and data analysis (Appendix 4; available at www.nsta.org/college/connections.aspx). The 10 questions addressed each of our student learning outcomes (Figure 1). Students took this survey before and after completing the statistics modules (Table 1). In October 2006, 103 Bio I students took the online survey; 82 Bio II students repeated these questions in May 2007 (ungraded both times, but with participation points for completion). We used Bonferroni-corrected permutation tests, paired by student to assess changes in self-reported confidence between the two online surveys (n = 78, after eliminating students who did not complete one of the two surveys). Permutation tests were appropriate for the data that lacked a normal distribution; Bonferroni corrections conservatively control for errors from multiple comparisons. These and our other research questions were tested using R software (Version 2.4.1 R Core Development Team, Vienna, Austria).

FIGURE 2

Laboratory activity to explore *t*-test concepts using jigsaw-style group investigations.

Instructions: Your group will be responsible for considering the answer to each question and then teaching the whole class, since each group has a different question. Within your group, you should have a facilitator, recorder, and two presenters. Take about 5 minutes to figure out your answers.

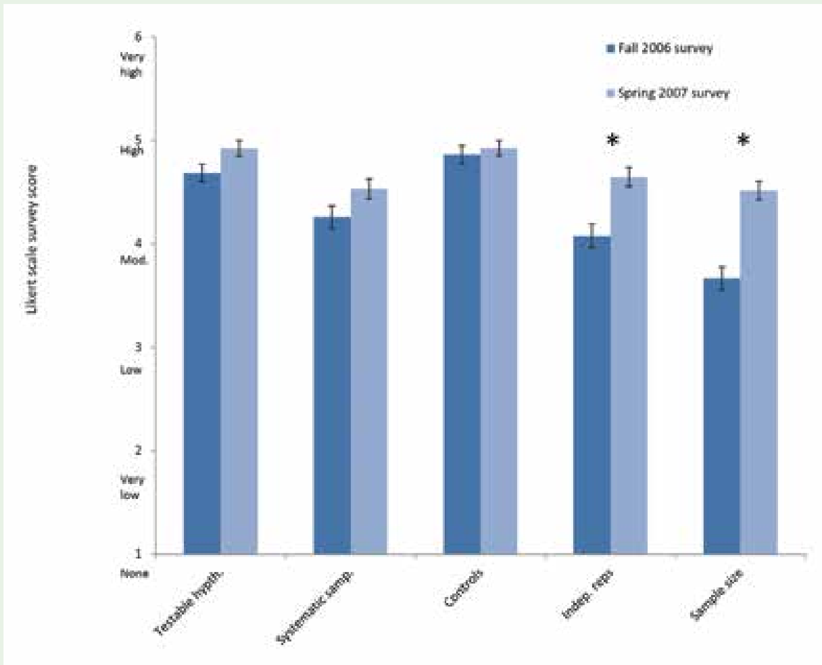
1. Label t = –4.56 (from prelab Question 8) on the *t* distribution graph and indicate where the corresponding *p*-value occurs on the distribution. Indicate where a *t*-score that has an associated *p*-value of 0.05 would fall on this graph.
2. What happens to the *t*-score (and associated *p*-value) as the sample size increases? Use Excel to run *t*-tests on each of these data sets:
   a) the tab called "means," which is the lab data that you used on your prelab assignment, and
   b) the tab called "Group 2," which includes the same lab data plus extra replicates that gave similar results.
3. What happens to the *t*-score (and associated *p*-value) as the sample variance decreases? Use Excel to run *t*-tests on each of these data sets:
   a) the tab called "means," which is the lab data that you used on your prelab assignment, and
   b) the tab called "Group 3," which includes similar data, but the replicates are more similar to each other (lower variance) within the same temperature group.
4. Use the figure to explain why the degrees of freedom are important in calculating a *p*-value.
5. Use Excel (and instructions on page 11 of your statistics primer) to run an independent two-sample *t*-test on our lab data. Write a sentence with the results, indicating *df* and two-tailed test.

Assessment of statistics appreciation

To evaluate whether student appreciation of statistics changed between the pre- and postmodule surveys (n = 47 students without prior statistical coursework who completed this question both times), we compared responses to the survey question, "Do you anticipate taking statistics courses in the future?" We assumed that those students who changed their plans to take more statistics after our lab courses had developed an appreciation for the importance and scope of statistics. We never discussed specific statistics courses with students during the biology labs. We used McNemar's tests (Argesti, 2002), paired by student, to compare the number of students whose plans changed from "no" to "yes" with the number of students whose plans changed from "yes" to "no." McNemar's test is a contingency table test for nominal data where the null hypothesis is that responses did not change from the first to second survey.

FIGURE 3

Comparison of mean (± 1 SE) pre- and postinstruction self-reported skill at determining: testable hypotheses, when systematic sampling is helpful; appropriate experimental controls, whether experimental replicates are independent; and appropriate sample sizes (Figure 1). Asterisks (*) indicate significant differences based on permutation tests paired by individual student (*n* = 78; *p* < 0.005).



Assessment of student performance

Two open-ended questions on the pre- and post-module surveys (Appendix 4) assessed how well students could apply three of our learning outcomes (Figure 1) to novel scenarios: when to use systematic versus random sampling, how indirect effects from uncontrolled variables can confound conclusions, and when to apply *t*-tests or analyses of variance (ANOVAs). We emphasized that students must complete the surveys individually outside of class. The second survey presented a different scenario from the first survey, and correct survey answers had not been discussed in class at any time. We assessed completion of outcomes using a simple binary scoring rubric. We used McNemar's tests, paired by student, to evaluate the number of students who met learning outcomes on the final survey in comparison with the first survey. Bonferroni corrections for conducting three separate tests led to a significance level of $p < 0.01$.

We assessed how many students correctly applied the stated data analysis learning outcomes in their research papers at the end of the four statistics modules (May 2007). Teaching assistants selected 25 group research papers randomly. Two of the authors (AJR and MAH) used a binary scoring rubric to record whether students achieved each learning outcome in these papers.

Results

Student confidence

After completing the two lab courses, students reported significantly greater confidence in two of the five experimental design learning outcomes (Figure 1): distinguishing independent replicates and determining appropriate sample sizes (Figure 3; $n = 78$). Of the eight data analysis skills included in the survey, students reported significantly higher confidence in six of the skills at the end of the lab sequence: determining when to use independent *t*-tests, paired *t*-tests, and ANOVAs; conducting independent *t*-tests; interpreting statistical tests; and stating conclusions of experimental results (Figure 4; $n = 78$). Other specific outcomes showed no change in self-reported student confidence.

Statistics appreciation

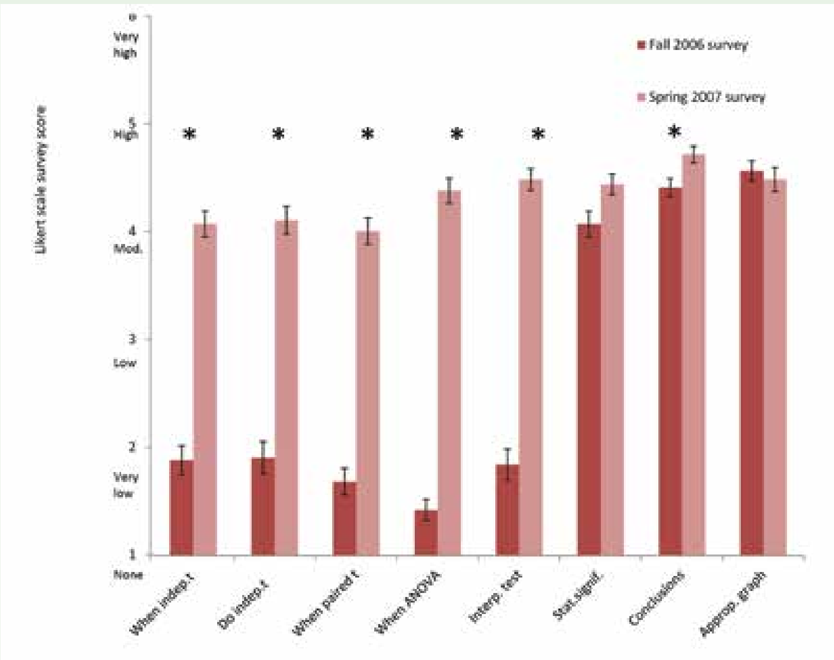
We saw evidence of increased student appreciation for the importance of statistics. Among students with no prior statistics coursework ($n = 47$), significantly more students (10) noted plans to take a statistics course after exposure to our statistics modules than those who no longer planned to take a statistics course after our modules (one student; McNemar's $\chi^2 = 7.4$, $df = 1$, $p = 0.007$). Among students who had already taken a statistics course (at any institution, $n = 30$), we found no change in their plans to take a statistics course.

Student performance

Student performance improved from 2% to 48% completion for one of the three learning outcomes tested on the pre- and postmodule surveys: identifying a relevant statistical test ($n = 65$; McNemar's $\chi^2 = 31$, $p < 0.0001$). Per-

FIGURE 4

Comparison of mean (± 1 SE) pre- and postinstruction self-reported skill at conducting independent and paired *t*-tests, conducting ANOVAs, interpreting statistical results, determining statistical significance, drawing conclusions, and selecting an appropriate graph to present data (Figure 1). Asterisks (*) indicate significant differences based on permutation tests paired by individual student (*n* = 78; *p* < 0.005).



centages of students correctly identifying systematic sampling and experimental complications from indirect effects both increased between the pre- and postmodule surveys (from 5% to 17% and from 21% to 25%, respectively), although individual learning gains were not significant.

On final research papers, most students correctly applied ANOVA tests, reached correct conclusions based on the data analysis, and drew appropriate biological conclusions from the data (Table 2). For data reporting procedures that we did not emphasize in class but assigned as reading in the statistics primer, students did not perform as well. This was despite the fact that the data reporting tasks required lower order thinking according to Bloom's taxonomy of educational objectives. Practice during class seemed to have helped students succeed at the higher order learning outcomes assessed in their papers: selecting statistical tests and making conclusions based on statistics.

Finally, anecdotal observations by the instructors of these lab courses contributed information about students' progress during the year. Students seemed to develop an ability to ask testable and sophisticated research questions as they began to understand more sophisticated experimental design and data analysis. Unfortunately, we also observed that students in research groups often split up aspects of the research so that not all students practiced all aspects of experimental design and analysis. Thus, individual-based assessment and accountability is important.

Discussion

Our three education research goals for the statistics modules were achieved: improving student confidence, appreciation of statistics, and performance on at least some of the learning outcomes. We documented greater appreciation of statistics on the basis of significant changes in students' plans to take a statistics course. According to pre- and postmodule surveys to assess student confidence, self-reported skills in two of five experimental design outcomes and six of eight data analysis outcomes improved significantly. The absence of significantly improved student confidence for two data analysis outcomes and three experimental design outcomes was associated with a perception of "moderate" to "high" skill levels in these outcomes even before the statistics learning modules. Most of these honors biology students apparently thought they already had sufficient experience with testable hypotheses, systematic sampling, statistical significance, and graph selection from high school science classes.

Self-reported skills do not always match assessed performance of those skills. Our final open-ended questions to assess student performance indicated that only 17% of students achieved the learning outcome of recognizing the need for systematic sampling, and only 25% of students achieved the outcome of controlling for indirect effects in experimental design. Nonetheless, their pre- and postmodule surveys indicated high confidence in these experimental design skills. When students assessed their own skills, they missed the full scope of the experimental design process. Performance scores were likely low because the open-ended scenario questions did not provide clues about relevant concepts, and they required higher level thinking (Crowe, Dirks, & Wenderoth, 2008). However, formal survey question validation may have improved the clarity of the assessment questions. We did not grade students on their answers to these questions because we wanted to avoid cheating temptations, but it is possible that students may not have been motivated to do their best work.

Student performance in recognizing statistically significant trends in data sets showed improvement, whereas self-reported surveys suggested no change due to moderately high initial confidence for this skill. Confronting a scenario in which they had to apply the concept was much more challenging than students anticipated when summarizing their own skills. Only 2% answered this question correctly on the premodules survey, whereas 48% answered correctly on the postmodule survey.

Two major factors likely account for the more modest improvement in experimental design outcomes compared with larger improvements in data analysis skills documented here. First, students have had exposure to some aspects of experimental design and graph selection since their science education in primary grades. This may lead to higher confidence and greater challenges with misconceptions. Second, experimental design includes less prescribed concepts than specific statistical tests and requires higher level critical thinking. For instance, experimental design for an ecological field-based experiment shares some basic principles with experimental design for an experiment in cellular biology, but there are also some context-specific aspects that are unique and require creative, critical thinking to diagnose potential confounding variables and aspects of time, development, and the environment. Critical thinking is a notoriously difficult skill for students to achieve because students are less able to appreciate their deficiencies in this nonprescriptive process than they are to appreciate deficiencies in specific knowledge (Halpern, 2002). Thus pretests to identify misconceptions about experimental design can be very useful in engaging students to relearn this process in progressively more sophisticated contexts while they are developing their own questions and experiments. Because we wanted to use the survey questions for a postteaching assessment, we did not return the initial survey to students with grades and comments. This likely would have helped students identify and correct their own weaknesses prior to beginning the subsequent experimental design modules.

Integrating statistics instruction in multiple contexts across the curriculum proved manageable and effective. Sequencing and lag times between our modules allowed students to focus first on variations in data as they developed experimental design skills, followed by hypothesis testing. This provided a sequence of learning so that students could build on prior skills and appreciate a need for more statistical training. The iterative nature of statistics instruction—through lab activities, homework assignments, and statistics readings—helped guide students through their inquiry-based labs. Focusing on experimental design

TABLE 2

Assessment of data analysis learning outcomes in final student research papers.

Criteria for research paper	Sample size[a]	Proportion correct	Bloom's level[b]
Recognize when ANOVA is useful	21	0.810	Application
Appropriate graph with variation around mean	25	0.880	Application
Reporting of ANOVA includes F, df, p	17	0.353	Knowledge
Reporting of t-test includes t, df, 1- or 2-tailed, p	18	0.333	Knowledge
Correct conclusion from p-value	25	0.920	Comprehension
Logical biological conclusion from data	25	0.840	Analysis

[a]Sample sizes were smaller for some of the outcomes recorded because some statistical tests were not applicable to all student projects, depending on the nature of the student hypothesis tested.

[b]Bloom's taxonomy of educational objectives (Bloom et al., 1956).

and data analysis during 4 weeks of the two-semester sequence appeared to be a minimum time commitment to help students achieve most of our learning outcomes. Homework assignments and jigsaw-style classroom instruction helped us reduce class time spent on statistics, which was desirable because of so much biology content deemed essential and so little student patience for "math"-related instruction. Perhaps most important, we also observed that students were motivated to learn statistical tools because of the immediate and relevant application to their own independent research data sets.

Improvements to science curriculum ideally include a system of coordinated prerequisite or corequisite statistics courses (Marsteller et al., 2010), but additional courses may not lead to students correctly applying their knowledge to inquiry science labs. Our results suggest that giving students assigned readings and the opportunity to apply statistics during science labs leads to less satisfactory learning gains than adding active learning statistics modules as part of the lab, similar to findings by Sirum and Humburg (2011). In the inquiry-based labs, where students developed and tested their own hypotheses, students were motivated to make careful decisions about replicates, sampling design, and relevant statistical tests.

As instructors, we also found that implementing this authentic process of science required extra time for teaching students how to use Microsoft Excel software for graphing and data analysis. A drop-in consulting session as students were analyzing their own data worked well. Tools online have made it easier for students to conduct statistical tests, but arranging and graphing data still require skills in Excel. Frequent updates to the software add to this learning curve, which some students respond to more readily than others. Excel applications could be added more explicitly to prelab assignments and learning modules.

Conclusion and implications

Successes and challenges from these statistics modules can inform science instructors who commonly struggle with the need for statistics support in their labs. Further research on effectiveness of the teaching should include more qualitative data. We observed anecdotally that motivation to learn statistics was enhanced through the timing of four modules presented when students needed them for the inquiry labs; documenting this change in motivation level would be an interesting avenue for future studies. Teaching concepts of variation among replicates prior to hypothesis testing helped students develop quantitative reasoning prior to application of statistical tools. Active lessons were useful in engaging and revealing misconceptions. The straightforward learning outcomes that relied exclusively on readings and homework were not met nearly as well as learning outcomes supported with face-to-face lab instruction. Students demonstrated relatively high confidence in their experimental design abilities but much room for improvement in some design considerations that require critical thinking such as systematic sampling. Specific learning outcomes that were the focus of these modules were designed to address common challenges observed in inquiry labs, although Excel software skills should be added to this list of common challenges. A limited number of outcomes enabled us to present the statistics modules during four short lessons across the two courses. Devoting more science time to statistics and coordinating across more courses is always challenging but beneficial. ■

References

Abrook, R., & Weyers, J. D. B. (1996). Teaching of statistics to UK undergraduate biology students in 1995. *Journal of Biological Education, 30,* 281–288.

American Association for the Advancement of Science. (2011). *Vision and change in undergraduate biology education: A call to action*. Available at http://visionandchange.org/files/2013/11/aaas-VISchange-web1113.pdf

Argesti, A. (2002). *Categorical data analysis.* New York, NY: Wiley.

Batzli, J. M. (2005). A unique approach? Four semesters of biology core curriculum. *CBE—Life Sciences Education, 4,* 125–128.

Bloom, B. S., Krathwohl, D. R., & Masia, B. B. (1956). *Taxonomy of educational objectives: The classification of educational goals.* New York, NY: David McKay.

Burger, L. W., & Leopold, B. D. (2001). Integrating mathematics and statistics into undergraduate wildlife programs. *Wildlife Society Bulletin, 29,* 1024–1030.

Conners, F. A., McCown, S. M., & Roskos-Ewoldsen, B. (1998). Unique challenges in teaching undergraduate statistics. *Teaching of Psychology, 25,* 40–42.

Crowe, A., Dirks, C., & Wenderoth, M. P. (2008). Biology in bloom: Implementing Bloom's taxonomy to enhance student learning in biology. *CBE—Life Sciences Education, 7,* 368–381.

Demir, A., Schmidt, F., & Abell, S. K. (2010). Science from the pond up: Using measurement to introduce inquiry. *Journal of College Science*

Teaching, 39(4), 23–27.
Garfield, J., Hogg, B., Schau, C., & Whittinghill, D. (2002). First courses in statistical science: The status of educational reform efforts. *Journal of Statistics Education, 10,* 456–467.
Goldstein, J., & Flynn, D. F. (2011). Integrating active learning and quantitative skills into undergraduate introductory biology curricula. *The American Biology Teacher, 73,* 454–461.
Halpern, D. (2002). *Thought and knowledge: An introduction to critical thinking*. Mahwah, NJ: Erlbaum.
Handelsman, J., Ebert-May, D., Beichner, R., Bruns, P., Chang, A., DeHaan, R., . . . Tilghman, S. M. (2004). Scientific teaching. *Science, 304*(5670), 521–522.
Hembree, R. (1990). The nature, effects, and relief of math anxiety. *Journal for Research in Mathematics Education, 21,* 33–46.
Higgins, J. J. (1999). Nonmathematical statistics: A new direction for the undergraduate discipline. *American Statistician, 53,* 1–6.
Hofstein, A., & Lunetta, V. N. (2004). The laboratory in science education: Foundations for the twenty-first century. *Science Education, 88,* 28–54.
Johnson, D., Johnson, R., & Smith, K. (1991). Jigsaw procedure. In *Active learning: Cooperation in the college classroom* (pp. 4–17). Edina, MN: Interaction Book Company.
Kendall, W. L., & Gould, W. R. (2002). An appeal to undergraduate wildlife programs: Send scientists to learn statistics. *Wildlife Society Bulletin, 30,* 623–627.
Leonard, W. (2000). How do college students best learn science? *Journal of College Science Teaching, 29,* 385–388.
Levesque, A. A. (2011). Using clickers to facilitate development of problem-solving skills. *CBE—Life Sciences Education, 10,* 406–417.
Maret, T. J., & Ziemba, R. E. (1997). Statistics and hypothesis testing in biology. *Journal of College Science Teaching, 26,* 283–285.
Marsteller, P., de Pillis, L., Findley, A., Joplin, K., Pelesko, J., Nelson, K., . . . Watkins, J. (2010). Toward integration: From quantitative biology to mathbio-biomath? *CBE—Life Sciences Education, 9,* 165–171.
National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics.* Reston, VA: Author.
National Research Council. (2003). *BIO 2010: Transforming undergraduate education for future research biologists.* Washington, DC: National Academies Press.
Sirum, K., & Humburg, J. (2011). The Experimental Design Ability Test (EDAT). *Bioscene: Journal of College Biology Teaching, 37,* 8–16.
Turner, D. E. (1981). Statistics across the curriculum. *Teaching Statistics, 3,* 5–7.

Alysa J. Remsburg (aremsburg@uwlax.edu) *is an Associate Lecturer in the Environmental Studies Program and River Studies Center at University of Wisconsin–La Crosse.* ***Michelle A. Harris*** *is a Faculty Associate and* ***Janet M. Batzli*** *is the Associate Director, both in the Biology Core Curriculum (Biocore) Program at the University of Wisconsin–Madison.*