Investigating the Effects of Group Response Systems on Learning Outcomes and Satisfaction in Accounting Education

Carla Carnaghan*

Alan Webb
School of Accountancy
University of Waterloo

June 2005

Copyright 2005 by Carla Carnaghan and Alan Webb
* Contact Author: e-mail: **cacarnag@uwaterloo.ca**, mailing address: School of Accountancy, University of Waterloo, Waterloo Ontario N2L 3G1 Canada

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. We thank Tony Atkinson, Irene Gordon, Duane Kennedy, Natalia Kotchetova, Bill Wright and participants at the University of Waterloo accounting research workshop, the 2005 European Accounting Congress, and the 2005 Canadian Academic Accounting Association Annual Meeting for helpful comments on earlier drafts. We further thank Tony Atkinson for allowing us access to his students for a further extension of this project. We also thank Sarah Fernandes for her excellent research assistance.

ABSTRACT

A review of vendor websites indicates numerous claims by proponents about the benefits of group response systems (GRS) on student participation, engagement and performance. GRS enable greater instructor-student interaction through individual student use of electronic response pads, with the student responses then aggregated and displayed by the instructor to provide immediate feedback on class understanding of a particular question. Despite their increasing use and prior research indicating student satisfaction with GRS, there is little empirical evidence of the effect of GRS on learning, measured directly or indirectly through participation or exam performance. This study examines the effect of GRS pedagogies on student learning and satisfaction in accounting education to address this issue.

We use a repeated measures experimental approach to examine class oral participation and performance on exams where the related classes have been conducted with and without use of the GRS. We also conduct surveys to compare within and between student perspectives on the technology and course with and without the use of the GRS. We find clear evidence of student satisfaction with GRS, but it has no effect on self-reported perceptions of the class more generally. Self-reported comfort with participation declines on a within-subjects basis once the GRS are removed, but there is no significant difference in this measure comparing GRS to non-GRS users during the period of technology usage. Analysis of objective participation measures indicates no effect of the GRS on the average number of questions answered, but a weakly significant decline in the average number of questions asked per student when the GRS was used. We also find that a smaller percentage of students participated in class when the GRS was used. We find some evidence that exam performance improvement is associated with GRS usage, but only for those exam questions most closely related to the questions displayed in class using the GRS. However, since these questions were displayed to all students, this results suggests the improved performance may be related to some GRS characteristic, rather than just to greater familiarity with the GRS displayed questions.

Our results are contrary to some of the claims being made by vendors of GRS that the technology can have dramatic effects on student engagement. We discuss possible reasons for these findings, and suggest additional research to address some possible limitations of the current study.

INTRODUCTION

Higher education is constantly under pressure to become both more effective and more efficient. One response to these pressures is deployment of information technology to improve both efficiency and effectiveness. While technology such as e-mail, course websites, and online chat rooms can be used to improve communications with large numbers of students without necessarily affecting pedagogy[1], other technologies are being experimented with specifically to improve student learning. The purpose of this study is to examine the effects of a technology called group response systems[2] (GRS) that is intended to directly improve student engagement and feedback and thus indirectly improve learning. Specifically, we explore the incremental effects of GRS, controlling for pedagogy, on both subjective measures of engagement and course satisfaction (as indicated by student survey responses) as well as objective measures based on changes in exam performance and student participation.

GRS comprise individual response pads that typically have wireless communication connections to a receiver, which is in turn connected to a computer/data projector combination. Instructors create various types of structured questions (such as multiple choice, true/false, and yes/no questions), using custom software to display them with the data projector. Each student uses his or her response pad to select an answer, which is transmitted to the receiver and recorded. The responses can then be automatically aggregated and displayed to provide immediate feedback on students' individual comprehension as well as the class's understanding.

---

[1] This statement is not intended to imply that such technologies can not be deployed with changes in pedagogy to improve learning, but rather that the intent in deployment of these technologies is often to provide another medium for communication, and thus to supplement office hours and in-class discussions.

[2] As noted by S. Draper on the website 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).

Vendors and users of GRS cite numerous benefits including improvements in student satisfaction, engagement, exam performance and interaction[3]. However, despite these claims, there is limited supporting evidence and little research on whether such benefits lead to improved learning. A number of studies suggest strong student (and instructor) satisfaction with the systems, but almost none have attempted to measure changes in student interaction or investigated objective improvements in learning incremental to the pedagogy being used (Judson and Sawada 2002; Roschelle, Penuel and Abrahamson 2004). Our study is designed to address this gap in the literature.

We expect that any learning effects of GRS will arise from the enhanced interactivity the technology provides. Learning frameworks such as those developed by Laurillard (1993; 2002) stress the importance of interactivity as one component in creating a "conversational framework", which includes student practice, instructor feedback, student reflection on the feedback, and tailoring of materials by the instructor to address the problems in understanding revealed by the student responses. We are not claiming that interactivity requires computer technology but some traditional approaches (such as regular hand-in assignments or involving the entire class in discussion) may be difficult to employ in an era of increasing class sizes and limited marking budgets. A benefit of GRS is that they may increase interactivity, regardless of class size, by having all students immediately respond to and receive feedback for every question, and by enabling instructors to focus on the problems revealed by the question responses. Consistent with Laurillard's (2002) framework, this may result in improved learning.

Our research design collects and analyses two kinds of data. We analyze student responses to surveys to determine student perceptions of the technology and the accounting

[3] Examples of these claims can be found at www.einstruction.com and www.gtco.com.

course more generally, and we collect and analyze objective measures of student participation and exam performance, using repeated measures of both items with and without the GRS system. Our study extends the GRS literature in three significant ways. First, we have a control group so that we can compare student general course perceptions with and without GRS. Second, we use an objective measure of student participation (as a proxy for engagement). Third, ours is the only study we are aware of that examines the effects of GRS on objective measures of learning incremental to pedagogy effects. The remainder of this paper is organized as follows: the next section provides a literature review summarizing what is known about the effects of GRS and our resulting research questions to extend the literature, followed by a discussion of our research design and then the analysis of our results. Finally, we summarize our findings, and provide some conclusions and possible limitations of our work.

PRIOR RESEARCH ON GROUP RESPONSE SYSTEMS AND DEVELOPMENT OF RESEARCH QUESTIONS

GRS have been used by a variety of disciplines in universities, including philosophy (logic), mathematics, statistics, engineering, architecture, physics, computing science, psychology, and medicine. Class sizes where the technology has been deployed range from 20 to 500 students per class. As noted earlier, many of the more recent studies of GRS have examined student satisfaction with the technology, with only a few investigating effects on learning as indicated by differences in exam performance.

Student Satisfaction

Judson and Sawada (2002) report that nearly all studies of GRS show high levels of student satisfaction with the technology. For example, Abrahamson (1999) provides results from

a study using GRS in an introductory physics course of 150 students. She notes that 90% of respondents claimed that they understood the subject better, enjoyed classes more, and a somewhat smaller percentage claimed they came to class better prepared and paid more attention in classes. The drop-out rate also decreased. Draper and Brown (2002) report that in a survey of students in a formal logic class of about 140 students where a GRS had been used, 77% of respondents rated the GRS as useful, very useful, or extremely useful. A further study by Draper and Brown (2004) on the results of using a GRS shows that roughly 80% of respondents in a first year computing science course reported that the benefits exceeded the disadvantages of using a GRS. Less than 5% of respondents indicated that there were more disadvantages than benefits. We have not located any studies reporting student dissatisfaction with the technology. However, the studies we have located have not generally used a control group to assess the non-GRS specific measures of satisfaction, such as class enjoyment and paying more attention in class.

Performance Effects

Judson and Sawada (2002) summarize the studies of GRS in the 1960's and 1970's which deployed the technology with traditional lecture approaches as finding that the technology was associated with no difference in student performance, despite student reports of strong satisfaction with the technology. Judson and Sawada note that a key difference between early studies of GRS and later studies is whether the technology was used to improve interactivity within the pedagogy being used. The early studies generally used a traditional lecture format, with the technology used as a means of counting and categorizing student responses and the summarized information provided largely for the instructor's benefit.

In contrast, recent GRS studies such as those done by Cutts et al (2004), Madill (2004,

Dufresne et al (1996) and Nicol and Boyle (2003) use the technology to enhance an interactive pedagogy. For example, Nicol and Boyle (2003) compare the effects of two types of interactive pedagogy when used with GRS, but did not examine differences in objective measures of performance (e.g., exam scores). One of the few studies cited as having objective measures of performance change is Poulis et al (1997). This study found that the 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. This study examined the use of GRS in conjunction with increased student discussion, although it is not clear if the comparison groups used student discussion without the GRS, or a more conventional lecture approach. Another study commonly cited as having demonstrated positive GRS learning effects is that of Hake (1998). Hake found substantially improved performance on standard physics exams by students who had been enrolled in "interactive engagement" courses relative to those enrolled in "traditional lecture" courses. However, it appears likely that many of the interactive engagement courses did not use GRS, so conclusions regarding the impact of GRS on learning based on this study are inappropriate.

The Importance of Interactivity

Judson and Sawada's conclusions along with the Hake (1998) study suggest the importance of considering pedagogy when trying to establish the learning effects of GRS. In particular, these and other GRS studies suggest that interactivity is a key determinant of learning. As noted in Dufresne et al, (1996, 3) the construction of student knowledge is "...facilitated by instruction, [but] not the direct consequence of instruction." According to these perspectives, it is the improvements in interactivity afforded by GRS that should improve learning.

While GRS may help create interactivity, it is not the only way to achieve it, and it is possible to use GRS without improving interactivity, as established by the GRS studies of the 1960's and 1970's. This raises questions as to the ways in which pedagogy may be interactive, and how GRS can contribute to each dimension. According to Borsook and Higginbotham-Wheat (1991) an interactive pedagogy has the following key elements: immediacy of response to questions; instructor responsiveness to the needs of the students; performance feedback (outcome and explanatory); and bi-directional communication. The degree of interactivity in different pedagogical approaches ranges from low (e.g., lectures with limited bi-directional communication) to high (e.g., tutorial sessions with immediate feedback and extensive bi-directional communication).

The value of an interactive pedagogy is indicated by studies such as Crouch and Mazur (2001). They report that over a 10-year period with multiple instructors, one form of an interactive pedagogy known as peer instruction, resulted in significant increases in learning on a standard exam relative to pre-course results on the same exam. The normalized gain in knowledge more than doubled relative to the gain provided by conventional lecture approaches. However, this and similar studies of interactive pedagogies did not investigate GRS.

Further reinforcement of the importance of interactivity and the ways in which technology can improve this factor arises from the learning framework suggested by Laurillard (1993, 2002). Laurillard's framework stresses the importance of interactivity between the instructor and student for a number of activities viewed as integral to learning, including agreement on learning goals related to the concept at hand, student practice and instructor feedback on their efforts. This should then be followed by student reflection and integration of their learning with their initial understanding, then by subsequent student articulation of this new

understanding and additional iterations of the framework.

Taken together, Laurillard's framework and Borsook and Higginbotham-Wheat's (1991) elements of interactivity suggest GRS could improve learning by improving interactivity in several ways:

1. By providing an opportunity and incentive for each student to practice their understanding by developing an answer to the posed question and entering it into the system using the response pad, which is then immediately followed by instructor feedback on the response (outcome feedback) and additional explanation (explanatory feedback).
2. By improving instructor understanding of the variations in understanding between the students and the instructor. This can then provide the basis for subsequent discussion and follow up questions by students and instructor, as well as for additional practice with the GRS.
3. By providing students with information to compare their understanding to that of their peers, which may act as an incentive to reflect on their learning and comprehension.

Synthesis of the Literature and Research Questions To Be Investigated

Research on learning suggests the value of interactive pedagogies across a number of disciplines, but typically does not consider technological supplements such as GRS. The existing research on GRS indicates a very favorable response to GRS technology by students (e.g., high student satisfaction), but little objective evidence of performance effects, perhaps because more traditional lecture formats were usually used. For those studies that do report some learning effects of GRS, it is difficult to disentangle the performance effects related to the use of an interactive pedagogy deployed *in conjunction* with the GRS, from that of the GRS itself.

While there is theory such as that provided by Laurillard (2002), and evidence (e.g., Hake

1998) that more interactive pedagogies do improve student learning, there is little research to support or refute the claim that GRS improve learning *beyond* the changes provided by moving to a more interactive pedagogy. As Judson and Sawada (p. 175) summarize their findings regarding GRS' effect on student performance, "... the issue of academic achievement remains open."

Given our review of the literature, we investigate whether GRS does have an incremental effect on student performance in the context of accounting education. We also investigate whether student satisfaction with GRS is similar in an accounting educational context to the high levels reported in other disciplines. Although *ex ante* we have no reason to expect accounting students' satisfaction with GRS to differ from the results documented in other disciplines, it is important to assess as a poor implementation of GRS (i.e., low satisfaction) could decrease or eliminate any learning benefits.

Our specific research questions are thus as follows:

1. Do GRS, when used with interactive pedagogies, improve student satisfaction (as measured by in-class surveys) relative to the use of interactive pedagogies?
2. Do GRS, when used with interactive pedagogies, improve objective indicators of greater student engagement, as measured by verbal class participation (using both counts of participation and average percentage of class participating) relative to the use of interactive pedagogies alone?
3. Do GRS, when used with interactive pedagogies, improve direct learning outcomes (measured by exam performance) relative to the use of interactive pedagogies alone?

RESEARCH DESIGN

This paper reports the results of one study, conducted as part of a larger examination of the effects of GRS in three different accounting courses. Since the study was conducted in a similar manner across the three courses, we focus on one course as an example of the research design and analysis, while noting major differences in findings across courses in the discussion of the results. This approach gives some sense of the generalizability of our findings across different courses. The focal course is an introductory management accounting course taught by one instructor in the fall of 2004, with an enrolment of 184 students. All students were enrolled in an accounting co-op honours program, resulting in a relatively homogeneous group. The GRS chosen for the study was developed by eInstruction Inc. The response pads resemble a television remote control, and enable students to answer true/false, yes/no, and multiple choice questions with up to five alternative answers.

Study Procedures

The course had four sections: three had approximately 40 students each and one had 72 students. The approach to deploying the GRS was to alternate usage of the system over the term, with two sections using the GRS for the first half of the term, while the other two sections used the system for the second half of the term. Each usage period was roughly five weeks in length. Each response pad had a numeric identifier, and was assigned to a particular student for the duration of usage. Students were told that the pads were being evaluated on a trial basis for consideration in future years and in other accounting courses.

This alternating usage approach facilitated a between-group comparison of performance across GRS and non-GRS sections, since each section was covering the same topics with the

same instructor so that differences in these factors should not drive the results from performing comparisons across sections. It also enabled analysis of within-subject changes in behavior corresponding to the GRS and non-GRS periods for each student to be analyzed, since every student used the GRS for some part of the course. Since GRS effects were measured for the first part of the term for one section and the last part of the term for the other section, effects due to maturation or fatigue should affect both conditions equally when comparing GRS effects on a within subjects basis. The five-week duration of use also make it unlikely that novelty effects alone would lead to greater satisfaction with the GRS.[4]

For each class, 4-6 questions were prepared ahead of time by the instructor on aspects of the material the students were to have read about prior to the class[5]. Questions were largely multiple choice, with a few true/false and polling type questions also used. Each class was conducted using an approach modeled after an interactive pedagogy known as peer instruction (Crouch and Mazur 2001). For both the GRS and non-GRS sections, asking the questions using the GRS software was interspersed with computer slide presentations about the topic. The material on the slides was discussed and then a related GRS question asked. For both the GRS and non-GRS sections, the students were encouraged to discuss the question with the students seated next to them before answering. Students were also free to use their textbooks or notes. For the GRS section, students used their response pads to answer the question when ready, with the instructor displaying the histogram of aggregated responses, including the correct response

[4] Our observation of student reactions to the technology indicate that during the first few classes in which it was used, students exhibited more excitement about using the response pads. Once the initial novelty subsided, students seemed to accept use of the GRS as a normal part of the in-class routine.

[5] The number of questions to use for 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 4-6 questions in an 80-minute class takes a significant amount of time.

highlighted, when all students had provided a response.[6] For the non-GRS sections the instructor asked for a volunteer to answer the question. In the GRS section, if the responses indicated a significant number of students were confused about the correct answer, a student volunteer was asked to explain their response[7], followed by further discussion. In the non-GRS section, if the student volunteer got the answer wrong, the instructor polled additional students until either the right answer was provided or the level of confusion indicated significant student difficulties. The instructor would display the right answer, followed by further discussion.

This approach was used to make the pedagogy between the GRS and non-GRS groups as similar as possible. Both groups had a similar degree of interactivity, and both groups were exposed to the GRS questions. The only difference between the groups was that in the GRS group all students were able to answer the question, and they saw the aggregate section response to that question displayed as a histogram. The "treatment effect", if any, should thus be related to this difference, which is solely related to the technology rather than to differences in underlying pedagogy. Our comparison across sections and within subjects should reflect the effects of the technology incremental to those of using a more interactive pedagogy.

To provide incentives to think about the responses selected with the GRS, 5% of the course mark in each course was based on GRS usage, and an additional 5% of the course mark was based on oral class participation (either asking or answering content related questions).

---

[6] The time allowed for students to enter their responses varied. For the GRS groups, the instructors prompted the remaining students to enter a response when in excess of 80% of students had finished responding. For the non-GRS groups, the instructors determined that the majority of students were ready to answer the question when most of the discussion between students had ended.

[7] An alternative approach would be to ask all the students to raise their hands and count the number selecting each possible response. However, this was rejected as impractical and time consuming.

Measures of Satisfaction and Learning Outcomes

To collect data on student satisfaction and self-reported effects on learning and in-class behavior, a survey was administered in all courses and sections in the middle of the term immediately after the GRS was switched to the other section(s), and at the end of term (see the Appendix for an annotated copy of the survey). Survey 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. The survey responses were anonymous to encourage students to answer honestly, with each question answered using a nine point Likert scale centered on zero, with the end points labeled "strongly agree" (4) and "strongly disagree" (-4) and the mid-point (0) labeled "neutral." Students were asked to write a unique identifier known to them but not the instructor on the first page of each of the two surveys to permit a within subjects analysis of the responses to the questions that appeared on each survey.

The average responses to questions specific to student use of the GRS (Questions 10-23) were compared to the scale midpoint of zero to determine student satisfaction with the technology.[8] The average responses to questions generally related to the course, such as whether the course was interesting, etc. (Questions 1-9) were compared across the GRS and non-GRS groups. Responses to these questions were also examined using a within subjects analysis to determine if use of GRS had any impact on students' perceptions about these general aspects of the course. The remaining end-of term questions (Questions 24-27) asked students to compare various aspects of the course and their own behavior across the GRS and non-GRS periods and were analyzed in the same manner as the GRS specific questions.

[8] Some survey questions (Questions 10, 12, 15, 17, and 22) were stated in negative form to detect students circling the same answer for each question without reading. However, to clarify the presentation and discussion of these items, they were reverse scored and reported in the positive.

The objective data on in-class participation was collected by a teaching assistant (TA) who attended every class and counted the number of questions asked and answered by each student, as well as whether they were present or absent. The participation data was then used to create two standardized participation measures called $Participation_{ask}$ and $Participation_{answer}$[9]. $Participation_{ask}$ is calculated as:

> $Participation_{ask}$ = (average number of questions asked for student *i* per class attended in part *p* of the course – average questions asked for all students for all course sections for part *p*)/ standard deviation of questions asked for all students in the course (i.e., all sections) for part *p*. "*P*" represents the first or second half of the course.

Negative (positive) values of $Participation_{ask}$ indicate lower (higher) average questions asked per class by a particular student than the course average per class for all students in all sections. To determine GRS effects, a difference measure called $DParticipation_{ask}$ was then calculated as $Participation_{ask1} - Participation_{ask2}$. If the GRS increases participation levels, then $DParticipation_{ask}$ should tend to have positive values for students who used the GRS in the first part of the course, and negative values for students who used the GRS in the second part of the course. The variables $Participation_{answer}$ and $DParticipation_{answer}$ were calculated in the same manner for questions answered orally by students.

The direct measure of GRS learning effects was based on midterm and final examination performance (each of which comprised both multiple choice and written answer questions) for each student. Performance for each student is compared for questions they covered when they were using the GRS versus questions covered when they were not using the GRS.

Three separate standardized measures of performance were calculated, intended to correspond to

[9] Because the number of questions *asked* by students (or course related comments) may be a better reflection of willingness to participate than the number of questions answered, $Participation_{ip}$ was calculated separately for the number of questions asked and the number of questions answered by each student.

performance on exam questions most closely related to the material covered using the GRS (*SPerformance*) [10], performance on all multiple choice exam questions (*MCPerformance*), and performance on the total exam respectively (*TPerformance*). *Sperformance* has the greatest likelihood of capturing a GRS effect on learning, if one exists. Since the GRS was used primarily with multiple-choice questions in each course, we believed there might be a higher correlation between performance on the multiple choice questions and usage of the GRS, which would be captured by *MCPerformance*. *TPerformance* captures the overall effects of the GRS on learning of the course topics. The calculation for *Sperformance* is:

> *SPerformance= (%* correct on GRS related questions for one student *i* – average *%* correct on the same questions for all students in all sections of the course)/standard deviation of the % correct on the same questions for the course.

The measures for *MCPerformance* and *TPerformance* were calculated in a similar way, using the percentage correct of all multiple choice questions and all questions respectively. Analogous to the approach taken with the participation measure to determine GRS effects, difference measures for each of these scores were calculated, called *DSPerformance, DMCPerformance, and DTPerformance* respectively. For example, *DSPerformance* was calculated as *Sperformance* for midterm questions – *Sperformance* for final exam questions. If use of a GRS improves student learning, then each of these difference measures should tend to have positive values for students who used the GRS in the first part of the course (since their performance relative to the average for the selected midterm exam questions will tend to be higher than their performance relative to the average for the selected final exam questions), and negative values for students who used the GRS in the second part of the course.

[10] GRS-related questions included on the mid-terms and final exam were essentially the same as those covered in class with changes made only to the parameter values.

ANALYSIS

Analysis of Survey Results: Student Self-Reported Effects of GRS

Table 1 shows the results of analyzing the student responses questions concerning the GRS (Questions 10-23). The response rate for the surveys is high (92%), because it was distributed and collected in class. For all questions, the average response was significantly different from 0 in the expected direction at a p-value of .001, indicating student views towards the effects and use of GRS were generally positive.[11] Average responses for the groups who used the GRS for the first versus second part of the course are similar (not tabulated), with the only differences in significance related to Question 16, which was significantly different from 0 at the .05 level for those using the pads for the second part of the course.[12]

<<< Insert Table 1 About Here >>>

The average responses indicate clear agreement with the idea that the GRS (described as "response pads" in the survey questions) should be used in other courses, as well as with the idea that the response pads were easy to use. The average responses also indicate clear agreement with the idea that the GRS and lectures were effectively integrated, and that the GRS were enjoyable to use. There was also agreement with the idea that the advantages of the GRS outweighed the disadvantages. Overall, the response to these questions are similar to those reported in previous studies in terms of student satisfaction and enjoyment, and there is no indication from the students' perspective of problems in how the GRS was implemented or their

[11] Although the mid-point of the scale (0) was labeled "neutral" not all respondents necessarily interpreted it as such. To provide a more stringent test of students' perceptions of GRS, we also compared the average responses for each question to 1. All comparisons remain significant at a p-value of .01 except for questions 20 and 21 in Table 1 ($p > .50$).

[12] We also compared responses for the Table 1 questions across the four sections of the course. No significant differences were found except for question 16 where the three smaller sections of the course (30-41 students per section) all agreed more strongly that the summarized answer feedback helped them track their progress in the course compared to students in the large section of the course (72 students).

use of the technology.

Students' written comments (provided on the survey) on the GRS include the following:

- "Good way to make sure everybody participates, because many are not comfortable. Good to identify problems class is having with subjects or particular ideas."
- "I believe that using the response pads is an ingenious idea and should be used in most courses; especially courses that require reading and are based on comprehension of "concepts" and not so much mathematical operations, etc. It really gives students a chance to gauge their performance on how well they've prepared or understood the topics."
- "My favorite course, even boring material is engaging and fun to learn."
- "Response pads can feel a little intimidating at times, but feel very rewarding when you've answered a question correctly."
- "I think the response pads create a strong incentive for students to stay on track with the course. It is a good method of allocating participation marks. I think it is more fair of an assessment than oral participation since it requires reading ahead of time."
- "I thoroughly enjoyed the response pads. It maintained my interest and encouraged participation and thinking."

There were few negative comments related to the GRS, with those that were provided often related to the stress of having to be prepared for class and to provide the correct answer to the GRS questions. Some sample negative comments included:

- "I believe that the response pad does encourage me to prepare for class. However it also has an effect of discouragement because when I had already tried my best to prepare for the class and still manage to answer the questions wrong, I worry about whether preparing and reading ahead is useful or not."
- "Response pads place a lot of pressure on the class. Made it a lot more stressful. It was also not an appropriate measure of knowledge since people often collaborated or didn't have the time to really think about it."
- "Response pads distracted me from the learning process. It definitely SHOULD NOT be used again next year."
- "Class felt more stressful with response pads because sometimes you don't have time to prepare for class to answer the questions."

Table 2, Panel A shows the average between-subject responses for the GRS versus non-GRS groups for the more general course satisfaction questions (survey questions 1-9). As shown in the table, there is no statistically significant difference in the responses for the GRS versus non-GRS groups. These results suggest that students generally did not perceive various aspects

of the course differently as a result of GRS use, despite their favorable views of the technology itself. We also checked for order effects by comparing the responses of the GRS sections to the non-GRS sections separately for the interim and end-of-term surveys. No differences were found except on the end-of-term survey where students who had not been using the GRS for the last half of the course indicated they were significantly ($p < .10$) less comfortable participating in class than students who had been using the GRS.

A within-subjects analysis of the responses to the general course questions was also performed. Of the 172 middle of term survey responses, 128 (74%) could be matched to corresponding end of term surveys for the same student. The results (non-tabulated) for only two of the nine survey questions are significantly different, with the rest showing no significant effects of GRS on students' perceptions of the course versus the same students' perceptions of the course when not using the GRS. Panels B and C of Table 2 explore these two significant within-subject differences further.

Panel B shows a weakly significant difference related to GRS use ($p < .10$) between the mean responses to the question regarding students' comfort participating in the course (Survey question 2). Panel B also shows that the strength of this effect depends on when the students used the GRS (first or second half of the course) since the interaction term (GRS First x GRS) is significant ($p < .05$). Further analysis of the mean differences (not tabulated) found that students who used the GRS first reported being significantly ($p < .01$) less comfortable participating during the second half of the course (means of 1.50 versus .86) while students who used the GRS second reported being similarly comfortable ($p > .70$) participating during each half of the course (means of 1.59 versus 1.66). In other words, use of the GRS first is associated with students reporting being less comfortable participating in class after the GRS were taken away, while no

such effect was reported by those who used the GRS for the second portion of the course.

A similar pattern of results emerges in Table 2, Panel C for the measure of the degree to which students reported being "comfortable answering oral questions." Further analysis of the significant interaction term ($p < .01$) indicates that students who used the GRS first reported being significantly ($p < .01$) less comfortable answering oral questions during the second half of the course (means of 1.33 versus .43) while students who used the GRS second reported being similarly comfortable ($p > .70$) answering oral questions during each half of the course (means of 1.05 versus 1.00).

<<< Insert Table 2 About Here >>>

Table 3 shows the responses to the end of term questions that asked the students to directly compare their understanding, course enjoyment, and comfort in participating for the part of the course that used the GRS to the part of the course that did not use the GRS (questions 24-27 on the survey). While the responses for understanding and enjoyment are significantly different from 0 (in the expected direction) with a p-value of less than .001 for both questions, the responses regarding comfort in answering or asking questions are not significantly different from 0.[13]

<<< Insert Table 3 About Here >>>

When the survey results are considered as a whole, the responses concerning participation effects of the GRS are somewhat contradictory. Overall the results in Table 1 suggest students enjoyed the GRS, felt it improved learning, and noticed no implementation problems. However, while students indicated that they felt more comfortable participating when

[13]We found no significant differences in the responses to these questions between students who used the GRS during the first half of the course and those who used the GRS during the second half of the course. Similarly, we found no differences in responses across the four sections of the course.

the response pads were used (Table 1, Question 14), they did not on average feel it improved oral participation relative to the part of the course where the GRS was not used (Table 3). Similarly there was no indication of self-reported improved comfort in participating in class when the response means of the GRS versus non-GRS groups were compared (Table 2, Panel A). It could be that the initial survey question asking whether the GRS made the student feel more comfortable participating in class was interpreted as meaning "participating with the response pad", while the other participation questions specifically refer to asking questions and answering oral questions, which the students did not feel were affected by GRS usage. Alternatively, the placement of the participation question among the other GRS specific questions may have created a "halo effect", which led to a positive response for this question but not for the other questions that were not as proximate to the GRS specific questions.

Prior research on student satisfaction with GRS has relied on survey results without the benefit of a control group. When the GRS specific (i.e. treatment group only) questions are compared to prior studies, the results reported here are very similar to those reported in the prior research. However, when the average responses for questions related generally to course enjoyment and conduct are compared between the GRS and non-GRS groups, the results do not support an overall improved perception of the course because of GRS use.

The results reported in Panels B and C of Table 2 suggests that the GRS has a negative effect on students' self-reported participation comfort level, but only for the period when the technology has been experienced and then removed. Students who did not use the GRS for the first half of the term showed no significant difference between surveys in their stated comfort with participation overall or with their comfort in answering oral questions. One possible interpretation is that students become less comfortable with conventional participation means

after they have used the technology, but other effects may also be at work. However, this result does not appear to be attributable to an overall dissatisfaction with the course resulting from the GRS being discontinued as the results for the other seven measures of satisfaction with the course do not significantly differ across the two surveys.

Overall these results do not suggest that the GRS is ineffective even when students express satisfaction with its use, but may mean that when an interactive pedagogy is used, the GRS has a limited impact on incrementally improving perceptions of the course in general.

Analysis of GRS Effects on Objective Measures of Participation

Table 4 Panel A reports the results of the analysis of variance conducted on the objective participation measure $DParticipation_{ask}$, with whether the student had the GRS system for the first part of the course included as an independent factor. As noted earlier, if the GRS increases average participation levels per student, then $DParticipation_{ask}$ should tend to have positive values for students who used the GRS in the first part of the course, and negative values for students who used the GRS in the second part of the course.

Panel A shows that use of the GRS did result in a significant difference in the average number of questions asked per student ($p < .10$). However, analysis of the mean values (not tabulated) indicates the result is opposite to what we expected. The mean of $DParticipation_{ask}$ for students who used the GRS in the first part of the course is -.16 compared to .10 for those that used it in the second part. This indicates students asked more questions when the GRS was *not* in use. To assess whether the use of the GRS had a significant impact on participation *within* each condition we compared each $DParticipation_{ask}$ score (-.16, and .10) to zero. Each score is significantly different from zero at the .05 level indicating that use of the GRS significantly

reduced the average number of questions asked within each condition.[14]

Further analysis of this result was done by splitting each class into quartiles based on total participation counts over the term in each class, and assigning each student to a particular quartile depending on their participation over the entire term relative to the class as a whole. The analysis of variance was then recalculated with the participation quartile classification as an additional factor. The intent of adding the participation quartile factor is to see if the GRS has a differential effect on students who are highly active participators versus those who tend to participate little. Panel B of Table 4 shows that adding this factor to the analysis of variance results in a somewhat more significant effect of GRS on asking questions ($p < .05$).

<<< Insert Table 4 About Here >>>

Table 5 provides further evidence on the effects of GRS, by examining its effects on the percentage of students asking questions in each class. The number of questions asked by the professor is used as a covariate to control for the possibility that more active questioning by the professor may generate questions from students. Whether the GRS was used in the first or second half of the course was also included as a between subjects variable to determine if it had any impact on the percentage of students asking a question.

Consistent with the results reported in Table 4, (Panels A and B) the analysis in Table 5 shows a significant effect of the GRS on the percentage of students participating in class ($p < .05$), with a comparison of means (not tabulated) showing a smaller percentage of the class participated when the GRS was used compared to when it was not (7.6 % versus 10.0%). The results in Table 5 also show that the main effect of GRS order was not significant nor was its

[14] We also tested for order effects by comparing the absolute values of the difference scores (.10 and .16); the difference is not significant.

interaction with GRS use (respectively, $p = .846$; $p = .540$).

<<< Insert Table 5 About Here >>>

A possible reason for the GRS apparently decreasing students' willingness to ask questions in class may be related to the difficulty of the multiple choice questions used with the GRS. The average percentage correct for all GRS questions was 84%, which means that the histograms displayed after each GRS question was completed showed that the majority of the class had responded correctly. A possible implication of this is that students are less likely to ask questions when GRS results show that a large majority of the class understands the concept being discussed. Conversely, no such feedback is available in the non-GRS sections so there is more uncertainty (among the students) about the extent to which a concept is understood. To evaluate this possibility, a correlation was calculated between the percentage of students asking a question in each class where the GRS was used and the overall score for that class on the GRS multiple choice questions. The correlation coefficient is negative (-.28) and significant ($p < .10$) indicating the more difficult the GRS questions, the greater the percentage of students asking questions. Thus it appears that a GRS can actually stifle discussion in classes where the feedback from the system indicates the majority of students understand the concepts being reviewed.

Finally, Table 4, Panel C shows that the GRS did not have a significant effect on participation as measured by the average number of non-GRS questions orally answered by students (i.e. students responding to questions asked by the professor rather than displayed via the GRS software). Further analysis (not tabulated) indicates that the percentage of students answering non-GRS questions in each class was also unaffected by the use of the GRS. These results may be attributable to the fact that five marks were given in the course for verbal participation. Thus students had some incentive to answer questions in class, which may have

reduced the likelihood of detecting any effects of the GRS.

Analysis of GRS Effects on Examination Performance

The final analysis was to determine the effects of the GRS on the three measures of direct learning outcomes. Table 6 shows the results of the analysis of variance performed on the standardized scores for each of the three measures. As noted earlier, positive values for these variables should tend to be associated with students who used the GRS in the first part of the course, and negative values for these variables should tend to be associated with students who used the GRS in the second part of the course.

<<< Insert Table 6 About Here >>>

The only significant result is the impact of GRS on performance as measured by exam scores for the multiple choice questions closely related to those displayed with the GRS during class. (Panel A, $p < .05$)[15]. The mean (not tabulated) for the *DSPerformance* "GRS First" groups is .24 compared to -.15 for the "GRS Second" groups. Examination of the raw means for the multiple choice questions (not tabulated) shows that the "GRS First" groups were two percentage points above (below) the class average for the first (second) mid-term. Similarly the "GRS Second" groups were one (two) percentage points below (above) the class average for the first (second) mid-term. To assess whether the use of the GRS had a significant impact on exam performance *within* each condition we compared each *DSPerformance* score (.24 and .-15) to zero. Each score is significantly different from zero (respectively at the .05 and .10 level) indicating that use of the GRS significantly improved performance on the GRS-related multiple choice exam questions *within* each condition.[16]

[15] Similar performance effects of the GRS were found for one of the other courses in which the GRS was used.
[16] We also tested for order effects by comparing the absolute values of the difference scores (.24 and .15); the difference is not significant.

The results for Panels B and C show that GRS did not have a significant effect on performance measured using either all multiple choice questions on the exam or the entire exam score (respectively p-values of .481 and .349). Means for *DMCPerformance* (*DTPerformance)* for the "GRS First" and "GRS Second" groups respectively are .06 and -.05 (.05 and -.07), which are consistent with the expected sign of the differences, but the magnitude of the difference is non-significant. Thus, the impact of the GRS appears to be limited to performance on questions very similar in nature to those employed when using the system. However, since all students (GRS and non-GRS users) saw the questions, the differential effects of the technology would appear to be associated with some characteristic of GRS, rather than greater familiarity by the GRS users with the GRS related questions.

CONCLUSIONS AND LIMITATIONS

A review of vendor websites (e.g., those of eInstruction and GTCO CalComp) suggests the GRS technology is becoming a popular tool with educators. However, the results of our research suggest the need to temper some of the enthusiastic claims attributed to use of GRS. Contrary to the claims of proponents, we do not find objective evidence of heightened student engagement as measured by verbal participation. While we do find evidence of positive learning outcomes associated with GRS, the effects are small and limited to those questions most directly related to those displayed by the GRS. We believe our results highlight the need for further research to determine if and how a GRS can affect learning outcomes.

Our study provides evidence that replicates previous findings and extends the literature on the effects of GRS on direct and indirect learning outcomes. Consistent with prior research,

average student responses to our survey questions suggest strong student satisfaction with the technology. This finding reduces the possibility that any non-significant results on our more objective measures of student engagement and learning arose from a poor implementation of the system.

An extension of prior research to include a control group found no difference between GRS and non-GRS groups in overall perceptions of the course, and no difference on a between subjects basis in student perceptions of willingness to ask or answer questions. Our between subjects design enable us to conclude that student reports of enjoyment of the GRS technology do not necessarily mean a course is generally regarded more favorably across any of several dimensions, when the effects of increased interactivity associated with changes in pedagogy are held constant.

Our study provides new insights regarding the effects of GRS on participation. The within subjects self-reported results that students are significantly less comfortable participating and less comfortable answering questions once they have had the GRS and it has been removed. Further research is needed to identify the cause of this outcome. However, it could be that the GRS creates a comfortable participation environment for students that they are most aware of when it is gone. The increased average number of questions asked and a higher percentage of students asking questions in the groups not using the GRS is contrary to what would be expected if GRS does heighten student engagement. Further research is needed to explore our finding regarding the relationship between GRS question difficulty and asking questions as one possible factor in GRS's effect on class participation.

Our study is the first that we are aware of to examine whether a GRS has any effect on learning outcomes, incremental to an interactive pedagogy. Whether removal of the 5% mark

incentive for the GRS related multiple choice questions covered in class would reduce the exam performance effects found is an empirical question.

Finally, differences in pedagogy is the other factor that should be explored in determining the effects of GRS. Despite our efforts, it is possible that the pedagogical approach we used was not interactive enough, or some other pedagogical characteristic would be more important in enabling a GRS effect. Further research is needed to investigate all of these factors before arriving at conclusions on the effects of GRS on student satisfaction and learning in accounting education.

REFERENCES

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 1999.

Bass , R. 1998. Engines of inquiry: Teaching, technology, and learner-centered approaches to culture and history. Available at http://www.georgetown.edu/crossroads/guide/ **engines** .html. Accessed September 13, 2004.

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. To appear *in Proceedings of the International Conference on Computers in Education 2004*, Melbourne, Australia Nov, 30th – Dec 3rd.

Davis, S. 2003. Observations in classrooms using a network of handheld devices. *Journal of Computer Assisted Learning* 19: 298-307.

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 . Accessed March 25, 2004.

Draper. S.W. and M.I. Brown. 2004. Increasing interactivity in lectures using an electronic voting system. *Journal of Computer Assisted Learning* 20: 81-94..

Draper, S.W., J. Cargill, and Q. Cutts . 2002. Electronically enhanced classroom interaction. Australian Journal of Educational Technology 18(1): 13-23.

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.

Hake. R.R. 1998. Interactive engagement versus traditional methods: A six thousand student survey of mechanics tests data for introductory physics courses. American Journal of Physics 53: 1043-1055.

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.

Laurillard. D. 1993. *Rethinking university teaching: A Framework for the Effective Use of Educational Technology*. Routledge: London.

Laurillard. D. 2002. *Rethinking university teaching: A Framework for the Effective Use of Educational Technology.* Second edition, Routledge: London.

Learning Media Unit. 2001. Student and staff feedback on using an electronic group response system in a mechanical engineering lecture at the University of Sheffield. *Learning Media Unit Evaluation Report Project 39, University of Sheffield.*

McCabe. M. 2003. Do mathematics interactive classrooms help academics engage learners? *MOSR Connections*.3(4):

McCabe, M. and I. Lucas. 2003. Teaching with CAA in an interactive classroom: Death by Powerpoint, life by discourse. *Proceedings of the 7th International Conference on CAA.*

Madill, B. 2004. Improving student interaction via a personal response system and peer instruction. Available at http://www.be.coventry.ac.uk/BPBNetwork/casestudy/uce_tla3i.htm . Accessed March 25, 2004.

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, M. Gilbert. 1997. Physics lecturing with audience paced feedback. Available from http://www.bedu.com/Publications/PhysLectAPF.pdf. Accessed September 13, 2004.

Purchase, H.C., C. Mitchell, and I. Ounis, 2004. Gauging students' understanding through interactive lectures. Available from http://www.psy.gla.ac.uk/~steve/ilig/papers/hcp1.pdf. Accessed September 13, 2004.

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 2004. Available at. http://www.ubiqcomputing.org/CATAALYST_AERA_Proposal.pdf. Accessed March 25, 2004

Stuart, S. and M. I. Brown. 2003. Traditional and non-traditional resources: Providing a well-supported learning environment. *Association of Learning Technology Journal* 11(3): 58-68.

Wit, E. 2002. Who wants to be...The use of a personal response system in statistics teaching. *MSOR Connections* 3(2): 5-11.

Table 1
Survey Response Means for GRS Specific Questions

Survey Question Question	Number of responses	Response means
10. Course does not focus too much on using response pads	172	1.959****
11. Lecture and response pads effectively integrated	172	2.358****
12. Response pads are easy to use	172	3.294****
13. Enough time to answer questions using response pads	171	1.658****
14. More comfortable participating when response pads used	170	2.038****
15. Response pads help learn material	171	1.751****
16. Summarized class answers help track progress	171	1.535****
17. Confident that response pads accurately record responses	171	1.912****
18. Enjoy using response pads	171	2.351****
19. Instructor clarifies correct solution for response pad questions	171	2.953****
20. Response pads encourage working harder to answer questions	172	1.250****
21. Response pads encourage working harder to prepare for class	171	0.825****
22. Response pads should be used in other courses	170	1.988****
23. Advantages of response pads outweigh disadvantages	170	2.250****

****, ***, **, and * refer to a significant difference from 0 at the .001, .01, .05, and .10 levels respectively, one tailed.

Table 2
Survey Response Means for General Course Questions for GRS versus non-GRS Groups

Panel A: Between Subjects Analysis[1]

Survey Question	Number of responses: GRS	GRS Response Mean	Std. Deviation	Number of responses: Non-GRS	Non-GRS Response Mean	Std. Deviation
1. Course interesting	172	1.884	1.4503	169	1.769	1.2464
2. Comfortable participating	172	1.526	1.7264	169	1.269	1.6894
3. Course not well organized	172	-2.337	1.5958	169	-2.411	1.4240
4. Material presented effectively	172	1.930	1.6243	169	1.896	1.3283
5. Do not feel comfortable asking questions	172	-0.965	1.9791	169	-1.101	1.7109
6. Classes help master course material	172	1.192	1.8554	169	1.249	1.6503
7. Trouble paying attention in class	172	-0.846	1.9797	169	-0.876	1.8329
8. Comfortable answering oral questions	172	1.047	1.8187	169	0.799	1.8276
9. Not required to think much in class about course concepts	172	-2.044	1.4856	169	-2.101	1.4440

Panel B: Within subjects analysis with response to "Comfortable Participating" as the repeated measure; n = 128

Source	Type III Sum of Squares	df	Mean Square	F	p
GRS[2]	5.025	1	5.025	3.458	.065
GRS First[3] x GRS	7.564	1	7.564	5.205	.024
Error	183.107	126	1.453		

Panel C: Within subjects analysis with responses to "Comfortable Answering Oral Questions" as the repeated measure; n = 128

Source	Type III Sum of Squares	df	Mean Square	F	p
GRS[2]	13.958	1	13.958	11.597	.001
GRS First[3] x GRS	11.083	1	11.083	9.209	.003
Error	151.651	126	1.204		

[1]Note: Minor differences in number of responses between the GRS and non-GRS groups reflect the number of students in class on the days the surveys were administered.
[2]GRS: the repeated measures factor indicating whether the GRS was being used
[3]GRS First: indicator of whether the student used the GRS during the first or second half of the course.

Table 3
Survey Response Means for End of Term Comparative Questions Comparing the GRS to the non-GRS Portions of the Course

Survey Question	Number of responses	Response means
24. Compared to the part of the course that did not use the response pads, I had a better understanding of material when response pads used	161	0.689****
25. Compared to the part of the course that did not use the response pads, the course was more enjoyable when response pads used	161	1.720****
26. Compared to the part of the course that did not use the response pads, I felt more comfortable asking questions when response pads used	161	-0.037
27. Compared to the part of the course that did not use the response pads, I felt more comfortable answering oral questions when response pads used	161	0.090

****, ***, **, and * refer to a significant difference from 0 at the .001, .01, .05, and .10 levels respectively, one tailed.

Table 4
Analysis of the Effects of GRS on Standardized zScores Based on Average Oral Participation Per Student Per Class

Panel A: Between Subject Results (N=179) with *DParticipation*$_{ask}$ as the dependent variable:[1]

Source	Type III Sum of Squares	df	Mean Square	F	p
GRS First[2]	2.649	1	2.649	3.211	.075
Error	146.035	177	.825		

Panel B: Between Subject Results with a Control for Level of Participation (N=186) and *DParticipation*$_{ask}$ as the dependent variable: [1]

Source	Type III Sum of Squares	df	Mean Square	F	p
GRS First	3.243	1	3.243	3.957	.048
Participation Quartile[3]	1.319	3	.440	.536	.658
GRS First x Quartile	4.059	3	1.353	1.651	.179
Error	140.135	171	.820		

Panel C: Between Subject Results (N=179) with *DParticipation*$_{answer}$ as the dependent variable:[4]

Source	Type III Sum of Squares	df	Mean Square	F	p
GRS First	.024	1	.0234	.055	.815
Error	47.672	177	.269		

[1] *DParticipation*$_{ask}$: calculated by taking the difference between the z-scores of the average number of questions asked per class per student when the GRS was used and the z-scores of the average number of questions asked per class per student when the GRS was not used.
[2] *GRS First*: a dummy variable indicating the portion of the course (first half or second half) the GRS was used.
[3] *Participation quartile*: each student was assigned to a quartile based on his or her overall level of participation for the course.
[4] *DParticipation*$_{answer}$: calculated by taking the difference between the z-scores of the average number of questions orally answered per class per student when the GRS was used and the z-scores of the average number of questions orally answered per class per student when the GRS was not used.

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

Between Subject Results with a Control for Number of Questions asked by Professor (N=88) and *Percent*$_{ask}$ as the dependent variable:[1]

Source	Type III Sum of Squares	df	Mean Square	F	p
GRS[2]	120.46	1	120.46	5.207	.025
GRS First[3]	.873	1	.873	.038	.846
GRS x GRS First	8.741	1	8.741	.378	.540
Prof. Questions[4]	6.604	1	6.604	.285	.595
Error	1920.206	83	23.135		

[1]*Percent*$_{ask}$: the percentage of students asking a question in each class. A separate observation for each section is included in the analysis (4 sections x 22 classes = 88 independent observations).
[2]GRS: a dummy variable indicating whether or not the GRS was in use for the class.
[3]GRS First: a dummy variable indicating for each section whether the GRS was used during the first or second half of the course.
[4]Prof. Questions: the number of questions asked by the professor during each class for each section.

Table 6
Analysis of the Effects of GRS on Standardized zScores Based on Exam Performance

Panel A: Between Subject Results with *DSPerformance* as the Dependent Variable (N=186) [1]:

Source	Type III Sum of Squares	df	Mean Square	F	p
GRS First	6.915	1	6.915	5.746	.018
Error	221.413	184	1.203		

Panel B: Between Subject Results with *DMCPerformance* as the Dependent Variable (N=186) [2]:

Source	Type III Sum of Squares	df	Mean Square	F	p
GRS First	.569	1	.569	.499	.481
Error	206.141	184	1.139		

Panel C: Between Subject Results with *DTPerformance* as the Dependent Variable (N=186) [3]:

Source	Type III Sum of Squares	df	Mean Square	F	p
GRS First	.656	1	.656	.881	.349
Error	135.629	184	.745		

[1] *DSPerformance*: the difference between the z-scores of the students' performance (%) on GRS-related exam multiple choice questions when the GRS was used and the z-scores of students' performance (%) on GRS-related exam multiple choice questions when the GRS was not used. Note that the same multiple choice questions were used for each section whether or not the GRS was in use.
[2] *DMCPerformance*: the difference between the z-scores of the students' performance (%) on all multiple choice exam questions when the GRS was used and the z-scores of students' performance (%) on all multiple choice exam questions when the GRS was not used.
[3] *DTPerformance*: the difference between the z-scores of the students' performance (%) on all exam questions when the GRS was used and the z-scores of students' performance (%) on all exam questions when the GRS was not used. The slight reduction in the sample size from the total class enrolment is due to students who missed the midterm exam, and thus could not have a difference score calculated for them.

APPENDIX
Example Survey Used to Collect Student Satisfaction and Subjective Measures of Learning Effects

Questions 1-9 were asked of all sections in both the middle and end of term survey. Questions 10-23 were asked only of these sections that had just completed using the GRS. Questions 24-27 were asked of all sections at the end of the term.

Please respond to <u>each</u> of the following questions by placing an "x" on the scale, wherever you feel appropriate. Please do <u>not</u> put your name on this document. This information will be compiled and given to the instructor, but all responses will be anonymous. This survey is not a substitute for the normal end of term course evaluation.

1. I find this course interesting.

-4	-3	-2	-1	0	1	2	3	4
strongly disagree		disagree		neutral		agree		strongly agree

2. I feel comfortable participating in this course.

-4	-3	-2	-1	0	1	2	3	4
strongly disagree		disagree		neutral		agree		strongly agree

3. This course is <u>not</u> well organized.

-4	-3	-2	-1	0	1	2	3	4
strongly disagree		disagree		neutral		agree		strongly agree

4. The course material is presented effectively.

-4	-3	-2	-1	0	1	2	3	4
strongly disagree		disagree		neutral		agree		strongly agree

5. I do <u>not</u> feel comfortable <u>asking</u> questions in this course.

-4	-3	-2	-1	0	1	2	3	4
strongly		disagree		neutral		agree		strongly

disagree agree

6. The classes help me master the course material.

-4 -3 -2 -1 0 1 2 3 4
| ---------- | ---------- | --------- | ---------- | ---------- | ---------- | --------- | ---------- |
strongly disagree neutral agree strongly
disagree agree

7. I have trouble paying attention in class.

-4 -3 -2 -1 0 1 2 3 4
| ---------- | ---------- | --------- | ---------- | ---------- | ---------- | --------- | ---------- |
strongly disagree neutral agree strongly
disagree agree

8. I feel comfortable answering oral questions in this course.

-4 -3 -2 -1 0 1 2 3 4
| ---------- | ---------- | --------- | ---------- | ---------- | ---------- | --------- | ---------- |
strongly disagree neutral agree strongly
disagree agree

9. I am not required to think much in class about course concepts.

-4 -3 -2 -1 0 1 2 3 4
| ---------- | ---------- | --------- | ---------- | ---------- | ---------- | --------- | ---------- |
strongly disagree neutral agree strongly
disagree agree

10. This course focuses too much on using the response pads.

-4 -3 -2 -1 0 1 2 3 4
| ---------- | ---------- | --------- | ---------- | ---------- | ---------- | --------- | ---------- |
strongly disagree neutral agree strongly
disagree agree

11. The lecture and response pad usage are effectively integrated.

-4 -3 -2 -1 0 1 2 3 4
| ---------- | ---------- | --------- | ---------- | ---------- | ---------- | --------- | ---------- |
strongly disagree neutral agree strongly
disagree agree

12. The response pads are not easy to use.

-4 -3 -2 -1 0 1 2 3 4
| ---------- | ---------- | --------- | ---------- | ---------- | ---------- | --------- | ---------- |

strongly disagree neutral agree strongly
disagree agree

13. I have enough time to answer the questions with the response pads.

-4 -3 -2 -1 0 1 2 3 4
| ---------- | ---------- | --------- | ---------- | ---------- | ---------- | --------- | ---------- |
strongly disagree neutral agree strongly
disagree agree

14. The response pads make me feel more comfortable participating in the course.

-4 -3 -2 -1 0 1 2 3 4
| ---------- | ---------- | --------- | ---------- | ---------- | ---------- | --------- | ---------- |
strongly disagree neutral agree strongly
disagree agree

15. Using the response pads does <u>not</u> help me learn the material in this course.

-4 -3 -2 -1 0 1 2 3 4
| ---------- | ---------- | --------- | ---------- | ---------- | ---------- | --------- | ---------- |
strongly disagree neutral agree strongly
disagree agree

16. Seeing the summarized class answers to response pad questions helps me track my progress in the course.

-4 -3 -2 -1 0 1 2 3 4
| ---------- | ---------- | --------- | ---------- | ---------- | ---------- | --------- | ---------- |
strongly disagree neutral agree strongly
disagree agree

17. I am <u>not</u> confident that the response pads accurately record my responses.

-4 -3 -2 -1 0 1 2 3 4
| ---------- | ---------- | --------- | ---------- | ---------- | ---------- | --------- | ---------- |
strongly disagree neutral agree strongly
disagree agree

18. I enjoy using the response pads to answer questions in this course.

-4 -3 -2 -1 0 1 2 3 4
| ---------- | ---------- | --------- | ---------- | ---------- | ---------- | --------- | ---------- |
strongly disagree neutral agree strongly
disagree agree

19. The instructor clarifies and explains the correct solution for questions answered with the response pads when a significant number of students have difficulty determining the correct answer.

-4	-3	-2	-1	0	1	2	3	4
strongly disagree		disagree		neutral		agree		strongly agree

20. Knowing the response pads will be used encourages me to work harder <u>to answer questions</u> in class.

-4	-3	-2	-1	0	1	2	3	4
strongly disagree		disagree		neutral		agree		strongly agree

21. Knowing the response pads will be used encourages me to work harder <u>to prepare</u> for class.

-4	-3	-2	-1	0	1	2	3	4
strongly disagree		disagree		neutral		agree		strongly agree

22. The response pads should <u>not</u> be used in other courses.

-4	-3	-2	-1	0	1	2	3	4
strongly disagree		disagree		neutral		agree		strongly agree

23. Overall, the advantages of using the response pads outweigh the disadvantages in this course.

-4	-3	-2	-1	0	1	2	3	4
strongly disagree		disagree		neutral		agree		strongly agree

24. Compared to the material in the part of the course that did <u>not</u> use the response pads, I have a better understanding of the material in this course when the response pads were used.

-4	-3	-2	-1	0	1	2	3	4
strongly disagree		disagree		neutral		agree		strongly agree

25. Compared to the part of the course that did <u>not</u> use the response pads, the part of this course that used the response pads was more enjoyable.

-4	-3	-2	-1	0	1	2	3	4
strongly disagree		disagree		neutral		agree		strongly agree

26. Compared to the part of the course that did <u>not</u> use the response pads, I felt more comfortable <u>asking</u> questions in the part of the course that did use the response pads.

-4	-3	-2	-1	0	1	2	3	4
strongly disagree		disagree		neutral		agree		strongly agree

27. Compared to the part of the course that did <u>not</u> use the response pads, I felt more comfortable <u>answering</u> <u>oral</u> questions in the part of the course that did use the response pads.

-4	-3	-2	-1	0	1	2	3	4
strongly disagree		disagree		neutral		agree		strongly agree

<u>Any other comments?</u>