

## Article

# Assessment of the Effects of Student Response Systems on Student Learning and Attitudes over a Broad Range of Biology Courses

Ralph W. Preszler, Angus Dawe, Charles B. Shuster, and Michèle Shuster

Department of Biology, New Mexico State University, Las Cruces, NM 88003-8001

Submitted September 14, 2006; Revised December 18, 2006; Accepted December 20, 2006  
Monitoring Editor: Marshall Sundberg

With the advent of wireless technology, new tools are available that are intended to enhance students' learning and attitudes. To assess the effectiveness of wireless student response systems in the biology curriculum at New Mexico State University, a combined study of student attitudes and performance was undertaken. A survey of students in six biology courses showed that strong majorities of students had favorable overall impressions of the use of student response systems and also thought that the technology improved their interest in the course, attendance, and understanding of course content. Students in lower-division courses had more strongly positive overall impressions than did students in upper-division courses. To assess the effects of the response systems on student learning, the number of in-class questions was varied within each course throughout the semester. Students' performance was compared on exam questions derived from lectures with low, medium, or high numbers of in-class questions. Increased use of the response systems in lecture had a positive influence on students' performance on exam questions across all six biology courses. Students not only have favorable opinions about the use of student response systems, increased use of these systems increases student learning.

## INTRODUCTION

Many efforts to improve science education involve enhancing the social context of learning. Traditional lectures and textbooks primarily rely on one-way communication from instructors and authors to students. This approach assumes that students are able to assimilate the information and integrate it into their existing understanding of a topic and are also able to adjust their understanding of related concepts. It also assumes that individual students are able to resolve conflicts between the information presented in the course and their preexisting understanding and beliefs. This extensive integration of knowledge within appropriate cognitive structures, the process of meaningful learning, is necessary to develop the ability to apply understanding of scientific knowledge to unique problems. Few students are able to accomplish this in the traditional structure of university science courses (National Research Council [NRC], 2000; Novak, 2002). As a result, universities are producing grad-

uates who are unable to apply a meaningful understanding of science to the solution of personal and societal issues (Boyer Commission on Educating Undergraduates, 1998). Our failure to help students learn to integrate scientific knowledge across cognitive domains also may account for the lack of graduates who are able to contribute to the information-rich, interdisciplinary process of modern biological research (NRC, 2003). Efforts to address this problem through increasing the frequency and quality of interactions within university science courses include enhancing student-to-instructor communications (Paschal, 2002), expanding instructional teams to include peer instructors (Tessier, 2004; Knight and Wood, 2005; Smith *et al.*, 2005), and promoting cooperative learning activities among students (Rao and DiCarlo, 2000). Implementation of these reforms ranges from the addition of interactive activities to traditional courses (Rao and DiCarlo, 2000; for review of methods, see Allen and Tanner, 2005) to the complete restructuring of courses (Klionsky, 2001; Udovic *et al.*, 2002; Roy, 2003; Knight and Wood, 2005). These efforts share a set of related objectives: developing more engaging courses that increase students' interest in and appreciation of science, helping

DOI: 10.1187/cbe.06-09-0190

Address correspondence to: Ralph W. Preszler (rpszle@nmsu.edu).

students learn to work more effectively in problem-solving teams, and improving students' understanding of scientific knowledge.

Educational technology provides many opportunities for enhancing communication and increasing interactivity within science courses. It can be used to support individualized learning activities, thereby opening up time in lecture courses for interactive sessions. For example, electronic student response systems have been used to quiz students at the beginning of lectures to encourage students to carefully read the textbook so that the basic information in the text does not need to be repeated in lecture (Paschal, 2002; Knight and Wood, 2005). However, the most exciting potential of educational technology is to facilitate interactions between students and instructors and among students (Jensen *et al.*, 2002; Novak, 2002; Wood, 2004). In this study, we have investigated wireless student response systems that may enhance student-to-instructor interactions by providing instructors with immediate formative assessments of students' understanding (Paschal, 2002). Additionally, these systems may facilitate cooperative learning among students (Dufresne *et al.*, 1996).

In a thorough review of student response systems, including current wireless systems (clickers) as well as earlier wired systems, Judson and Sawada (2002) concluded that three decades of studies have consistently documented positive student evaluations of the use of student response systems in lecture courses. They found that when used in the stimulus-response, operant conditioning, approach to learning prevalent in the 1960s and 1970s, students claimed student response systems increased their level of interest and their learning; however, this increased interest was not reflected in measures of student performance. When coupled with constructivist, student-centered learning strategies in physics, the use of wireless student response systems (clickers) has been associated with measurable conceptual gains. Judson and Sawada (2002) also concluded that it was simply too early to tell whether clickers would enhance learning in other sciences, but they recommended that this technology should be used to promote student-centered learning. Since the review of Judson and Sawada (2002), the use of clickers to promote learning in biology has produced mixed results. Knight and Wood (2005) increased student learning by using peer instructors and clickers to facilitate cooperative learning in a developmental biology course. Studies of the use of clickers in microbiology (Suchman *et al.*, 2006) and physiology (Paschal, 2002) describe some encouraging trends, but they did not find clearly significant effects of clickers on students' exam performance. Currently, there is strong evidence that students have positive opinions about the use of clickers in biology courses; however, the effects of clickers on student learning remain unclear.

In this study, we have explored student opinions regarding the use of clickers and the effects of the frequency of clicker use on student performance. A challenge of using student-centered teaching methods, such as clickers, is that the application of these methods should be centered on and responsive to the learning environment of each group of students. Each method needs to be compatible with the students' approach to learning, the instructor's teaching style, and the course content. We investigated the robustness of using clickers by evaluating their effectiveness across six

biology courses ranging from freshman- to senior-level classes. We also explored how students' opinions about this technology vary in association with their course grades and with their preference for studying individually or cooperatively. In addition to investigating students' opinions about the use of clickers in biology courses, we investigated the impact of clickers on students' exam performance. This impact was tested by systematically varying the frequency of clicker questions among lectures within six biology courses and comparing this variation in clicker question frequency to students' performance on exam questions from lectures associated with different numbers of clicker questions. This comparison allowed us to determine whether the frequency of clicker use within a course influenced students' performance on exams as well as whether this influence was consistent across courses.

Student response systems enable each student with a uniquely registered clicker to answer questions posed by the instructor by sending his or her response to the receiver attached to the instructor's computer. This approach engages each student as he or she commits to the answer and has the potential to provide a dramatic improvement over the traditional approach of the instructor orally posing a question and waiting while most students passively watch the few who are willing to respond. The instructor can immediately display a histogram of the class responses while maintaining the anonymity of each individual's answer. Projecting the histogram of class responses allows students to gain a better understanding of their class, and it provides the instructor with an immediate measure of students' understanding of a topic. The instructor can use this rapid formative assessment to move on if most students understand; if students are confused, the class can break into a discussion or a review. When clickers are used to stimulate discussion and cooperative problem solving, they may enhance meaningful learning in lecture courses of any size (Judson and Sawada, 2002). We often allowed students to discuss the problem before transmitting their individual answers to clicker questions; we also used the histogram of students' responses as a catalyst to trigger discussion. Therefore, our lectures that included more clicker questions also incorporated more interactions and cooperative learning among students. We have evaluated the hypothesis that students learn more effectively during these higher clicker frequency, more interactive lectures. Specifically, we tested the prediction that average student scores on exam questions will increase with low, medium, to high frequencies of clicker use in corresponding lectures.

## METHODS

In fall 2005, six courses in the Department of Biology at New Mexico State University (NMSU) used wireless student response units. The number of clicker questions posed was systematically varied within courses. This variation in clicker questions was subsequently compared with students' performance on exam questions tied to low-, medium-, or high-clicker lectures. At the end of the semester, students' opinions about the use of these clickers were surveyed. The surveys were analyzed to determine students' impressions of the use of clickers in biology courses and to explore how these opinions varied among courses, among students who preferred to study individually for exams in contrast to those who studied cooperatively, and to evaluate variation in the opinions of students based on

their course grades. Our protocol for using human subjects in this research was approved by the NMSU Institutional Review Board (NMSU IRB 6136).

### Biology Courses and Student Response Units

The six biology courses participating in this study included four lower-division courses (Biology 101, 111, 211, and 219) and two upper-division courses (Biology 311 and 377). Table 1 describes the demographic characteristics of each course. Biology 101 is a non-majors introductory biology course that uses an issues-based approach to help students develop an understanding of the application of scientific knowledge to personal and societal issues. Although this is a freshman-level course, many of the students in Biology 101 are not freshmen. Biology 111 and 211 are introductory biology courses that include science majors as well as students with a professional need for training in biology, such as the prenursing students in Biology 211. Biology 219 is an introductory microbiology course specifically developed to meet the needs of allied health (prenursing and related fields) majors. Biology 311 is an upper-division general microbiology course targeting, but not limited to, science majors. Last, Biology 377 is a cell biology course composed primarily of senior biology, microbiology, and biochemistry majors. Biology 101, 211, and 219 were taught by the same instructor. This instructor had used clickers in three courses before this study. Biology 111 and Biology 311 were taught by instructors who were using clickers for the first time. The instructor of Biology 377 had used clickers in one course before this study.

All courses used the most current radio frequency Classroom Performance System available from eInstruction (Denton, TX) (<http://www.einstruction.com/>) in fall 2005. Students purchased the remote handsets at the university bookstore, except for Biology 111 students who were given the handsets. Students in all courses registered their wireless remotes at the eInstruction website and paid a registration fee at the same website. Instructors downloaded eInstruction software and student registration information onto the laptop computers that they used in lecture. Two of the six courses (Biology 311 and 377) used Macintosh computers in lecture. Instructors presented multiple-choice questions in lecture on PowerPoint slides, often allowing students to discuss the question before transmitting their answers. Receivers connected to the instructors' laptops recorded each student's response. Immediately after the question, instructors presented a histogram of the distribution of students' responses. Instructors adjusted their lectures on-the-fly to students' comprehension of each concept explored with a clicker question. If many students did not answer the question correctly, the instructors moved to a discussion or review; if students generally understood the material, instructors moved on to the next topic.

The eInstruction software recorded each student's response to each question. After class, instructors set each correct answer at 100% of the possible points, each incorrect answer at 80%, and unanswered questions at 0%. Although all courses were using the wireless response units by the second week of classes, students did not receive credit for their responses until near the time of the first exam. This delay in assigning credit to clicker responses was made to provide students with sufficient time to purchase and register their clickers. The proportion of total course points associated with

**Table 1.** Course descriptions

Student category	Course (Biology)					
	101	111	211	219	311	377
Demographic						
Female (%)	75	65	71	85	57	67
Ethnic minority (%)	61	52	56	44	49	63
Class size						
Originally enrolled	236	147	127	56	91	64
Completed exams	206	170	115	50	86	58
Completed survey	130	140	115	50	86	58
Completed survey	130	152	68	40	71	50
Completed survey		98				
Class (% by course)						
Freshman	26	48	7	0	0	0
Sophomore	39	31	47	29	4	0
Junior	22	12	30	45	30	10
Senior	13	8	15	25	65	90
Discipline (% by course)						
Agriculture	3	16	15	0	19	10
Arts & Letters	10	0	2	0	0	0
Biochemistry & Molecular Biology	0	3	7	0	19	49
Biology	0	14	14	0	29	36
DABCC	1	6	11	0	1	0
Education	24	7	2	2	8	0
Health & Social Services	11	6	30	98	14	0
Social Sciences	23	5	0	0	3	2
Other	6	7	5	0	6	2
Undeclared	22	36	15	0	1	2

Gender, class size, ethnicity (percentage of underrepresented ethnic minority), class, and discipline are listed for each course. Two class sizes of each type are listed for Biology 111 because it was taught in two sections. Discipline categories includes some individual colleges (Agriculture, Dona Anna Branch Community College [DABCC], Education, and Health and Social Services), groups of related departments within the College of Arts and Sciences (Arts & Letters includes arts as well as various English and communications majors; Biochemistry and Molecular Biology; and Social Sciences) as well as the individual Department of Biology.

clicker points varied among courses from 8.5 to 19% (Biology 101, 10.4%; 111, 19.23%; 211, 9.6%; 219, 8.5%; 311, 10.9%; and 377, 13.5%).

## Survey

This study evaluated the responses of approximately 550 students (the number of responses varied slightly between questions) to 12 multiple-choice questions about the use of wireless student response units in biology courses (Appendix A in the Supplemental Material includes the text of each question and its answer categories). These questions included three questions about students' overall impression of clickers, three questions exploring students' impressions of specific benefits of using clickers, three questions related to students' recommendations regarding how instructors should use clickers, and two questions asking about the frequencies of technical problems students experienced while using clickers. One question asked whether students prefer to study individually or cooperatively while preparing for exams; responses to this question as well as course and students' course grades were used to explore variation in students' opinions about the use of clickers.

The students were introduced to the survey in lecture, and the survey was posted on the course websites 7 d before finals week. The survey was left open for 2 wk. Both announcements in lecture and a preamble at the entrance to the surveys in the websites explained that participation in the surveys was completely voluntary and that responses would be reported at the population, rather than individual student, level. Students' participation rates (percentage of students taking the final exam who completed the clicker survey) ranged from 60 to 85% (Biology 101, 60.7%; 111, 72.5%; 211, 61.3%; 219, 80.0%; 311, 84.5%; and 377, 84.7%).

**Analytical Methods: Survey Responses among Courses.** For each of the 12 survey questions, stepwise chi-square contingency table analyses were used to compare the distribution of students' answers across courses. Because very few students chose the more negative responses to some questions, these answer categories had to be pooled to maintain adequate subsample sizes for the chi-square analyses. Appendix A in the Supplemental Material indicates which survey answer categories were pooled during the analyses. The first step of each of these analyses asked whether the distribution of students' responses across the answer categories differed among the six courses. If the answers to a survey question were not significantly different among the six courses, the pooled answers across all courses were described. If the distribution of answers differed among courses, the course with the most divergent distribution of answers was removed, and the course by answer interaction of the remaining five courses was analyzed. This stepwise removal of courses was repeated until the remaining courses were not significantly different; the answers of these remaining courses were then pooled and described. If more than one course had been removed during this initial sequence of contingency table analyses, the stepwise process was repeated starting with the removed courses to determine whether they differed in students' responses to the survey question. If students' responses to a question separated into two groups corresponding to lower- and upper-division courses, a follow-up analysis limited to students who earned an "A" grade was conducted. Probability values equal to or  $<0.01$  were considered significant in these stepwise contingency table analyses of the survey questions across courses. Because of the many factors that may influence differences in students' responses among courses, we used the 0.01  $\alpha$  level to focus on larger, more meaningful, differences in the distribution of students' answers among courses.

The two upper-division courses were the only courses using Macintosh computers; these computers were found to be less compatible with eInstruction software. Therefore, consideration of differences by academic level in students' responses to survey questions was confounded by the increased number of in-class problems by using clickers in upper-division courses. When the chi-square contingency table analyses of students' responses by course grouped responses to a survey question into upper- and lower-

division courses, post hoc log-linear analyses were used to compare the amount of variation in students' responses to the survey questions that was associated with academic level to that which was associated with the frequency of in-class problems by using clickers. The log-linear models were based on three-way tables of student survey responses by academic level (upper- or lower-division course) and by the frequency of in-class problems by using clickers (no problems, 1 or 2 problems, 3 or more problems). The fit of the full log-linear models including the three-way interaction as well as all two-way interactions and main effects was estimated with a chi-square value. The fit of these full models was compared with that of the models after removing the two-way interactions involving students' responses by academic level; the full models also were compared with that of the models after removing the two-way interactions involving students' responses by the frequency of in-class problems (Systat Software, Inc., 2002). The magnitude of the change in the fit of the models was used as an indication of the amount of variation in students' responses associated with academic level in comparison with that associated with in-class problems.

**Analytical Methods: Survey Responses by Student Grades.** To determine whether students' opinions about the use of clickers differed among students receiving different course grades, it was first necessary to determine whether grades represented similar performance categories relative to peers across the six courses. Before the analyses of the relationships between grades and survey responses, stepwise chi-square contingency table analyses also were used to determine which courses had roughly similar grade ("A," "B," "C," "D," and "F") distributions and so could be included in the same analysis of the relationships between student survey opinions and their course grades. This initial analysis used the more rigorous  $\alpha$  value of 0.01, because the goal was to group courses with roughly similar grade distributions rather than to identify all differences in grades across courses. After grouping courses with similar grade distributions, for those groups with sufficient subsample sizes, simple chi-square contingency table analyses were used to determine whether students' answers to survey questions differed by course grades. The more standard  $\alpha$  level of 0.05 was used in this analysis because any significant differences by grade in students' responses to survey questions was of interest.

**Analytical Methods: Survey Responses by Exam Study Preference.** Simple chi-square contingency table analyses ( $\alpha = 0.05$ ) also were used to determine whether variation in responses to the survey question asking whether students preferred to study for exams individually or cooperatively was associated with variation in students' responses to survey questions about clickers.

## Effects of Clicker Frequency on Content Acquisition

The number of clicker questions asked in each lecture was systematically varied to evaluate the effects of the frequency of clicker questions on students' ability to answer associated exam questions. Within each course, students experienced low-, medium-, and high-clicker lectures. The number of clicker questions associated with these three levels was set within each course, but it was allowed to vary slightly between courses so that each instructor could use clickers at frequencies that were within his or her comfort level (Biology 101, 211, 219, and 311: low-clicker frequency lectures, 0–2 clicker questions; medium, 3–4 questions; high, 5–6; Biology 111: low, 0–1; medium, 2–3; high, 4–6; and Biology 377: low, 0–1; medium, 2; high, 3–5). To avoid biasing variation in clicker frequency in a pattern that might covary with students' performance, such as lecture topic, day of the week, or proximity to exams, clicker frequencies followed a temporal Latin square design. For example, if there were nine lectures before an exam, the clicker level of the first lecture would be randomly determined (e.g., medium) and the remaining lectures of a course with Monday, Wednesday, and Friday lectures would follow a pattern similar to a  $3 \times 3$  Su Do Ku puzzle (e.g., week 1: medium, high, low; week 2: high, low, me-

dium; and week 3: low, medium, high). Four courses (Biology 101, 111, 211, and 219) followed this pattern throughout the semester with the minor exception of setting 1 of the Latin squares in each course so that a low (no-clicker) day coincided with a guest lecture. Instructors of the remaining two courses (Biology 311 and 377) found that following a predetermined clicker level interfered with developing optimal lectures for each topic. Although the instructors of these two courses continued to vary clicker frequency, they did not follow a predetermined sequence. During the set of lectures before exam 1 of each course, students were purchasing and registering clickers, and the instructors and students were adjusting to the use of clickers. The experiment began after exam 1 and ran through the last week of the semester.

### Analytical Methods: Content Acquisition

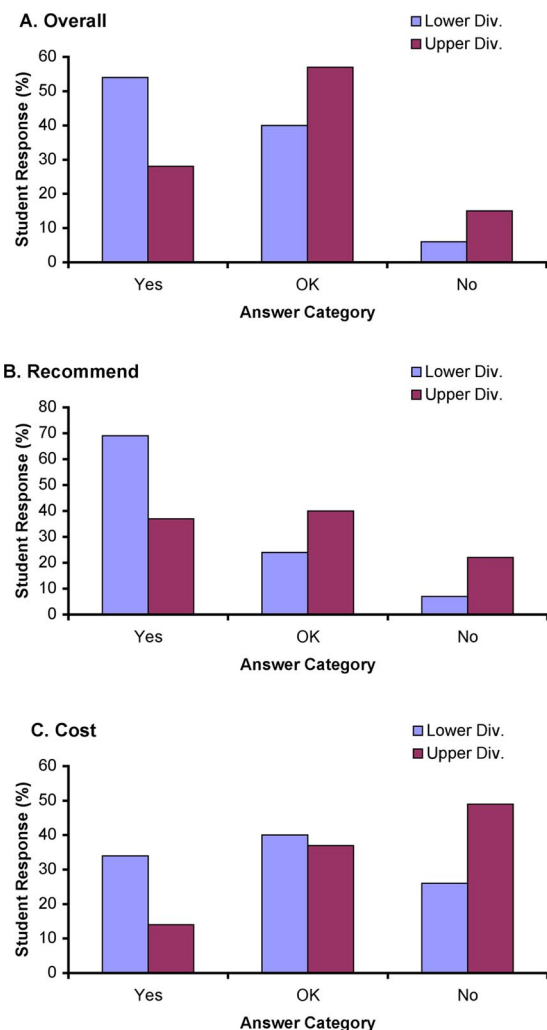
The effect of variation in the frequency of clicker questions on students' exam performance across all courses was evaluated with an analysis of variance (ANOVA). The dependent variable was the percentage of students who correctly answered multiple-choice exam questions tied to specific lectures. The analyses included questions from the second midterm exams through noncomprehensive questions in the final exams. Across all courses, the analysis of exams included 659 exam questions. The ANOVA used three main effects to investigate variation in students' performance across these exam questions: 1) course (6 different biology courses), 2) exams (exams 2, 3, and 4), and 3) number of clicker questions in associated lecture (low, medium, or high). This analysis also included each two-way interaction term that included the effects of clickers: 1) clicker by exam, which determined whether the effects of clickers on scores varied among exams; and 2) clicker by course, which determined whether the effects of clickers on exam scores varied among courses. Although the main effect of clickers in the ANOVA was used to determine whether exam scores significantly differed among the three levels of clicker use, polynomial contrasts were used to ask more specific questions. These contrasts evaluated the significance of linear and quadratic variation among the mean exam question scores of the three clicker levels. The linear contrast determined whether there was a significant linear increase (or decrease) in exam scores associated with low-, medium-, and high-clicker use. This linear contrast was the statistical test most closely associated with our hypothesis that content acquisition increases with increased use of clickers. The quadratic contrast determined if exam scores significantly varied among clicker levels in a curved quadratic pattern.

## RESULTS

### Survey Responses among Courses

**General Value.** Students' responses to the three questions probing their general view of the value of using clickers in biology courses were positive. Across all six courses (averaged across lower- and upper-division courses, which are shown separately in Figure 1), 48% of students categorizing the overall value of clickers found them to be a great addition to the class (the "yes" category in Figure 1A; see Appendix A in the Supplemental Material for complete survey text), whereas only 8% found them to be a distraction or very detrimental. When asked whether they would recommend a clicker class, 62% selected yes, absolutely, and only 10% would not recommend a clicker course (Figure 1B averaged across lower- and upper-division responses). When asked whether clickers were worth the cost, 69% responded yes or probably, whereas 31% felt they were not (Figure 1C averaged across lower- and upper-division responses).

The distribution of students' responses to each of these three questions significantly differed between lower- and upper-division courses (Table 2, General value questions). As shown in Table 2, the chi-square contingency table analyses showed that the distribution of students' answers across the three answer categories was significantly different among the courses until the lower- and upper-division courses were partitioned into separate analyses. These distributions of students' answers across the three answer categories, for each of the questions addressing students' general impression of clickers, did not differ among the four lower-division courses, nor did they vary between the two upper-division courses (Table 2, General value questions). They only differed between lower- and upper-division courses. Students in the lower-division courses had more favorable responses to each of the three questions evaluating their general impression of clickers than did students in upper-division courses (Figure 1). Follow-up analyses, limited to "A" stu-



**Figure 1.** Each panel represents students' responses to individual survey questions. The bars within each panel show the percentage of lower-division students and the percentage of upper-division students who chose each answer category.

**Table 2.** Stepwise chi-square contingency table analyses of student responses to survey questions across courses

Survey questions by category	All courses <sup>a</sup>	Analyses of courses remaining after the sequential removal of courses <sup>b</sup>	Stepwise analyses of removed courses <sup>c</sup>
General Value			
Overall, I found the clickers to be	$\chi^2 = 33.59$ , $df = 10$ , $p < 0.001$ Note: Because this $p$ value is $<0.01$ , it indicates a significant difference among the six courses. <sup>d</sup>	101, 111, 211, 219, 311, (377), $p = 0.002$ Note: This step indicates that after removing the course with the most divergent answers (377), the remaining courses (101, 111, 211, 219, 311) are still significantly different. 101, 111, 211, 219, (311, 377), $p = 0.634$ Note: This step indicates that after removing the two most divergent courses (311, 377), the distribution of student answers to this question do not differ significantly among the remaining courses (101, 111, 211, 219).	311, 377, $p = 0.738$ Note: This indicates that students in the courses removed in the previous analyses (311, 377) had similar answers to this question.
I would recommend a clicker class to a friend	$\chi^2 = 65.54$ , $df = 10$ , $p < 0.001$	101, 111, 211, 219, 311, (311), $p = 0.002$ ; 101, 111, 211, 219, (311, 377), $p = 0.030$	311, 377, $p = 0.110$
The cost of ... the clicker...	$\chi^2 = 42.47$ , $df = 10$ , $p < 0.001$	101, 111, 211, 219, 311, (377), $p = 0.001$ ; 101, 111, 211, 219, (311, 377), $p = 0.118$	311, 377, $p = 0.057$
Benefits of Clickers			
The clickers kept me more interested	$\chi^2 = 21.15$ , $df = 15$ , $p = 0.132$		
The clickers made it more likely that I would attend class	$\chi^2 = 29.92$ , $df = 15$ , $p = 0.012$	111, 211, 219, 311, 377 (101), $p = 0.229$	
The clickers helped me understand and/or learn ...	$\chi^2 = 31.44$ , $df = 15$ , $p = 0.008$	101, 111, 211, 311, 377 (219), $p = 0.039$	
How should we use clickers?			
How many clicker questions do you recommend ...	$\chi^2 = 27.93$ , $df = 10$ , $p = 0.002$	101, 111, 211, 219, 311, (377), $p = 0.041$	
The number of clicker points is	$\chi^2 = 21.13$ , $df = 10$ , $p = 0.020$		
I prefer to answer clicker questions	$\chi^2 = 47.48$ , $df = 10$ , $p < 0.001$	101, 111, 211, 219, 377, (311), $p = 0.003$ ; 101, 111, 211, 219, (377, 311), $p = 0.102$	311, 377, $p = 0.061$
Problems with clicker use			
Did you have any problems registering your clicker?	$\chi^2 = 28.81$ , $df = 10$ , $p = 0.001$	101, 211, 219, 311, 377, (111), $p = 0.267$	
... problems getting your clicker to work in class?	$\chi^2 = 83.66$ , $df = 15$ , $p < 0.001$	101, 111, 211, 219, 311, (377), $p < 0.001$ ; 101, 111, 211, 219, (311, 377), $p = 0.005$ ; 101, 111, 211, (219, 311, 377), $p < 0.053$	219, 311, 377, $p < 0.001$ ; 311, 377, $p = 0.604$

<sup>a</sup> Column 2 describes the significance of the initial step determining whether student answers to a question differed among the six courses.

<sup>b</sup> If there was a significant difference among the six courses (column 2), the first line associated with each question in the third column shows the results of the analysis of the remaining five courses (listing the removed column in parentheses). If the remaining five courses differed, the second line in the third column describes the results of the analysis of the remaining four courses. In the one case of a significant difference among the remaining four courses, the third line in column 3 showed the analysis of the remaining three courses.

<sup>c</sup> The far-right column describes the significance of differences among student answers in the courses removed from the first group.

<sup>a-c</sup> In these analyses,  $p$  values equal to or  $<0.01$  are considered significant.

<sup>d</sup> Notes in the results of the analyses of this question provide an illustrated example of how to interpret this table.

dents, of lower-division courses in comparison with upper-division answers to these three questions revealed similar patterns as the analyses that included all students. Lower-division answers differed from upper-division answers in each case (overall assessment,  $\chi^2 = 8.39$ ,  $df = 2$ ,  $p = 0.015$ ; recommend to a friend,  $\chi^2 = 17.85$ ,  $df = 2$ ,  $p < 0.001$ ; cost of clickers,  $\chi^2 = 8.22$ ,  $df = 2$ ,  $p = 0.016$ ). For each of these three questions, the responses of "A" students were similar to those illustrated for the entire classes in Figure 1.

**Specific Benefits.** Eighty-one percent of students across all courses agreed or strongly agreed that using clickers increased their interest in their course; only 5% of the students did not think clickers improved their interest in the course (Figure 2A). Students' responses to this question about the effect of clickers on their interest in their course did not vary among biology courses (Table 2, Benefits of clickers, analysis of "The clickers kept me more interested").

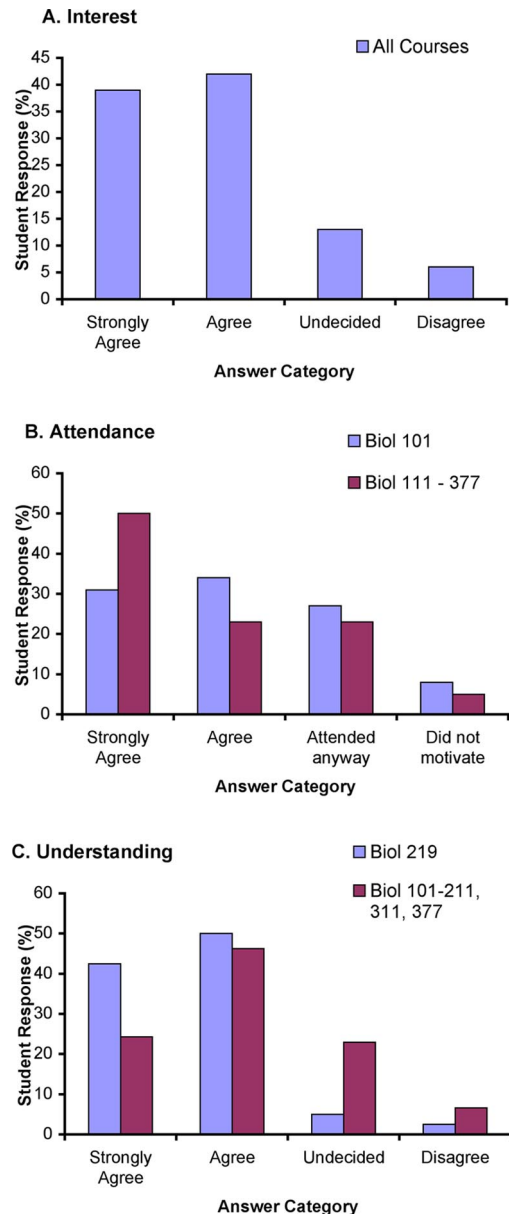
A majority of students responded that using clickers in their course improved their attendance (71% of students across all courses selected strongly agree or agree to the statement that clickers made it more likely that they would attend class). The responses of students in the nonscience majors introductory biology course (Biology 101) differed from, and were slightly less positive than, students in the remaining five courses (Table 2, Benefits of clickers, analyses of "The clickers made it more likely that I would attend class"; and Figure 2B).

Averaged across all courses, 70% of students agreed or strongly agreed with the statement that using clickers improved their understanding of course material; only 6% of students disagreed with this statement. Students in the microbiology course for allied health majors (Biology 219) had a significantly more positive response to this question (Table 2, Benefits of clickers, analysis of "The clickers helped me understand and/or learn . . ."). Ninety-two percent of these students claimed that using clickers improved their understanding of course content (Figure 2C combining strongly agree and agree responses of Biology 219 students).

**How Should We Use Clickers?** Students in the most advanced course in this study, Biology 377, had a different distribution of answers regarding the optimal number of clicker questions in a lecture than did students in the other five courses (Table 2, How should we use clickers?, analysis of "How many clicker questions do you recommend . . ."). A majority of students in all courses recommended three or four questions per lecture (Figure 3A). A majority of the remaining students in Biology 377 preferred one to two questions over five to six questions; in contrast, a majority of the remaining students in Biology 101–311 preferred five to six questions over one to two questions (Figure 3A). Although the percentage of the total course grade derived from clicker questions varied among the courses, there was no significant variation among courses in students' satisfaction with the number of points associated with clicker questions (Table 2, How should we use clickers?, analysis of "The number of clicker points is . . ."). Overall, 71% of the students agreed with the number of points allocated to clicker questions in their particular course.

The third question about how we should use clickers in our courses asked students whether they preferred to an-

swer clicker questions without discussion, after discussion with classmates, or first by themselves and then again after discussion. The responses to this survey question significantly varied between lower-division and upper-division courses (Table 2, How should we use clickers?, analysis of "I prefer to answer clicker questions"). Students in upper-division courses showed a stronger preference for an-

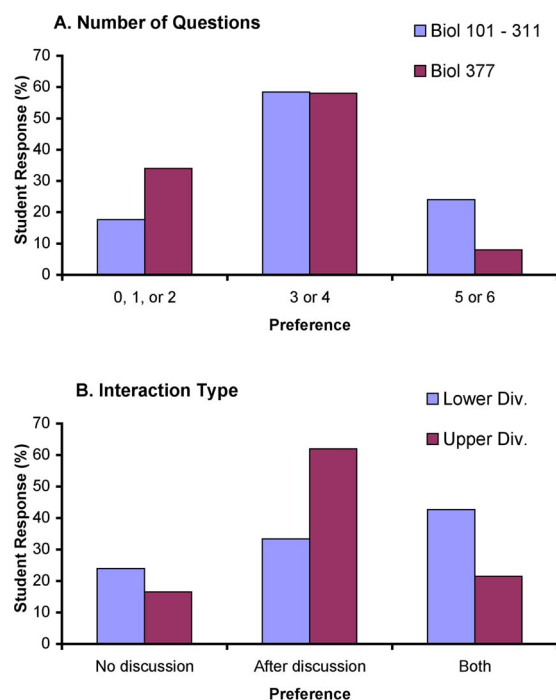


**Figure 2.** Each panel illustrates students' responses to individual questions about specific benefits associated with the use of clickers. (A) Percentage of students' responses in each answer category across all courses. (B) Responses of students in Biology 101 separately from those of students in the remaining five courses. (C) Responses of students in Biology 219 separately from those of students in the remaining courses.

swering questions after discussion with their classmates (Figure 3B).

**Mechanical Problems with Registering, Paying for, and Using Clickers.** Students in Biology 111, the course with the highest proportion of freshmen (Table 1), had significantly more problems with Web-based registration and payment than did students in the other courses (Table 2, Problems with clicker use, analysis of “Did you have any problems registering your clicker?”). Nearly one-third of Biology 111 students had problems registering or paying for their clickers (Figure 4A). In the other courses, 84% of students had no problems registering or paying for their clickers (Figure 4A). There also were significant differences among courses in the frequency of students who had problems in class with their clickers (Table 2, Problems with clicker use, analysis of “. . . problems getting your clicker to work in class?”). This category includes problems with students’ individual wireless response units or with the instructors’ receivers and associated software. More students experienced these problems in our two upper-division courses; students in Biology 219 had the fewest in-class problems using clickers (Figure 4B).

**Class Level and Frequency of In-Class Problems.** As described above, students’ responses to four survey questions differed by class level (the 3 questions exploring students’ general impression of clickers and the question about their interaction preference). Also, as described in the preceding section, students in upper-division courses, which used Macintosh computers, had more in-class problems using

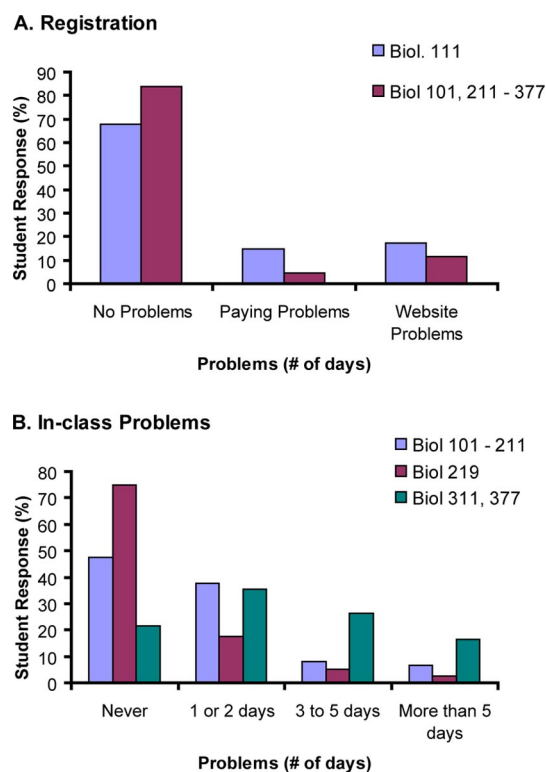


**Figure 3.** (A) Percentage of students from Biology 101 through Biology 311 who prefer specific number of clicker questions in each lecture compared with the preferences of students in Biology 377. (B) Interaction preferences of students in lower- and upper-division courses.

clickers. The results of the log-linear analyses show that the differences in the responses of students to the four survey questions is better explained by considering course level than by considering the frequency of in-class problems. In each of these four questions, more of the variation in students’ responses to the survey questions was associated with their academic level than with their estimate of the frequency of in-class problems. As a result, when academic level was removed from the log-linear models predicting students’ answers the fit of the models changed more significantly than when the frequency of in-class problems was removed (Table 3).

### Survey Responses by Grades

**Course Grade Distributions.** To determine whether we could pool students across courses when analyzing the relationships between students’ course grades and their survey responses, we first determined whether grades represented the same performance category, relative to peers, across the six courses. Contingency table analyses revealed significant differences in grade distributions across all six courses ( $\chi^2 = 86.47$ ,  $df = 20$ ,  $p < 0.001$ ). Stepwise removal of courses revealed that the two microbiology courses (Biology 219 and 311) had similar grade distributions ( $\chi^2 = 4.42$ ,  $df = 2$ ,  $p = 0.110$ ; “C,” “D,” and “F” grade categories were pooled



**Figure 4.** (A) Percentage of students in Biology 111, and in the remaining courses, who had no problems or had problems registering or paying online for their clickers. (B) Percentage of students in Biology 101, 111, and 211 in contrast to students in Biology 219, or students in the upper-division courses, who had various frequencies of problems using their clickers in class.



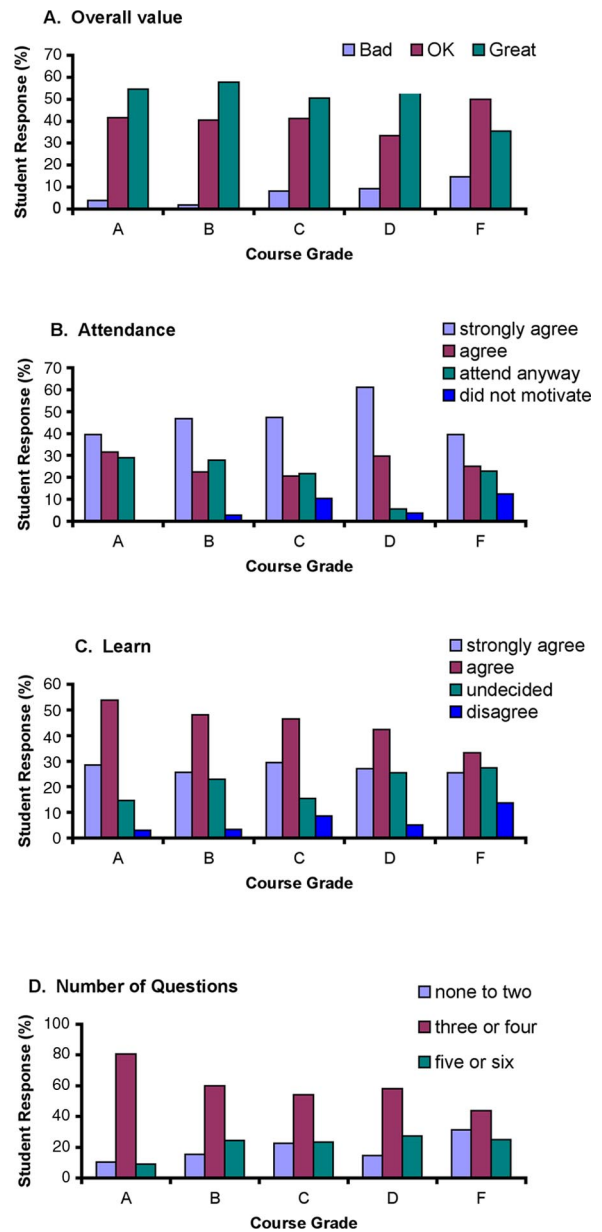
**Table 3.** Log-linear analyses of variation in student responses associated with academic level and frequency of in-class problems with clickers

Survey question	Change in fit of the log-linear model due to removal of the two-way interactions of each variable with each survey question	
	Class problems	Academic level of course
Overall	$\chi^2 = 17.32, df = 4, p = 0.0017$	$\chi^2 = 54.35, df = 2, p < 0.0001$
Recommend clickers	$\chi^2 = 17.46, df = 4, p = 0.0016$	$\chi^2 = 80.08, df = 2, p < 0.0001$
Cost of clickers	$\chi^2 = 10.33, df = 4, p = 0.0352$	$\chi^2 = 18.07, df = 2, p < 0.0031$
Interaction preference	$\chi^2 = 9.27, df = 4, p = 0.0547$	$\chi^2 = 54.35, df = 2, p = 0.0001$

to maintain sufficient subsample sizes). Analyses of the remaining courses (Biology 101, 111, 211, and 377) found that their grade distributions significantly differed ( $\chi^2 = 52.40, df = 12, p < 0.001$ ) until Biology 377 was removed. The remaining three courses (Biology 101, 111, and 211) had similar grade distributions ( $\chi^2 = 1.74, df = 8.00, p < 0.011$ ). These analyses identified three groups with similar course grade distributions: group 1: Biology 101, 111, and 211; group 2: Biology 219 and 311; and group 3: Biology 377. Only group 1 had sufficient numbers of students across all grades to allow contingency table analyses of associations between grades and students' responses to survey questions. Therefore, the following analyses of the associations between grades and survey answers were limited to students from Biology 101, 111, and 211.

**General Value.** Students' opinion of the overall value of clickers differed by their course grade ( $\chi^2 = 15.90, df = 8, p = 0.044$ ). The percentage of students who felt that clickers were a distraction or were detrimental to the class gradually increased as grades decreased, although it never reached >15% of students in a grade category (Figure 5A). A majority of students claimed that clickers were a great addition to their course in each grade category, except for students who failed the course (Figure 5A, category labeled "bad"). However, when asked whether they would recommend a course using clickers to a friend, student responses did not vary by their course grade ( $\chi^2 = 6.37, df = 8, p = 0.606$ ). There also was no difference among students in the five different grade categories in their assessment of the cost of clickers ( $\chi^2 = 12.97, df = 8, p = 0.113$ ).

**Specific Benefits.** Students consistently felt that clickers improved their interest in the courses; this opinion did not vary by course grades ( $\chi^2 = 11.20, df = 12, p = 0.512$ ). Students' opinions of the effect of clicker use on their attendance did vary by their course grades ( $\chi^2 = 30.7, df = 12, p = 0.002$ ). The percentage of students who strongly agreed that clickers motivated them to attend lecture increased from "A" to "D" students; as the percentage who claimed they would have attended anyway declined, these trends reversed with students who failed the course (Figure 5B). Students' opinions of the influence of clickers on their ability to learn the course material also varied by grade ( $\chi^2 = 21.40, df = 12, p = 0.045$ ). Whereas 25–30% strongly agreed that clickers helped them learn across all grades, the number of students who simply agreed that clickers helped them learn declined with their course grade (Figure 5C).



**Figure 5.** Each panel represents the pooled responses of students from Biology 101, 111, and 211 to specific survey responses. The bars illustrate the percentage of students in each grade category who chose each answer category.

**How Should We Use Clickers?** Students' recommendations regarding the number of clicker questions that we should ask in each lecture varied by grade ( $\chi^2 = 24.28$ ,  $df = 8$ ,  $p = 0.002$ ). Students who earned an "A" in Biology 101, 111, or 211 more consistently (81%) recommended three or four clicker questions, whereas students in each of the "B" through "F" grade categories had increasingly more diverse recommendations (Figure 5D). Students' recommendations regarding the proportion of course points associated with clickers did not vary by grade ( $\chi^2 = 11.62$ ,  $df = 8$ ,  $p = 0.169$ ). Students' preference for answering clicker questions individually or after discussion with classmates also did not vary by their course grade ( $\chi^2 = 9.63$ ,  $df = 8$ ,  $p = 0.292$ ).

**Mechanical Problems with Registering, Paying for, and Using Clickers.** Although the proportion of students who had problems registering or paying for their clickers online did not vary by course grade ( $\chi^2 = 8.90$ ,  $df = 8$ ,  $p = 0.351$ ), the frequency of problems using clickers in class did vary across course grades ( $\chi^2 = 16.42$ ,  $df = 8$ ,  $p = 0.037$ ). Students who failed the class were more than twice as likely to claim that they had trouble using their clickers in three or more lectures. Nearly one third of these students had this type of problem in three or more lectures.

### Survey Responses by Study Preference

A slight majority (55%) of students preferred to study individually for exams. Many students (42%) preferred to study in groups of two or three students. Very few students (3%) preferred to study for exams with four or more students. Because so few students preferred to study in larger groups, when analyzing the relationships between exam study preference and responses to the clicker survey, we collapsed the study preference responses into two categories: 1) prefer to study individually and 2) prefer to study cooperatively. Students' responses to each of the three questions assessing their general impression of clickers did not vary between students who preferred to study for exams individually and those who preferred to study cooperatively (overall value:  $\chi^2 = 3.02$ ,  $df = 2$ ,  $p = 0.221$ ; recommend clickers:  $\chi^2 = 2.36$ ,  $df = 2$ ,  $p = 0.307$ ; and cost:  $\chi^2 = 1.12$ ,  $df = 2$ ,  $p = 0.571$ ). Students' responses to each of the questions exploring specific benefits of clickers also did not vary by exam study preference (interest:  $\chi^2 = 1.23$ ,  $df = 3$ ,  $p = 0.746$ ; attendance:  $\chi^2 = 3.15$ ,  $df = 3$ ,  $p = 0.370$ ; and understanding:  $\chi^2 = 0.363$ ,  $df = 3$ ,  $p = 0.948$ ). Students' recommendations regarding the number of clicker questions and the number of points associated with clickers did not vary between students who preferred to study for exams individually and those who preferred to study with other students (number of questions:  $\chi^2 = 1.06$ ,  $df = 3$ ,  $p = 0.787$ ; points:  $\chi^2 = 4.85$ ,  $df = 2$ ,  $p = 0.088$ ). However, students who preferred to study individually for exams also were more likely to prefer to answer clicker questions without discussion than were students who preferred to study cooperatively for exams ( $\chi^2 = 23.82$ ,  $df = 2$ ,  $p < 0.001$ ). Of students who preferred to study individually, 29.8% also preferred to answer clicker questions without discussion; of students who preferred to study cooperatively for exams, only 13.4% preferred to answer clicker questions without discussion.

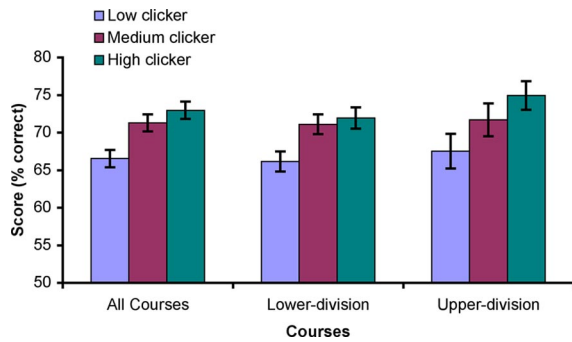
### Content Acquisition

The significance of the effects of course, exams, and clickers on scores on exam questions is shown in Table 4. The ANOVA was run across all courses as well as separately for lower- and upper-division courses. The test of the main effect of clickers on exam scores, which asks whether exam scores differed among the three clicker levels, was significant when all courses were included in the analysis, was significant when only lower-division courses were included, and was inconclusive when only the two upper-division courses were included (Table 4, Analyses of clickers). The linear contrasts of the effects of clickers on exam scores asked the more specific question of whether there was a consistent linear change in exam scores from low-, to medium-, to high-clicker lectures. When considered across all courses as well as separately for lower- and upper-division courses, there was a significant linear increase in exam scores across the three levels of clicker frequencies (Table 4, Analyses of linear contrast of clickers; and Figure 6). Although exam scores generally differed among courses, the positive linear relationship between clicker frequency and exam scores was consistent across all courses (Table 4, analyses of Clickers  $\times$  course).

**Table 4.** Analyses of variance of effects on student scores on exam questions

Effects by course group	F statistic	df	Probability <sup>a</sup>
<b>All courses</b>			
Course	4.15	5,635	0.001*
Exam	0.04	2,635	0.957
Clickers	8.10	2,635	<0.001*
Clickers $\times$ course	0.54	10,635	0.866
Clickers $\times$ exam	0.15	4,635	0.965
Linear contrast of clickers	14.92	1,635	<0.001*
Quadratic contrast of clickers	2.05	1,635	0.153
<b>Lower-division courses</b>			
Course	6.96	3,428	<0.001*
Exam	0.11	2,428	0.898
Clickers	5.56	2,428	0.004*
Clickers $\times$ course	0.76	6,428	0.599
Clickers $\times$ exam	0.16	4,428	0.957
Linear contrast of clickers	8.77	1,428	0.003*
Quadratic contrast of clickers	3.14	1,428	0.077
<b>Upper-division courses</b>			
Course	0.33	1,201	0.566
Exam	0.05	2,201	0.956
Clickers	2.91	2,201	0.057 <i>i</i>
Clickers $\times$ course	0.29	2,201	0.748
Clickers $\times$ exam	0.61	4,201	0.659
Linear contrast of clickers	5.78	1,201	0.017*
Quadratic contrast of clickers	<0.001	1,201	0.985

<sup>a</sup> All significant effects are indicated with an asterisk. The test of the main effect of clickers in upper-division courses is inconclusive (*i*); remaining effects are insignificant ( $\alpha = 0.05$ ).



**Figure 6.** Each cluster of three bars shows the effects of clicker levels on student performance on associated exam questions. The height of the bars is the mean score on exam questions, and the line shows the SE above and below the mean.

## DISCUSSION

Our three survey questions probing students' general impression of clickers revealed that students in all six biology courses had positive overall opinions about the use of clickers. However, students in lower-division courses had more positive overall impressions than students in upper-division courses (Figure 1, A–C). Studies of clickers in lower-division courses describe strongly positive student evaluations of the value of clickers (Elliott, 2003; Hatch *et al.*, 2005). In contrast, studies of the use of clickers in upper-division courses have generated more varied student responses (strongly positive: Casanova, 1971; Knight and Wood, 2005; and mixed opinions: Slain *et al.*, 2004). Studies of the use of clickers in multiple courses describe more favorable student opinions in the first semester of two-semester courses in physics and in business (Cue, 1998) and describe more favorable student opinions in nonmajors compared with science majors courses (Dufresne *et al.*, 1996). Our analyses of lower-division compared with upper-division courses were confounded by the use of Macintosh computers only in the upper-division courses. Students experienced increased in-class problems using clickers in these courses due to compatibility problems between eInstruction and Macintosh systems (also see Hatch *et al.*, 2005 regarding problems using infrared clickers with Macintosh computers). However, the log-linear analyses showed that course level had a stronger impact on student's opinions than did their number of in-class problems using clickers (Table 3, the significance of changes of the fit of the model associated with removing academic level are greater than those associated with removing the frequency of in-class problems), suggesting that the course-level effect is not simply an artifact of the type of computer used by the instructors. Additionally, our results indicate that the lower general impression of clickers expressed by students in upper-division courses is not due to these courses having a higher proportion of successful science students. Fortunately, assigning lower overall value to the use of clickers does not seem to be an initial preference of successful students as they work their way through introductory science courses. Students who earned an "A" in the lower-division courses had much stronger overall impressions of clickers than did students who earned an "A" in

the upper-division courses. Students' survey responses seldom varied across students' course grades in our introductory courses; when there was significant variation in responses across grade categories, it was due to more negative responses of students with very low grades. Collectively, these results suggest that initially students have positive overall impressions of clickers across grade levels, but students who have successfully progressed to upper-division science courses place less value on the use of clickers. Most of these students reached their upper-division courses without having consistently used clickers in their previous courses (this was the first clicker class for approximately two thirds of the students in Biology 311 and 377). Clickers represented a pedagogical change for these students. Whereas upper-division students' overall assessment of the value of clickers was not as strong, they readily agreed (as strongly as lower-division students) that clickers made lectures more interesting, and improved their attendance and their understanding of the material.

Although upper-division students may be hesitant to embrace more interactive instruction methods, they may be the students most in need of the type of learning associated with these methods. There is often a disconnect between students' perceptions of the course learning objectives (based on experience taking term-recognition exams) and the instructor's expectations of the level of student learning (Bloom, 1956; NRC, 2000; Zoller, 2000). As students progress from introductory to advanced science courses, they are expected to learn at increasingly difficult levels progressing from term recognition and simple memorization to synthesis and interpretation. Unfortunately, this transition may not be clearly previewed to students. Therefore, students attempt to apply strategies sufficient for accomplishing tasks associated with lower-order learning skills to more challenging learning tasks in upper-division courses. In longitudinal analyses of student exam performance in Pathophysiology and Therapeutics in a Doctor of Pharmacy program, Slain *et al.* (2004) found that students did not perform better on simple memorization questions when the course was taught using clickers to encourage discussion than when the course was taught by the same instructors without clickers; however, they did perform better on analytical questions when the course was taught using clickers. This difference in the impact of interactive learning methods on lower- and higher-order learning is not limited to advanced students. In a convincing analysis of the effects of cooperative learning on ninth-grade earth science students, Chang and Mao (1999) found that cooperative learning did not improve students' performance on questions requiring only lower-level cognitive skills, but cooperative learning did improve student performance on higher-level questions. Our survey results suggest that upper-division students who have successfully progressed through science courses, many of which are taught using traditional methods, are less enthusiastic about new pedagogical methods associated with increasing interactions in the lecture course. However, if they avoid adopting cooperative learning methods and continue to apply traditional study methods to the more challenging expectations of advanced courses, and ultimately to professional positions, they may struggle with the higher-order thinking skills required to be successful in these courses and positions. We suggest that the solution to this dilemma is the

consistent use of interactive learning strategies throughout the curriculum with a concurrent emphasis on the development of, and assessment of, higher-level thinking skills.

Although the positive effect of clickers on student performance was consistent across all courses, a higher proportion of students in the microbiology course for allied health majors (Biology 219) correctly thought that using clickers improved their understanding. This course is primarily composed of prenursing students who hope not only to pass their required microbiology course but also to learn the material well enough to pass the National Council Licensure Examination. This strong extrinsic motivation of prenursing students to succeed in the course may be paired with less confidence in their ability to learn science than is typical of science majors. Interactive teaching methods, such as the use of clickers, may be particularly effective at improving the confidence level, as well as the performance, of nonscience majors who are under strong pressure to succeed in a science course. Instructors may need to introduce and present the use of clickers differently depending on the nature of their course: students in upper-division science courses may need more convincing that the effective use of clickers will improve their grades than students in lower-division science courses, and students in applied science courses may be most receptive. Fortunately, we found only minor differences of opinion about the use of clickers when students were categorized by their course grade, or by their preference for studying individually or cooperatively. Instructors may need to present clickers differently between courses; however, they will be received in a similar manner by most students within each course.

Although we found differences among courses in students' opinions about the use of clickers, clickers improved student performance in all courses. Students in both lower- and upper-division courses performed significantly better on exam questions associated with high-clicker-use lectures than on exam questions associated with low clicker use (Figure 6). This conclusion adds a new level of specificity to our understanding of the effects of clicker use on student performance. Latitudinal comparisons of concurrent sections of courses taught with and without wired student response units have found no improvement on exam performance associated with the use of this technology (Bessler and Nisbet, 1971; Casanova, 1971; Brown, 1972). As Judson and Sawada (2002) point out, classrooms in the 1970s, such as those in the latitudinal studies discussed above, tended to use wired student response units to enhance student-to-instructor communication, but not to enhance discussion and cooperative learning among students. More recent longitudinal studies comparing courses taught with clickers to the same courses taught without clickers have shown significant conceptual gains associated with clicker use (Knight and Wood, 2005); significant gains limited to specific courses and question types (Slain *et al.*, 2004); and no significant differences in performance, but some encouraging trends in the face of confounding factors (Paschal, 2002). The strength of these latitudinal and longitudinal comparisons is that they evaluate the total effect of a course innovation on student performance; the limitation is that they do not necessarily isolate the effects of a specific method, such as clickers. By varying clicker frequency within courses, our experimental design was not confounded by extraneous fac-

tors contributing to variation in student performance between sections of a course in latitudinal comparisons and between semesters in longitudinal comparisons. However, the fine scale of our comparisons did not allow us to consider course-level effects of the use of clickers on student performance due to factors such as improved overall attendance or student attitudes. In spite of excluding these course-level impacts of using clickers, but perhaps due to the increased specificity of our experimental design, we found that increased clicker use across a variety of biology courses resulted in improved performance on exam questions. Although our study experimentally demonstrated that increased use of clickers is associated with increased learning, this study was not designed to reveal the underlying mechanism of this effect. However, establishing this relationship helps set the stage for investigations of more mechanistic hypotheses. For example, this positive impact of clicker use on student learning may be due to a variety of factors associated with using clickers: increased student engagement associated with each student committing to an answer rather than many passively waiting while a few students answer verbal questions; higher attention levels due to more periodic breaks from lecture; immediate formative assessments benefiting both instructors and students throughout the lectures; and increased discussion among students and between students and the instructor of concepts that students find challenging. Our study also reveals variation in students' impression of clickers, although not in the effects of clickers on their learning, between courses. Future research could investigate explanations of these differences among courses. For example, if the lower overall impression of clickers expressed by upper-division students is due to presenting them with a change in pedagogy when they have been successful in more traditionally taught courses, then this difference should attenuate as students who have been introduced to these methods in their lower-division courses begin to move into the upper-division courses. Students' impression of clickers also may be influenced by class size, causing students in large lower-division courses to express particularly strong preferences for the use of clickers. The strong preference for clicker use shown by students in our microbiology course for allied health majors suggests that the nature of students' motivation and their attitudes toward science may influence their responses to the use of clickers in science classrooms. Although the mechanisms that mediate the impacts of clickers on student learning are not well established, we recommend using clickers to improve student attitudes and learning.

## REFERENCES

- Allen, D., and Tanner, K. (2005). Infusing active learning into the large-enrollment biology class: seven strategies, from the simple to complex. *Cell Biol. Educ.* 4, 262–268.
- Bessler, W. C., and Nisbet, J. J. (1971). The use of an electronic response system in teaching biology. *Sci. Educ.* 55, 275–284.
- Bloom, B. S. (1956). *Taxonomy of Educational Objectives—The Classification of Educational Goals, Handbook I: Cognitive Domain*, New York: David McKay.
- Boyer Commission on Educating Undergraduates (1998). *Reinventing Undergraduate Education: A Blueprint for America's Research*

- Universities. SUNY, Stony Brook, NY. <http://naples.cc.sunysb.edu/pres/boyer.nsf> (accessed 11 January 2007).
- Brown, J. D. (1972). An evaluation of the Spitz student response system in teaching a course in logical and mathematical concepts. *J. Exp. Educ.* 40, 12–20.
- Casanova, J. (1971). An instructional experiment in organic chemistry. *J. Chem. Educ.* 48, 453–455.
- Chang, C. Y., and Mao, S. L. (1999). The effects on students' cognitive achievement when using the cooperative learning method in earth science classrooms. *Sch. Sci. Math.* 99, 374–379.
- Cue, N. (1998). A universal learning tool for classrooms? In: Proceedings of the First Quality in Teaching and Learning Conference. Hong Kong, SAR, China. <http://celt.ust.hk/ideas> (accessed 11 January 2007).
- Dufresne, R. J., Gerace, W. J., Leonard, W. J., Mestre, J. P., and Wenk, L. (1996). Classtalk: a classroom communication system for active learning. *J. Comp. High. Educ.* 7, 3–47.
- Elliott, C. (2003). Using a personal response system in economics teaching. *Int. Rev. Econ. Educ.* 1, 80–86.
- Hatch, J., Jensen, M., and Moore, R. (2005). Manna from heaven or “clickers” from hell. *J. Coll. Sci. Teach.* 34, 36–39.
- Jensen, M., Moore, R., and Hatch, J. (2002). Electronic cooperative quizzes. *Am. Biol. Teach.* 64, 169–174.
- Judson, E., and Sawada, D. (2002). Learning from the past and present: electronic response systems in college lecture halls. *J. Comput. Math. Sci. Teach.* 21, 167–181.
- Knight, J. K., and Wood, W. G. (2005). Teaching more by lecturing less. *Cell Biol. Educ.* 4, 298–310.
- Klionsky, D. J. (2001). Constructing knowledge in the lecture hall. *J. Coll. Sci. Teach.* 31, 246–251.
- National Research Council (2000). *How People Learn: Brain, Mind Experience, and School*. Washington, DC: National Academies Press.
- NRC (2003). *Bio 2010, transforming undergraduate education for future research biologists*. Washington, DC: National Academies Press.
- Novak, J. D. (2002). Meaningful learning: the essential factor for conceptual change in limited or inappropriate propositional hierarchies leading to empowerment of learners. *Sci. Educ.* 86, 548–571.
- Paschal, C. B. (2002). Formative assessment in physiology teaching using a wireless classroom communication system. *Adv. Physiol. Educ.* 26, 299–308.
- Rao, S. P., and DiCarlo, S. E. (2000). Peer instruction improves performance on quizzes. *Adv. Physiol. Educ.* 24, 51–55.
- Roy, H. (2003). Studio vs. interactive lecture demonstration—effects on student learning. *Bioscience* 29, 3–6.
- Slain, D., Abate, M., Hodges, B. M., Stamatakis, M. K., and Wolak, S. (2004). An interactive response system to promote active learning in the doctor of pharmacy curriculum. *Am. J. Pharm. Educ.* 68, 1–9.
- Smith, A. C., Stewart, R., Shields, P., Hayes-Klosteridis, J., Robinson, P., and Yuan, R. (2005). Introductory biology courses: a framework to support active learning in large enrollment introductory science courses. *Cell Biol. Educ.* 4, 143–156.
- Suchman, E., Uchiyama, K., Smith, R., and Bender, K. (2006). Evaluating the impact of a classroom response system in a microbiology course. *Microbiol. Educ.* 7, 3–11.
- Systat Software, Inc. (2002). *SYSTAT 10.2, Statistics I*. Point Richmond, CA: Systat Software, Inc.
- Udovic, D., Morris, D., Dickman, A., Postlethwait, J., and Wetherwax, P. (2002). Workshop biology: demonstrating the effectiveness of active learning in an introductory biology course. *Bioscience* 52, 272–281.
- Tessier, J. (2004). Using peer teaching to promote learning in biology. *J. Coll. Sci. Teach.* 33, 16–19.
- Wood, W. B. (2004). Clickers: a teaching gimmick that works. *Dev. Cell* 7, 796–798.
- Zoller, U. (2000). Teaching tomorrow's college science courses—are we getting it right? *J. Coll. Sci. Teach.* 29, 409–414.