ISSUES IN ACCOUNTING EDUCATION
Vol. 22, No. 3
August 2007
pp. 391–409

Investigating the Effects of Group Response Systems on Student Satisfaction, Learning, and Engagement in Accounting Education

Carla Carnaghan and Alan Webb

ABSTRACT: We examine group response systems (GRS) as an educational tool. We use an experimental approach and student survey data to assess vendors' claims that GRS improve student engagement and feedback, and thus improve learning. A key part of our design involves controlling for effects of moving to a more interactive pedagogy that have been found to affect learning.

For a management accounting course, we find only limited GRS learning effects, as proxied by exam performance. Contrary to our expectations, we find a decline in engagement, as proxied by student oral participation, when GRS are used. We also find little evidence that GRS lead to greater student satisfaction with the course. We do find support for student satisfaction with GRS, from which we infer that implementation problems are not driving our results. In summary, we find little support for vendor claims, when controlling for changes in pedagogy.

Keywords: group response systems; learning; clickers; interactive pedagogy; engagement; education.

Data Availability: Available from authors upon request.

INTRODUCTION

The purpose of this study is to examine the effectiveness of an educational technology, group response systems[1] (GRS), or "clickers," as an educational tool. Vendors and satisfied users of GRS cite numerous benefits including improvements in student satisfaction, engagement, exam performance, and interaction.[2] However, there is limited

Carla Carnaghan is an Assistant Professor at the University of Lethbridge and Alan Webb is an Associate Professor at the University of Waterloo.

We thank the editor, an associate editor, and two anonymous reviewers for their many suggestions. We also thank Tony Atkinson, Irene Gordon, Duane Kennedy, Natalia Kotchetova, Kate Mooney, Bill Wright, and participants at the University of Waterloo accounting research workshop, the 2005 European Accounting Congress, the 2005 Annual Meeting of the Canadian Academic Accounting Association, the 2005 American Accounting Association Annual Meeting, and the University of Waterloo Learning through Teaching and Technology (LT3) seminar for helpful comments on an earlier draft. We also thank Sarah Fernandes for her excellent research assistance. We thank the Gordon H. Cowperthwaite Centre for Professional Accounting Education, the University of Waterloo Learning Initiatives Fund, and the University of Waterloo Faculty of Arts for their financial support.

[1] As noted by Steve Draper (http://www.psy.gla.ac.uk/~steve/ilig), these systems are also referred to as electronic voting systems, personal response systems (PRS), and classroom communication systems (CCS).
[2] Examples of these claims can be found at http://www.einstruction.com and http://www.gtco.com.

evidence or theoretical rationale provided for these claims (Judson and Sawada 2002; Roschelle et al. 2004). Given the interest in this technology by accounting educators and the investment in time and resources to implement GRS in the classroom, we believe research utilizing objective measures on the value of this technology for accounting education, as well as more generally, is needed.[3] We use prior research as well as the conversational framework theory of learning (Laurillard 1993, 2002) to posit why and how GRS could have such results, and then conduct an experiment using both student self-reports and objective measures to test our hypotheses.

Our study extends the GRS literature in three significant ways. First, ours is the only study we are aware of that examines the effects of GRS on objective measures of learning, incremental to pedagogical changes that can accompany the implementation of GRS. Our results show only modest GRS effects on student learning, as proxied by exam scores for questions similar to those asked in class with the GRS, relative to the pedagogical improvements. Second, we use objective measures of student participation (as a proxy for engagement), rather than relying solely on student self-reported participation like prior GRS studies, which may be biased given student satisfaction with the technology. Surprisingly, our results suggest that GRS usage can suppress oral participation. Finally, our within-subjects research design permits comparisons of students' general course satisfaction with and without GRS. Prior GRS studies show that students using GRS have favorable course perceptions, but have generally not compared these to the same course perceptions of non-GRS users. We find that although students were satisfied with our implementation of GRS technology, this had a limited effect on their general satisfaction with the course.

The remainder of this paper is organized as follows: the next section briefly explains the GRS technology and reviews prior research that has examined the effects of GRS in education. We then describe relevant theory to explain the potential role for GRS in improving student learning, engagement, and satisfaction, and state our resulting hypotheses. This is followed by a description of our research design and subsequent analysis of our results. Finally, we summarize our findings and provide some conclusions and possible limitations.

BACKGROUND AND HYPOTHESES

An Overview of GRS Technology

GRS comprise both software and hardware components. The hardware includes individual response pads that typically utilize wireless communication to a receiver, which is in turn connected to a computer/data projector combination. Instructors use GRS software to create various types of questions (such as multiple choice, true/false, and quantitative response questions) and display them using the software and a data projector. Students use their response pads to select an answer, which is transmitted to the receiver and recorded. Responses can then be automatically aggregated and displayed by the software to provide immediate feedback on students' individual comprehension as well as the class's understanding.

Prior Research on GRS Learning Effects

Judson and Sawada (2002) summarize the studies of GRS in the 1960s and 1970s as finding that the technology was associated with no difference in student performance,

[3] Examples of this interest include effective learning strategies forum session papers at both the 2004 and 2005 American Accounting Association (AAA) Annual Meeting (Segovia 2004; Tietz 2005), and discussion in July 2005 on the Accounting Education Using Computers and Multimedia (AECM) mailing list (AECM 2005).

despite student satisfaction with the technology. Judson and Sawada (2002) note that a key difference between early studies of GRS and later studies is whether the technology was viewed as a tool for providing student feedback or instructor feedback. The early studies generally used a traditional lecture format, with the technology used as a means of categorizing student responses largely for the instructor's benefit. In contrast, more recent GRS studies have examined the technology as a means for facilitating student feedback (e.g., Abrahamson 1999; Cutts et al. 2004; Dufresne et al. 1996; Nicol and Boyle 2003).

While these studies tend to find improved learning, much of the research has two potential limitations. First, the learning effects are largely self-reported. For example, Abrahamson (1999) notes that in a class employing GRS, 90 percent of respondents claimed that they understood the subject better, and a somewhat smaller percentage claimed they came to class better prepared and paid more attention. However, objective measures of learning effects are rarely provided. It is possible that student satisfaction with the technology leads to a "halo effect," which causes students to believe GRS improves their learning.

Second, it is often difficult to attribute the reported performance effects solely to GRS, since GRS deployment was often done currently with making the pedagogy more interactive, relative to a traditional lecture approach. More interactive pedagogies are associated with increased learning (e.g., Mazur 1997; Weaver and Qi 2005), and introducing the two changes together makes it difficult to determine the individual effects of GRS.

For example, one of the few GRS studies cited as using objective measures of performance is Poulis et al. (1997). They find that GRS pass rates were higher in six of the seven topics covered in the course, and the standard deviation of the pass rates were also smaller, suggesting more consistent understanding of the material. However, while the GRS were deployed in conjunction with increased student discussion, it is not clear if the comparison groups used similar student discussions without GRS, or a more conventional lecture approach.

Theoretical Rationale for GRS Effects on Learning and Engagement, and Related Hypotheses

The "Conversational Framework" is an influential theory concerning how technology affects learning in higher education, developed by Laurillard (1993, 2002). Laurillard (2002) categorizes the learning and teaching activities that comprise the framework as being of one of four types. *Discursive* activities are those related to describing the participants' conceptions of the topic goal. *Adaptive* activities are those related to actions by the participants in light of the descriptions of the task goals. *Interactive* activities are those involved in setting the task goal; planning and actions to achieve the task goal; and creating and responding to feedback about the task goal. Finally, *reflective* activities consider the interaction leading to new conceptions about the topic goal. As summarized by Laurillard (2002, 71), an effective learning dialogue allows the instructor and student to understand each other's "conceptions [of material], and the variations between them, and these in turn will determine the focus for further dialogue."

We hypothesize that GRS will particularly improve *interactive activities* and thus improve learning for the following reasons. First, GRS enhance the information provided to students, both by providing immediate feedback on their understanding and enabling them to compare their understanding to that of their classmates. As noted by Bangert-Drowns et al. (1991, 213), "researchers for a long time have advocated and verified the importance of feedback to learning." Second, the *immediacy* aspect of the GRS feedback is also important, with a meta-analysis by Kulik and Kulik (1988) showing that studies using actual classroom-based quizzes and learning materials (as opposed to experiment-based assessments and learning materials) have usually found immediate feedback to be more effective than delayed feedback.

Third, a key difference of GRS is that every student must attempt the problem and provide the answer *before* receiving feedback. This is important because Bangert-Drowns et al. (1991) note that when feedback is provided before students generate responses, the presence of correct answers can "short-circuit" learning by preventing students from practicing information retrieval, integration, or elaboration. In other words, students just copy the answer without attempting to understand. Weaver and Qi (2005) provide further support for the importance of active participation, citing a number of studies that have found an effect from students' active involvement in learning, with benefits derived from critical thinking and better retention of information. Finally, Vollmeyer and Rheinberg (2005) find that students anticipating feedback used better problem-solving strategies even before feedback was provided, leading to more knowledge acquisition over fewer attempts.

In summary, appropriate deployment of GRS should improve interactivity, by ensuring that students are actively involved in learning and have attempted the problem before receiving immediate feedback. These findings lead to our first hypothesis:

H1: Students using GRS will have improved learning relative to students learning the same material without GRS.

A benefit of having active learning and immediate feedback in the classroom is that students and instructors can improve the quality of reflective, discursive, and adaptive activities while a particular topic is still fresh. By providing timely and unambiguous feedback, particularly when the feedback indicates problems in understanding, GRS should encourage students and instructors to ask questions, and to restate learning goals and techniques in response to the feedback. Support for this view in part stems from Wong and Weiner (1981), who find that unexpected outcomes lead individuals to ask "why" questions, and Hastie (1984), who finds similarly that unexpected events lead to causal reasoning. We feel GRS create unexpected outcomes for that proportion of the students who feel they understand course material but who receive GRS feedback that their answers are incorrect. This should lead to improved engagement in the form of increased participation in the classroom, as instructors and students engage in more discussion to clarify expectations and topic materials. This leads to our next hypothesis:

H2: Students will be more engaged in classes when GRS are used relative to classes when GRS are not used.

Prior Research on Satisfaction with GRS and Its Impact on Course Satisfaction

Judson and Sawada (2002) report that nearly all studies of GRS show high levels of student satisfaction with the technology and with the course using the technology. For example, Draper and Brown (2002) report that in a survey of students in a formal logic class of about 140 students where GRS had been used, 77 percent of respondents rated the GRS as useful, very useful, or extremely useful. Abrahamson (1999) notes that 90 percent of respondents in an introductory physics class claimed to have enjoyed classes more, and the dropout rate also decreased.

We have not located any studies reporting student dissatisfaction with the technology. However, it is important to verify student satisfaction with GRS, as poor deployment, leading to student dissatisfaction, could have negative consequences for our earlier hypotheses. In addition, prior research has not generally used a control group to assess the significance of self-reported assessments of the course. We expect a halo effect from enjoyment of the technology to lead to more positive course perceptions. However, our use of a control group enables us to better determine if differences in course satisfaction can be associated specifically with GRS use, rather than arising from use of a more interactive pedagogy. We thus replicate and extend prior research in this area with our last hypotheses:

H3: Students will be more satisfied with a course when the GRS is in use, than when it is not.

H4: Students using a GRS will be satisfied with the technology.

METHOD

The research was conducted in an introductory management accounting course taught by one of the co-investigators; the course had four sections. Three sections had about 40 or fewer students each, while the other had 72 students. We use surveys conducted both in the middle and at the end of the term to evaluate student satisfaction with GRS and the course, as well as self-reported learning and engagement effects. We also create more objective measures of learning and engagement using performance on midterm exams and oral participation over the term. Details on our research design, measures, and participants follow.

Research Design

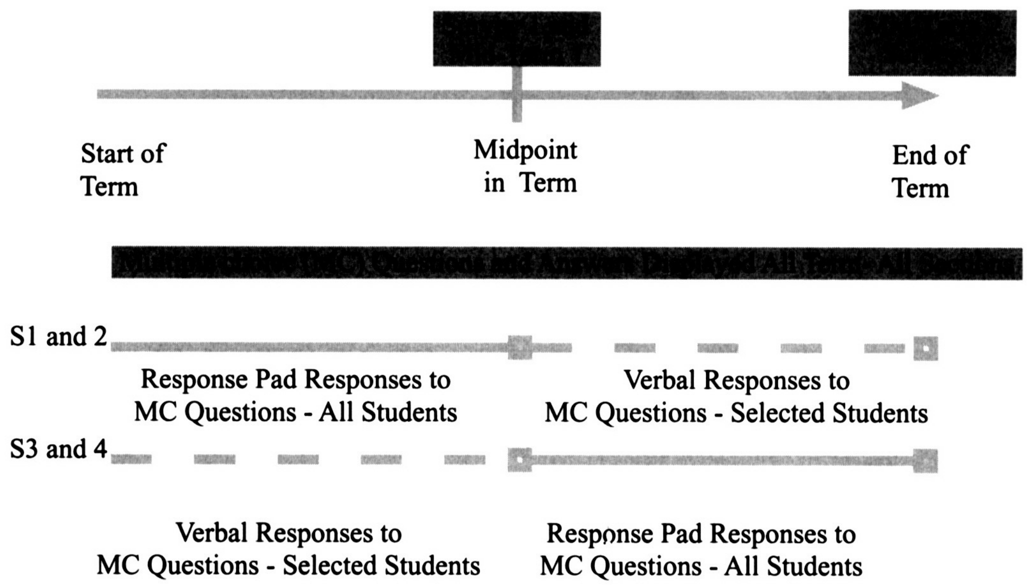
Figure 1 summarizes our within-subjects research design, counterbalancing for order of GRS usage. The within-subjects factor, GRS usage, is manipulated by having *all* students use the response pads for approximately one-half of the course, with the remaining portion of the course completed without the response pads. As shown in Figure 1, we counterbalance the order of GRS usage by having sections 1 and 2 (3 and 4) use the GRS for the first (second) portion of the term. We test for, and find no significant section effects and therefore do not discuss this factor further. Each GRS usage period was roughly five weeks in length, which makes it unlikely that novelty effects alone would lead to greater satisfaction with the GRS. Our observation of student reactions to the response pads indicated that once the initial novelty subsided, students seemed to accept use of the GRS as a normal part of the in-class routine.

Manipulating GRS usage on a within-subjects basis permits analysis of its impact while controlling for individual differences (e.g., interest in the subject matter) that might influence outcomes. Because we were unable to randomly assign students to sections, this is an important feature of our design. Counterbalancing order permits us to assess whether the order in which the GRS is used (first versus second portion of the term) has any impact on its effects. Potential factors contributing to order effects include student fatigue and other timing-related issues such as differing degrees of topic difficulty.

Implementation Procedures

Each class was conducted using an approach modeled after an interactive pedagogy known as peer instruction (Crouch and Mazur 2001). We adopted this approach for two reasons. First, we wanted to use our within-subjects design to address the issue noted earlier in prior research (e.g., Cutts et al. 2004; Dufresne et al. 1996) where GRS usage and a

FIGURE 1
Research Design



change in pedagogical approach were implemented simultaneously, making it difficult to isolate the effects of the technology itself. Second, those universities and instructors debating whether to invest in GRS technology may be interested in whether GRS provides incremental learning benefits beyond those provided by implementing an interactive pedagogy, which requires little expenditure.

Students were encouraged to ask questions, discuss issues, and answer questions posed by the instructor on assigned readings and problems. Several short-answer assigned problems (qualitative and quantitative) were covered each class and students led the discussion of each. In addition, the instructor prepared four to six multiple-choice questions on various aspects of the assigned material. The number of questions used in each class session was based on prior studies. Covering too many questions results in a student perception that the technology is being over-emphasized, and once time is allowed for discussion, covering even four to six questions in an 80-minute class takes a significant amount of time.

For *all* sections, each multiple-choice question was displayed after the related material had been covered. Exposure of *all* sections to the *same* multiple-choice questions ensures that any effects we find are not driven simply by increased familiarity with certain questions or differences in opportunities to think about certain aspects of the material. Interspersing the questions between topics resulted in a relatively even distribution of the questions across the 80-minute class. Consistent with the interactive pedagogy, for *all* sections, students were encouraged to discuss the questions with each other before answering. Students were also free to use their textbooks or notes.

Each response pad has a numeric identifier and was assigned to a particular student for the duration of usage. For the sections using GRS, all students used their response pads to answer the displayed question, with a response pad ID grid at the bottom of the data projector display indicating which pad IDs had responses recorded. A histogram of aggregated responses was then displayed, with the correct response highlighted.

In the GRS-using sections, if a significant number of students responded incorrectly, a volunteer was asked to explain his or her response, followed by further discussion. For the non-GRS-using sections, the instructor asked for a volunteer to answer the displayed question. If the student volunteer got the answer wrong, then the instructor polled additional students until either the correct answer was provided or the level of confusion indicated significant student difficulties. The instructor would then display the correct answer, followed by further discussion of the correct approach to the problem.

The interactive pedagogical approach for all in-class activities was nearly identical across all sections. However, unavoidably, the sections using GRS were somewhat more interactive with respect to the four to six multiple-choice questions covered each class since the instructor could use the response histogram feedback as a basis for stimulating further discussion of a topic. No such feedback was available in the sections not using GRS, so the degree to which further discussion occurred was largely dependent on the answer given by the volunteer and any questions raised by other students related to that answer.

Thus, the *only* differences between the condition using the GRS and that of not using the GRS are related solely to characteristics of the technology: the means by which students could respond to these questions (i.e., response pad use); whether all students had attempted the question before receiving feedback (as indicated by the response pad ID grid); whether feedback was provided about the percentage of students who chose the right answer (i.e., the histogram); and any changes in interactivity associated with the feedback. Consequently, the "treatment effect," if any, can be attributed to GRS-related effects, rather than fundamental differences in the pedagogical approaches employed during the periods when the GRS was in use versus not in use.

Dependent Variables

Objective Measures of Learning Effects

To evaluate H1, we use student performance on two midterm examinations, each of which comprised both multiple choice and short-answer questions. We develop two measures of learning effects corresponding to performance on exam questions most closely related to the in-class multiple choice questions eligible for coverage with GRS (*Related Multiple-Choice*); and performance on all other questions (*All Other*). We examine performance on the two midterm examinations separately to ensure that effects hold consistently for the term. If GRS usage improves learning, then, on average, both of these measures should be higher in the portion of the course when the GRS was used.

Related Multiple-Choice questions are defined as those that are essentially the same (only parameter values are changed) as the multiple-choice questions covered in *all* sections (with and without GRS in use). Our measure is the number of points achieved on the *Related Multiple-Choice,* with a maximum of 14 points possible on each midterm for eight separate questions.

All Other questions is measured in an analogous way, with a maximum of 86 points possible on each midterm, and captures any potential overall effects of GRS usage on learning, potentially because of differences in student efforts to prepare for class or changes in student engagement.

Objectives Measures of Engagement

To evaluate H2, we use instances of in-class participation as our proxy for student engagement. The data were collected by a teaching assistant (TA), who recorded attendance for every class and counted the number of questions asked and answered by each student. To ensure reliability in counting, students had name cards and were assigned to particular seats, and the TA used a seating chart to track participation by person.

We use participation count data for the number of questions asked (*Ask*) and the number of question answered (*Answer*) to analyze the effects of GRS on student engagement. If the GRS increases participation levels, then *Ask* and *Answer* should be higher in the portion of the course when the GRS was used than when it was not.

GRS Implementation Quality and Course Satisfaction and Subjective Measures of Learning and Engagement

Surveys were used to capture a variety of self-reported measures related to students' satisfaction with the technology and with the course, learning effects, and engagement effects. A survey about general course perceptions and students' preparation for class was administered to all students in each section immediately after the GRS were switched to the other sections (around the middle of term) and again at the end-of-term. Questions specific to GRS use were asked only on the surveys administered to the sections that had just finished using the GRS.

The questions were modeled on those used in prior studies of GRS to improve comparability as well as to capture multiple aspects of potential GRS interactivity effects (e.g., Abrahamson 1999; Draper and Brown 2004). Each question used a nine-point Likert scale centered on 0, with the end points labeled "strongly agree" (4) and "strongly disagree" (−4), with the mid-point (0) labeled "neutral." To encourage honest responses, anonymity was assured. However, students did record a unique identifier (known to them but not the instructor) on each of the two surveys to permit a within-subjects analysis of their responses for those questions asked on both surveys.

Participants and Incentives

A total of 186 second-year undergraduate students were enrolled in the course. All students were in an accounting co-op honors program, resulting in a relatively homogeneous group with respect to ability and academic background. The average age was about 20 years and 70 percent of the class was female. Students were told that the response pads were being evaluated on a trial basis for use in future years and in other accounting courses.

To provide an incentive to bring the response pad to class and to think about the responses when using it, 5 percent of the course mark in each section was based on GRS usage. In keeping with the interactive pedagogical approach employed, an additional 5 percent of the course mark was based on oral class participation (either asking or answering content related questions).

RESULTS

Impact of GRS on Learning

Hypothesis 1 predicts that GRS usage will have a positive impact on students' learning. We test our prediction using both the subjective measures of GRS learning effects reported by students and a between-subjects analysis of exam performance. Results for students' self-reported perceptions of GRS learning effects are summarized in Table 1. The overall response rate for the two surveys is high (92 percent) because they were conducted in class. The number of responses (n = 171 or 172) is less than 186 because some students were absent on the days the survey was administered or did not answer a question.

Items 1 and 2 indicate that students felt GRS usage helped them *learn the material* and the summarized class answers (histograms) helped them *track their progress* (mean

TABLE 1
GRS Effects on Self-Reported Measures of Student Learning

Item[a]	Survey Question	n	Mean[b]	Standard Deviation
1.	Response pads help in learning material[c]	171	1.75***	1.70
2.	Summarized class answers help track progress	171	1.53***	1.64
3.	Response pads encouraged working harder to answer questions	172	1.25***	1.80
4.	Response pads encouraged working harder to prepare for class	171	0.82***	1.72

[a] These questions appeared only on the survey conducted at the middle of the term (end-of-term) for students who used the GRS first (second). For these items, the number of responses is fewer than 186 (total students enrolled in course) because of absentees on the two days the surveys were conducted. Questions with fewer than 172 responses had data missing.
[b] Means were compared (one-tailed t-tests) to the scale mid-point of 0; p-values are reported as follows: ***, **, * $p < 0.001$, $p < 0.01$, and $p < 0.05$, respectively. One-tailed tests are appropriate given the directional nature of our prediction.
[c] Question was stated in the negative form on the survey.

responses of 1.75 and 1.53, respectively). Responses to questions regarding related learning behaviors indicated that students also felt the GRS encouraged working harder to answer questions and to prepare for class (with means of 1.25 and 0.82, respectively). All means in Table 1 are significantly greater than the scale mid-point of 0 ($p < 0.001$), which we interpret as evidence that students perceived the GRS to have a positive impact on learning.[4]

Table 2 reports the analyses of our objective measures of learning effects using exam performance as the proxy. We separately report performance for the *GRS-Related Multiple-Choice* questions and *All Other* questions on the exam. The pattern of means for each measure on each exam indicates that students who had been using the GRS for that portion of the term outperformed students who had not been using the technology. We use t-tests to compare the performance of students who had been using the GRS to those who had not, for each exam. The only significant difference is on the GRS related multiple-choice questions on the second exam; students who had used the GRS scored significantly higher ($p < .06$) than those who had not (means; 11 out of 14, 10.5 out of 14, respectively).[5]

Overall, support for H1 is limited. While students' self-reported assessment of the impact of GRS on learning and learning behaviors suggests a significant positive effect, learning, when proxied by exam performance, indicates GRS effects are limited to exam questions similar to those employed when using the system in class. Since all students saw the same questions during lectures, the effects observed appear to be associated with the GRS, rather than greater familiarity by the GRS users with the questions.

Impact of GRS Usage on Engagement

Hypothesis 2 predicts that students will be more engaged when GRS are used. Table 3 summarizes the analyses used to test this hypothesis for students' self-reported measures

[4] All p-values reported for directional predictions are one-tailed. We test for order effects (GRS used first versus second) in all of our analyses. However, to simplify the results presentation, we only report *Order* (descriptive statistics and results of statistical tests) when we find significant effects involving order of GRS usage.
[5] We also conducted paired t-tests comparing, on a within-subjects basis, performance with and without the GRS. Results (not tabulated) indicate students performed better on the *GRS-related multiple-choice* questions when they had been using the GRS compared to when they had not ($p < 0.05$). No significant difference was found for the *All Other* questions measure.

TABLE 2
Analysis of GRS Effects on Objective Measures of Student Learning

Exam Results (n = 183):[a]

Performance Measure	GRS in Use		GRS Not in Use			
	Mean	Standard Deviation	Mean	Standard Deviation	t-statistic[c]	p
Exam One[b]						
Related multiple-choice (out of 14)	9.7	2.2	9.4	2.4	1.00	0.159
All other questions (out 86)	64.3	12.3	63.4	12.2	0.46	0.322
Exam Two						
Related multiple-choice (out of 14)	11.0	2.1	10.5	1.9	1.56	0.060
All other questions (out of 86)	73.5	8.7	72.7	7.3	0.62	0.266

[a] The number of students used in the analysis is 183 because three students missed at least one of the two mid-terms.
[b] *Related multiple-choice* questions were similar to the multiple choice questions covered in class using the GRS. *All other questions* covered topics different from those included in the *Related multiple-choice* questions.
[c] For each exam, t-tests (one-tailed) are used to compare the exam scores for students who had used the GRS with those who had not used the GRS for that portion of the course.

of participation as one proxy for engagement. Students responded to these survey questions immediately after using the GRS and again after the period when they had not been using the GRS. Panel A of Table 3 provides descriptive statistics on students' subjective ratings of comfort participating in class, averaged across the responses to the three questions used in our survey.[6] Overall, collapsing across the two order conditions, students were more comfortable participating when the GRS was in use (mean 1.21) than when it was not (1.07). However, examination of the four cell means reported in Table 3 (Panel A) indicates the pattern of results is different for students who used the GRS first compared to those who used it second.

Panel B of Table 3 reports the results of the repeated measures ANOVA used to test H2. *GRS* usage is the within-subjects factor and *Order* of usage is the between-subjects variable. Results show not only a significant main effect for *GRS* usage ($p < 0.05$), but also a significant interaction between *GRS* usage and *Order* ($p < 0.001$). To interpret the interaction we test the simple effects within each order condition. Students who used the GRS first were more comfortable participating with the GRS in use (mean 1.37) compared to when it was not (mean 0.66) and this difference is significant ($p < 0.001$). Students who used the GRS second were equally comfortable participating with (mean 1.11) and without (mean 1.34) the technology in place ($p > 0.15$). Thus students who had the GRS taken away felt significantly less comfortable participating in the absence of the technology.

Panels A and B of Table 4 report descriptive statistics for our two objective measures of student participation. In Panel A, *Ask* (*Answer*) represents the average number of questions asked per student (answered) pooled across the 11 classes when the GRS was in use, versus the same measure when the GRS was not in use. We report results in Table 4 only for those students who asked (n = 84) or answered (n = 160) at least one question at *any* time during the 22 classes. This enables us to separately consider GRS effects on students

[6] We asked students to rate their comfort for the following items: *participating in class*, *asking questions*, and *answering questions*. Because of the high inter-item correlations among the three measures ($r = .41$ or more, p-values < 0.001) we average them to simplify reporting.

TABLE 3
Analysis of GRS Effects on Self-Reported Student Participation in Class

Panel A: Comparison of Student Views on Participation Collected in the GRS versus Non-GRS Portion of Term: Descriptive Statistics (n = 128)

	GRS in Use		GRS Not in Use		
Average Comfort Participating[a]	**Mean**	**Standard Deviation**	**Mean**	**Standard Deviation**	**Overall**
Used GRS first (n = 51)[b]	1.37	1.50	0.66	1.56	1.01
Used GRS second (n = 77)[b]	1.11	1.40	1.34	1.44	1.22
Overall	1.21	1.44	1.07	1.52	

Panel B: Repeated Measures ANOVA, Dependent Variable: *Average Comfort Participating*

Source	**Sum of Squares**	**df.**	**Mean Square**	**F**	**p**
GRS[c]	3.562	1	3.562	4.278	.020[d]
GRS × *Order*[e]	13.451	1	13.451	16.155	.001
Error	104.910	126	0.833		
Between-Subjects					
Order	2.688	1	2.688	0.777	.380
Error	435.749	126	3.458		

[a] Three items were used to evaluate students' comfort participating and appeared on the surveys conducted both at the middle of the term and at the end of the term. The items asked students to self-assess their comfort: *participating*, *answering questions*, and *asking questions*. Given the high inter-item correlations (all Pearson correlations > .41, $p < 0.001$) among the three items, to simplify the presentation we report the average of the individual values in Panel A and use it as the dependent variable in our Panel B ANOVA.
[b] Used GRS first (second) refers to the group of students who used the GRS during the first (second) part of the term. The total number of students (n = 128) is less than 172 (Table 1) because: (1) some students failed to record the same unique identifier on each survey and (2) some students missed one of the classes in which the surveys were conducted.
[c] *GRS*: within-subjects factor measured twice, GRS in use and GRS not in use.
[d] One-tailed test.
[e] *Order*: indicator of whether the student used the GRS during the first or second portion of the course.

who are inclined to participate, while in Table 5 we focus on GRS effects on the percentage of students participating in each class. The proportion of the total class who asked (45 percent) or answered (86 percent) at least one question during the term is approximately the same across all four sections of the course. Results (not tabulated) using *all* students, including those who never asked or answered a question, are qualitatively similar to those reported in Table 4.

The descriptive results show that, on average, students each asked (answered) 2.3 (5.9) questions over the 11 classes when the GRS was in use compared to 3.2 each (6.2) over the 11 classes when the GRS was not in use. While the participation activity results appear small on an individual student basis, they are more meaningful when aggregated for all students on a total participation per class basis. For example, the average of 6.2 questions answered for the 11 classes where students did not use the GRS translates into approximately 20 questions *per class* (6.2/11 × 35) being answered by students in a section with 35 participating students.

TABLE 4
Analysis of GRS Effects on Objective Measures of Student Participation in Class

Panel A: Average Number of Questions Asked per Student for 11 Classes Held under Each Condition: Descriptive Statistics (n = 84)

	GRS in Use		GRS Not in Use		
Participation Measure: *Ask*[a]	**Mean**	**Standard Deviation**	**Mean**	**Standard Deviation**	**Overall**
Used GRS first	2.79	4.46	3.33	5.40	3.06
Used GRS second	1.94	2.64	3.11	4.31	2.53
Overall	2.27	3.47	3.20	4.74	

Panel B: Average Number of Questions Answered per Student for 11 Classes Held under Each Condition: Descriptive Statistics (n = 160)

	GRS in Use		GRS Not in Use		
Participation Measure: *Answer*[a]	**Mean**	**Standard Deviation**	**Mean**	**Standard Deviation**	**Overall**
Used GRS first	8.18	8.81	5.85	7.50	7.01
Used GRS second	4.25	4.88	6.45	5.84	5.35
Overall	5.90	7.06	6.20	6.57	

Panel C: Generalized Negative Binomial Regression with *Ask* as Dependent Variable (n = 168; 84 × 2 observations per student)

	Estimate	**Standard Error**	**Z-value**	**p**
Constant	0.54	0.21	2.50	0.01
GRS[b]	−0.48	0.13	−3.68	0.001[c]
Order[d]	−0.05	0.38	−0.13	0.90
GRS × *Order*	0.30	0.18	1.66	0.10

Panel D: Generalized Negative Binomial Regression with *Answer* as Dependent Variable (n = 320; 160 × 2 observations per student)

	Estimate	**Standard Error**	**Z-value**	**p**
Constant	1.86	0.09	19.90	0.001
GRS	−0.42	0.08	−5.49	0.001[c]
Order	−0.10	0.18	−0.54	0.60
GRS × *Order*	0.75	0.11	6.74	0.001

[a] *Ask* (*Answer*) represents the average number of questions asked (answered) by students pooled across the 11 classes when the GRS was in use and then across the 11 classes when the GRS was not in use. Only students who asked or answered at least one question at any point during the term are included in the analysis.
[b] *GRS*: coded 1 (0) if the GRS was in use (not in use).
[c] One-tailed test.
[d] *Order*: indicator of whether the student used the GRS during the first or second portion of the course.

We use generalized negative binomial regressions (Cameron and Trivedi 1998) to test the effects of GRS usage on the counts of the numbers of questions asked (answered) by each student. Negative binomial regressions are reasonable alternatives for estimating models of count data when goodness-of-fit tests indicate Poisson regressions are inappropriate because of over-dispersion, as is our case. The estimated standard errors were adjusted for intragroup correlation, as our sample has two observations per student (one when the student was using GRS, and one without). These results are reported in Panels C and D, Table 4.

Contrary to H2, the significant negative coefficients ($p < 0.001$) on *GRS* usage indicates that use of the technology *reduced* the likelihood that students would ask or answer a question in class. There are significant positive *GRS* × *Order* interactions affecting both asking and answering questions. Cell means reported in Panel A of Table 4 indicate that students who used the GRS first asked an average of 2.79 questions over the 11 classes of use, compared to an average of 3.33 questions when the GRS were not in use, while those using the GRS second asked 1.94 questions when the GRS were in use, versus 3.11 when the technology was not in use. The negative effect was thus greater for those who used it second. Cell means reported in Panel B indicate that students who used the GRS first answered an average of 8.18 questions over the 11 classes of use, compared to an average of 5.85 questions when the GRS were not in use, while those using the GRS second answered 4.25 questions when the GRS were in use, versus 6.45 when the technology was not in use. Thus, while the GRS increased the average number of questions for those using the technology first, it decreased the average number of questions for those using GRS second. This pattern holds when medians are considered rather than means. We conclude that despite the mixed findings reported in Panel B, the overall results reported in Table 4 suggests the effect of GRS on participation is negative.

To further examine the effects of GRS usage on participation, we use the average percentage of students asking questions ($Ask_{percent}$) in each class as the dependent variable. This shifts the focus of the analysis to class-level effects of GRS usage and also allows us to control for the number of questions the instructor asked each class, which may have impacted students' willingness to participate.[7] We use an ANOVA with *GRS* usage as the between-subjects factor, and questions posed by the instructor (*Instructor Questions*) as a covariate. Results reported in Table 5 (Panel B) show a significant effect of *GRS* usage on $Ask_{percent}$ ($p < 0.05$), with the means in Panel A showing a smaller percentage of the class asked questions when the GRS was used (7.6 percent with GRS; 10.0 percent without GRS). An equivalent analysis (not tabulated) of the percentage of students answering questions showed no significant effects for GRS usage. Overall the results reported in Tables 4 and 5 indicate that GRS usage reduced the extent to which individual students asked and answered questions in class, and on average reduced the percentage of students in attendance willing to raise a question.

To provide further insight on this issue, we also examined on a *post hoc* basis whether the effect of GRS usage on Ask_{zip} is related to the difficulty of the multiple choice questions used in-class. On average, students (all sections combined) scored 84 percent on these questions, which means the GRS histograms showed a large majority of the class had responded correctly. Therefore, students may have been less likely to ask questions when feedback indicated most of the class understood the concept. Since similar feedback was

[7] Because *Ask* is based on the average number of questions asked per student for an entire portion of the course (with or without the GRS in use), the number of questions asked per class by the instructor cannot be used as a covariate in the analysis reported in Table 4.

TABLE 5
Analysis of GRS Effects on the Percentage of Students Participating in Class

Panel A: Descriptive Statistics

	GRS in Use		GRS Not in Use		
Participation Measure[a]	**Mean**	**Standard Deviation**	**Mean**	**Standard Deviation**	**Mean Difference**
Ask Percentage	7.6%	4.5	10.0%	4.9	−2.4

Panel B: ANOVA, Dependent Variable: *Ask* Percentage (n = 88[b])

Source	Type III Sum of Squares	df	Mean Square	F	p
GRS[c]	120.402	1	120.402	5.303	.024[d]
Instructor Questions[e]	7.436	1	7.436	.328	.569
Error	1929.827	85	22.704		

[a] The percentage of students in attendance in a section that asked a question that class.
[b] A separate observation for each class in each section for the entire term is included in the analysis: 4 sections × 22 classes = 88 independent observations. There were 24 classes in the term, but the first and last classes were dropped as they consisted primarily of introductory and review material respectively.
[c] *GRS*: a dummy variable indicating whether the GRS was in use for the class.
[d] One-tailed test.
[e] *Instructor Questions*: the number of questions asked by the professor during each class for each section. The same teaching assistant who tracked oral participation also recorded the number of questions the instructor asked each class.

unavailable when the GRS was not used, students would be more uncertain about the extent to which their peers understood a concept and perhaps more willing to raise questions. To evaluate this possibility, a correlation was calculated between the percentage of students asking a question in each class the GRS was used, and the overall score for that class on the multiple-choice questions. The correlation coefficient is negative (−.28) and marginally significant ($p < 0.10$), indicating the more difficult the questions, the greater the percentage of students asking questions. Thus, it appears that a GRS may discourage discussion in classes where the feedback from the system indicates the majority of students understand the concepts being reviewed.

Impact of GRS on Students' General Course Perceptions

Hypothesis 3 predicts that students using GRS will have more positive general perceptions about the course. Table 6 reports descriptive results and the analysis used to test this hypothesis. Our results are based on the responses to the surveys administered after *each* of the first and second halves of the course. Descriptive statistics for differences in ratings for the GRS versus non-GRS period of usage are reported in Table 6, Panel A. To simplify the presentation, we report only results for those questions where we find a significant effect for GRS usage.[8]

[8] We also asked students for their perceptions of: the course organization, the presentation of the materials in class, the value of classes in helping them master the material, and how easy it was for them to pay attention in class. Results for these measures (not tabulated) show no significant main effects for *GRS* usage and the interaction between *GRS* usage and *Order* is not significant.

TABLE 6
Analysis of GRS Effects on Students' General Course Perceptions

Panel A: Comparison of Student General Course Perceptions Collected in the GRS versus Non-GRS Portion of Term—Descriptive Statistics (n = 128)

	GRS in Use		GRS Not in Use		
	Mean	Standard Deviation	Mean	Standard Deviation	Overall
Course Interesting[a]					
Used GRS first (n = 51)[b]	2.01	1.41	1.44	1.18	1.72
Used GRS second (n = 77)[b]	1.94	1.43	2.15	1.13	2.04
Overall	1.97	1.42	1.87	1.19	
Think in class about course concepts					
Used GRS first (n = 51)	2.46	1.32	1.53	1.69	1.99
Used GRS second (n = 77)	1.89	1.47	2.34	1.28	2.11
Overall	2.12	1.44	2.02	1.50	

Panel B: Repeated Measures ANOVA, Dependent Variable: *Course Interesting*

Within Subjects	Sum of Squares	df	Mean Square	F	p
GRS[c]	1.997	1	1.997	2.127	.074[d]
GRS × *Order*[e]	9.247	1	9.247	9.846	.002
Error	118.343	126	0.939		
Between Subjects					
Order	6.282	1	6.282	2.614	.108
Error	302.746	126	2.403		

Panel C: Repeated Measures ANOVA, Dependent Variable: *Think in class about course concepts*

Source	Sum of Squares	df.	Mean Square	F	p
GRS	3.488	1	3.488	2.270	.067[d]
GRS × *Order*	29.464	1	29.464	19.181	.001
Error	193.550	126	1.204		
Between Subjects					
Order	0.910	1	0.910	0.353	.553
Error	324.519	126	2.586		

[a] These questions appeared on both the surveys conducted at the middle of the term and at the end of the term.
[b] Used GRS first (second) refers to the group of students who used the GRS during the first (second) part of the term. The total number of students (n = 128) is less than 172 (Table 1) because: (1) some students failed to record the same unique identifier on each survey and (2) some students missed one of the classes in which the surveys were conducted.
[c] *GRS*: the within-subjects factor measured twice, GRS in use and GRS not in use.
[d] One-tailed test.
[e] *Order*: indicator of whether the student used the GRS during the first or second portion of the course.

Panel B of Table 6 indicates a significant main effect for *GRS* usage ($p < 0.10$) and a significant interaction between *GRS* usage and *Order* ($p < 0.005$) with the extent to which students found the *course interesting* as the dependent variable. Testing simple effects using the cell means reported in Panel A shows students who used the GRS first found the portion of the course with the GRS (mean 2.01) significantly ($p < 0.005$) more interesting than the portion of the course without the GRS (mean 1.44). Students who used the GRS second found both portions of the course equally *interesting* (means: with GRS 1.94; without GRS 2.15).

Results using students' beliefs about how much they were required to *think in class about course concepts* as the dependent variable are reported in Panel C, Table 6. Again we find a significant main effect for *GRS* usage ($p < 0.10$) and a significant interaction between *GRS* usage and *Order* ($p < 0.001$). Testing simple effects using the cell means in Panel A shows that students who used the GRS first believed it significantly ($p < 0.001$) increased the amount they were required to *think in-class* (means: with GRS 2.46, without GRS 1.53). However, students who used the GRS second indicated they were required to *think in-class* significantly less ($p < 0.05$, not tabulated) when the system was in use (means: with GRS 1.89, without GRS 2.34).

Overall, we find little support for H3. We find significant main effects for *GRS* usage for two measures of general course perceptions. However, analysis of the significant *GRS* × *Order* interaction indicates the effect of *GRS* usage on students' *course interest* is attributable to students who used the system for a period of time and then had it removed. Further, the extent to which GRS usage impacts students' perceptions about how much they were required to *think in-class* depends on whether the technology was deployed for the first or second part of the class. Overall, we conclude that the GRS has limited incremental impact on improving course satisfaction.

Student Satisfaction with GRS

Hypothesis 4 predicts that students will have positive perceptions of the GRS technology. Responses to the measures designed to evaluate students' reactions to our implementation of the GRS technology are summarized in Table 7. These questions appeared on the mid-term (end-of-term) survey only for students who used the GRS during the first (second) part of the term. The results suggest, as predicted, students reacted favorably to the GRS. The strongest level of agreement was for Items 1 and 2, indicating they believed the response pads were *easy to use* and that the *instructor clarified the correct solution for each response pad question* (means of 3.29 and 2.95, respectively). For all measures in Table 7, the mean is significantly greater than the scale mid-point of 0 (all p-values < 0.001), which we interpret as evidence supporting H4.[9] We also conclude from these results that our failure to find support for some of our other predictions is not attributable to a poor implementation of the GRS technology in the course.

CONCLUSIONS AND LIMITATIONS

A review of vendor websites (e.g., those of eInstruction and GTCO CalComp) and recent offerings of GRS optionally bundled with textbooks by publishers such as McGraw-Hill suggests that GRS technology is becoming a popular tool in university education. The inclusions of GRS-related papers in recent AAA Annual Meeting forums (i.e., Segovia 2004; Tietz 2005) suggest accounting educators are also interested in this topic.

The intent of our research is to begin to evaluate the claims concerning the effects of GRS on student course satisfaction, learning, and engagement, and thus help to inform the

[9] To provide a more stringent test of students' perceptions of GRS, we also compared the average responses for each measures to 1. All p-values < 0.01.

TABLE 7
Student Perceptions of GRS Implementation

Item	Survey Question Number/Description[a,b]	n[c]	Mean[d]	Standard Deviation
1.	Response pads are easy to use[e]	172	3.29***	1.08
2.	Instructor clarified correct solution for response pad questions	171	2.95***	1.18
3.	Lecture and response pads effectively integrated	172	2.36***	1.36
4.	Enjoy using the response pads	171	2.35***	1.54
5.	Advantages of response pads outweigh disadvantages	170	2.25***	1.37
6.	Response pads should be used in other courses[e]	170	1.99***	1.73
7.	Course does not focus too much on using response pads[e]	172	1.96***	1.60
8.	Confident that response pads accurately record responses	171	1.91***	1.97
9.	Enough time to answer questions using response pads	171	1.66***	1.81

[a] All questions required students to agree with the statement using a nine-point Likert scale centered on 0 with end-points labeled "strongly agree" (4) and "strongly disagree" (–4), with the mid-point labeled "neutral." All survey measures reported in subsequent tables used this same scale.
[b] These questions appeared only on the survey conducted at the middle of the term (end of the term) for the students who used the GRS first (second). The order of GRS use factor has been collapsed within the reported values.
[c] The number of responses is fewer than 186 (total students enrolled in course) because of absentees on the two days the surveys were conducted. Questions with fewer than 172 responses had data missing.
[d] Means were compared (one-tailed t-tests) to the scale mid-point of 0; p-values are reported as follows: ***, **, * $p < 0.001$, $p < 0.01$, and $p < 0.05$, respectively. One-tailed tests are appropriate given the directional nature of our predictions.
[e] Question was stated in the negative form on the survey.

decisions of those considering this technology. Our results suggest current claims of GRS effects on learning should be viewed cautiously.

Consistent with prior research, average student responses to our survey questions suggest strong student satisfaction with the technology. Also consistent with prior research, we find that students claim positive effects of GRS on learning. However, our extensions of prior studies to include more objective measures of learning, as proxied by improvements in exam performance, suggest that learning effects are limited. We find positive learning effects *only* for the exam questions most similar to the in-class questions eligible for coverage with the GRS. Our objective measures of student participation, as proxies for student engagement, also suggest that students tend to ask fewer questions when GRS is in use, although we find mixed effects of GRS on students answering questions. The positive student response to the GRS technology reduces the possibility that the relatively weak GRS effects on student engagement and learning arose from a poor implementation of the technology.

Our extension of prior research to include a control group resulted in a finding of limited changes in students' general course satisfaction as a result of GRS usage. In those cases where student perceptions of the course did change, it often seemed related to a decline in ratings when the technology was taken away, rather than an improved perception from providing the technology midway through the semester. These findings are similar to those of decreased comfort participating when the GRS was removed for students who used the technology first. It could be that the GRS creates an interesting and comfortable participation environment for students that they are most aware of when the system is gone.

Overall, we conclude that student reports of enjoyment of the GRS technology do not necessarily mean a course is generally regarded more favorably.

Finally, our control for changes in pedagogy while implementing GRS help us disentangle technology effects from other learning effects obtainable through changing pedagogy alone. While this creates a challenging benchmark to use in evaluating the effects of GRS, it seems appropriate so that learning improvements achievable through other means are not attributed solely to GRS.

We acknowledge that this study provides only preliminary conclusions about the effects of GRS on student learning, engagement, and satisfaction. Like all studies, ours has limitations and provides various opportunities for further investigation. First, social desirability issues may have influenced students' responses to some of our survey questions. For example, students (with or without the GRS) may not have been willing to acknowledge any difficulty *paying attention in class* or express concerns about the degree to which *course material was presented effectively*. While these effects are inherent to the survey method, in our case they bias against finding support for our predictions. Second, it is possible that demand effects could have influenced some of our results. For example, students may have indicated being satisfied with the technology on the surveys because they believed this was the *desired* response.

Finally, our research design does not permit us to make inferences about the effectiveness of GRS usage on topics of differing levels of difficulty since we asked only one or two questions per topic on each exam. Further research is needed to examine whether the difficulty of the material has any impact on the relation between exam performance and GRS usage. Additional research is also needed to further explore if GRS can affect learning outcomes. Perhaps some pedagogical characteristic, other than interactivity, would be more suitable in enabling a GRS effect. Examining the interplay between different types of pedagogy and GRS may be a useful area for further study. Further research is needed before arriving at firmer conclusions on the effects of GRS on student satisfaction and learning in accounting education.

REFERENCES

Accounting Education Using Computers and Multimedia (AECM). 2005. Archives of discussions. Available at: http://pacioli.loyola.edu/aecm. (Use keyword "clickers.")

Abrahamson, A. 1999. Teaching with classroom communication systems: What it involves and why it works. Paper presented at the 7th International Workshop on New Trends in Physics Teaching, Puebla, Mexico, May.

Bangert-Drowns, R. L., C. C. Kulik, J. A. Kulik, and M. T. Morgan. 1991. The instructional effect of feedback in test-like events. *Review of Educational Research* 61 (2): 213–238.

Cameron, A. C., and P. K. Trivedi. 1998. *Regression Analysis of Count Data*. New York, NY: Cambridge University Press.

Crouch, C. H., and E. Mazur. 2001. Peer instruction: Ten years of experience and results. *American Journal of Physics* 69: 970–977.

Cutts, Q., A. Carbone, and K. van Haaster. 2004. Using an electronic voting system to promote active reflection on coursework feedback. In *Proceedings of the International Conference on Computers in Education 2004*, Melbourne, Australia, November 30–December 3.

Draper, S. W., and M. I. Brown. 2002. Use of the PRS (personal response system) handsets at Glasgow University. Available at: http://www.psy.gla.ac.uk/~steve/ilig/interim.html.

———, and ———. 2004. Increasing interactivity in lectures using an electronic voting system. *Journal of Computer Assisted Learning* 20: 81–94.

Dufresne, R. J., W. J. Gerace, W. J. Leonard, J. P. Mestre, and L. Wenk. 1996. Classtalk: A classroom communication system for active learning. *Journal of Computing in Higher Education* 7: 3–47.

Hastie, R. 1984. Causes and effects of causal attribution. *Journal of Personality and Social Psychology* 46: 44–56.

Judson, E., and D. Sawada. 2002. Learning from past and present: Electronic response systems in college lecture halls. *Journal of Computers in Mathematics and Science Teaching* 21 (2): 167–181.

Kulik, J. A., and C. C. Kulik. 1988. Timing of feedback and verbal learning. *Review of Educational Research* 58 (1): 79–97.

Laurillard, D. 1993. *Rethinking University Teaching: A Framework for the Effective Use of Educational Technology*. London, U.K.: Routledge.

———. 2002. *Rethinking University Teaching: A Framework for the Effective Use of Educational Technology.* Second edition. London, U.K.: Routledge.

Mazur, E. 1997. *Peer Instruction: A User's Manual.* Upper Saddle River, NJ: Prentice Hall.

Nicol, D. J., and J. T Boyle. 2003. Peer instruction versus class-wide discussion in large classes: A comparison of two interaction methods in the wired classroom. *Studies in Higher Education* 28 (4): 458–473.

Poulis, J., C. Massen, E. Robens, and M. Gilbert. 1997. Physics lecturing with audience paced feedback. Available at: http://www.bedu.com/Publications/PhysLectAPF.pdf.

Roschelle, J., W. R. Penuel, and L. Abrahamson. 2004. Classroom response and communication systems: Research review and theory. Paper presented at the Annual Meeting of the American Educational Research Association, San Diego, CA, April. Available at: http://www.ubiqcomputing.org/CATAALYST_AERA_Proposal.pdf.

Segovia, J. R. 2004. Who wants to learn accounting? The use of personal response systems in introductory accounting. Presented at the American Accounting Association Annual Meeting, Orlando, FL, August 8–11.

Tietz, W. 2005. Using audience response systems ("clickers") in introductory accounting classes. Presented at the American Accounting Association Annual Meeting, San Francisco, CA, August 7–10.

Vollmeyer, R., and F. Rheinberg. 2005. A surprising effect of feedback on learning. *Learning and Instruction* 15 (6): 589–602.

Weaver, R. R., and J. Qi. 2005. Classroom organization and participation: College students' perceptions. *The Journal of Higher Education* 76 (5): 570–601.

Wong, P. T. P., and B. Weiner. 1981. When people ask "why" questions, and the heuristics of attributional search. *Journal of Personality and Social Psychology* 40: 650–663.