The Report committee for Jason Robert Crandall

Certifies that this is the approved version of the following report:

Clickers and Classroom Engagement: The Impact of Audience Response Systems on Student Attentiveness and Engagement

APPROVED BY
SUPERVISING COMMITTEE:

Supervisor: ____________________________________
Daniel Robinson

____
Tiffany Whittaker

Clickers and Classroom Engagement: The Impact of Audience Response Systems on Student Attentiveness and Engagement

by

Jason Robert Crandall, B.A.

Report

Presented to the Faculty of the Graduate School

of the University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Master of Arts

The University of Texas at Austin

May 2011

Clickers and Classroom Engagement: The Impact of Audience Response Systems on Student Attentiveness and Engagement

by

Jason Robert Crandall, M.A.

The University of Texas at Austin, 2011

SUPERVISOR: Daniel Robinson

Student engagement is a critical component of effective classroom instruction. Many socio-constructive pedagogies, including active learning and peer instruction, depend upon students not only paying attention to, but actively shaping, the learning environment. Student response devices, such as clickers, are thought to increase student engagement by providing students with regular opportunities to check their comprehension or express their opinions. Claims of increased student engagement due to clicker use are often based upon student self-reports and have only a small correlation with observed learning gains or other measures of attentiveness. This paper compared self-report data, pre- and post-test scores, and a direct test of attentiveness to investigate what effect clicker use has on student engagement. Analyses showed that subjects using clickers were significantly slower to respond to tests of attentiveness than subjects in other conditions. This suggests that using clickers affects what students are able or willing to attend to during a lecture.

Table of Contents

List of Tables

List of Figures

Introduction

Although the tools and methods available to classroom instructors have changed considerably in the last few decades, the basic challenges remain the same. One of the most persistent challenges faced by instructors is reliably determining how much students are learning at a given point in the lecture. Failure to do so can result in too much time covering concepts that are well-understood or too little time covering concepts that are misunderstood. Either mistake will prevent instructors from using class time as effectively or efficiently as they could if they were able to reliably determine student comprehension during lectures. Quizzes, tests, free-writing, classroom discussion, and comprehension questions all provide some insight and their use is well-supported and well-documented. In addition to these traditional techniques, instructional technologies have been developed that can supplement and enhance classroom instruction.

One such technology is the student response system or “clicker” (alternately referred to as audience response system or classroom response system). Clickers allow a practically unlimited number of students to simultaneously and anonymously respond to questions during lectures. Typically, each student either purchases or is provided with a personal response device. Depending on the model, clickers can either be used to select from a range of multiple choice answers or to input numerical and text answers. These responses are instantaneously collected, calculated, and prepared for display, if the instructor chooses.

Such features can be collectively deemed “classroom affordances” and include anonymous feedback (when displayed to students; instructors can choose to record how students respond separately), real-time comprehension assessment, and instantaneous grading. Clickers have has also been linked to numerous pedagogical approaches, including team-based learning (Michaelsen, Knight, & Fink, 2004), think-pair-share (Lyman, 1981), agile teaching (Beatty, Gerace, Leonard, & Dufresne, 2006), peer instruction (Mazur, 1997), and others. (For a much more comprehensive list of pedagogical approaches used with clickers, see Bruff, 2009).

Of course, not all researchers are sanguine about the prospect of clicker use in the classroom. Michael Bugeja has published several pieces detailing concerns about clicker use. In a commentary for the *Chronicle of Higher Education*, Bugeja (2008) writes that:

> Ira David Socol, a scholar of technology in special ***education*** at Michigan State University, states, ‘The idea of wasting money on a device no more sophisticated pedagogically than raising your hand drives me nuts, whether it is students' money or the university's.’ Cellphones, he says, can perform the same tasks as ***clickers*** with more interactivity and less inefficiency. (para. 7, emphasis original)

Bugeja goes on to question the prevailing notion that, with regard to instructional technology, “if you can identify a benefit, you can justify the expense” (para. 19). He concludes that “professors need to realize that technology comes at a price, even when advertised as ‘free’” (Bugeja, 2008, para. 34).

Nonetheless, such concerns seem minor when compared to the sheer logistical effect of being able to gather so much data so quickly and so privately. Clicker use is rapidly becoming commonplace in large-scale classrooms, particularly those in Science, Technology, Engineering and Mathematics (collectively, STEM) disciplines. However, clicker use has also been studied in courses such as management information systems (Nelson & Hauck, 2008), accounting (Edmonds & Edmonds, 2008), library science (Corcos & Monty, 2008; Hoffman & Goodwin, 2006), psychology (Mayer, et al., 2008; Cleary, 2008), and communications (Jackson & Trees, 2007). As clicker use grows in popularity, it is increasingly important to understand if, and how, it affects student learning.

Integrative Analysis

This analysis will provide a basic overview of the pedagogical approaches that contributed to the development and application of clicker technology. Specifically, it will discuss how the theory of socio-constructivism has informed the development of student-centered and active learning classroom environments, as well as specific approaches such as peer instruction and think-pair-share. Attempts to apply these approaches in large classrooms illuminated the need for a tool such as clickers to facilitate rapid and reliable teacher-student interaction. Finally, the analysis will cover current clicker research and studies of student engagement in order to demonstrate the importance of better understanding the effect of clickers on engagement.

Theoretical Background

Much of the modern thinking on student-centered classrooms and socio-constructivism can be traced back to the writing of Lev Vygotsky, a Russian psychologist whose writings from 1925 to 1934 continue to influence modern educational psychology. Vygotsky was concerned with the relationship between speech and thought and the social nature of learning and wrote extensively on these topics. As Vygotsky's work was translated and introduced to Western academia during the 1970s, it began to influence classroom teaching, curricular theory, and developmental psychology.

Two of Vygotsky's concepts which are perhaps most relevant to the topics of student engagement and active learning are the Zone of Proximal Development (ZPD) and scaffolding (Vygotsky, 1978). The ZPD is defined as the range of tasks in a given learning domain that a learner can accomplish only with the measured assistance of someone who is more proficient in that domain. Vygotsky maintained that the majority of instruction should take place within this zone as the student will eventually learn to do these tasks alone and continually expand his or her ZPD into progressively challenging tasks. Furthermore, tasks within the ZPD are neither so difficult as to be demoralizing nor so easy as to be uninspiring.

A closely related concept is the process of scaffolding (Vygotsky, 1978). Scaffolding is the process by which experienced learners or instructors modify tasks that are too difficult for learners to complete unassisted. Such modifications can include completing the task together, simplifying the task, carefully guiding the student through each step of the task, or completing challenging portions of the task in advance before the student begins. In each case, the degree of scaffolding required is inversely proportional to the student's task ability; ultimately, the fully competent student should complete the task with no scaffolding. Tasks for which the student requires some degree of scaffolding are considered to be within the student's ZPD.

As mentioned above, Vygotsky's writings were highly influential in the development of socio-constructivism. Socio-constructivists believe that knowledge about the world is constructed through shared experience and that the meaning of learning artifacts can only be established through interaction among teacher, learner, and subject. One of the most important methods of such interaction is classroom discussion. Discussion is the simplest way to establish how much students understand, to create new knowledge about learning artifacts, to diagnose a learner's ZPD, and to determine the appropriate amount of instructional scaffolding. Pintrich (2003) writes:

> Many motivational theories, as well as cognitive theories (including Vygotskian models), stress the importance of providing tasks that are within the range of competence for students. The tasks should be neither too easy nor too difficult, but challenge students in appropriate ways (Brophy, 1999; Pintrich & Schunk, 2002). (p. 672)

An ideally socio-constructivist classroom would provide sufficient discussion to achieve all of these purposes.

Of course, most classrooms are far from ideal. Nystrand's landmark book, *Opening Dialogue: Understanding the Dynamics of Language and Learning in the English classroom*, investigated classroom discussion during a three-year period in over sixty classrooms with more than 2400 students (Nystrand, 1997). Nystrand found that instructors averaged three minutes of classroom discussion per sixty minutes of instruction. McKeachie discusses various barriers to classroom discussion, including "students' feeling that they are not learning", "the instructor's tendency to tell a student the answer" too quickly, instructor discomfort, and the difficulty of "appraising the group's progress", and "be[ing] aware of barriers. . . that are blocking learning" (McKeachie & Svinicki, 2006, pp. 44-45). Allowing true discussions is particularly challenging in large-scale classrooms, where only a minority of viewpoints can be entertained and some students are prone to monopolize discussion while others withdraw altogether.

Fortunately, student discussion is not the only way to incorporate socio-constructivist tenets into large-scale classrooms. In 1987, Arthur Chickering and Zelda Gamson compiled a list of principles for effective undergraduate education that remains useful today (Seven principles for good practice in undergraduate education). These principles particularly emphasize the frequency and quality of student-faculty contact and student-student contact. As Guerrero (2009) writes, “Each of these principles rests on the belief that students benefit from an instructor’s ability to design a learning environment that considers the students’ activity level, cooperation, diversity, expectations, interactions, and personal responsibility for learning” (p. 7). According to Chickering & Gamson (1987), good undergraduate teaching practice:

1. Encourages student-faculty contact
2. Encourages cooperation among students
3. Encourages active learning
4. Gives prompt feedback
5. Emphasizes time on task
6. Communicates high expectations
7. Respects diverse talents and ways of learning (pp. 1-2)

Each of these points warrants brief consideration in terms of their individual impact on creating a positive learning atmosphere and a student-centered classroom.

Student-faculty contact is described as the "most important factor in student motivation and involvement" (Chickering & Gamson, 1987, p. 1). Students who have frequent and diverse contact with faculty tend to be more involved and more successful than other students (McKeachie & Svinicki, 2006; Edmonds & Edmonds, 2008). One of the difficulties in a large-scale classroom is the limitation on the variety and frequency of student-faculty contact. Nearly all contact is unidirectional (from teacher to student) and is seldom personalized. Office hours, effectively moderated discussions, lab sessions, and interactions immediately before or after class can help address this difficulty.

In addition to contact with faculty, contact with other students can be highly beneficial to students. Chickering and Gamson write, "Good learning, like good work, is collaborative and social, not competitive and isolated" (1987, p. 1). Thus, effective undergraduate learning environments should seek to *encourage cooperation among students*. Various instructional methods, such as team-based learning (Michaelsen et al., 2004), depend heavily on the learning benefits derived from small-group interactions. Even a less formal approach, such as asking students to discuss questions with their neighbors prior to initiating class-wide discussion, allows students to create knowledge together and receive low-stakes feedback from their peers.

Active learning has been defined as "instructional activities involving students in doing things and thinking about what they are doing" (Bonwell & Eison, 1991, p. 1). Bonwell and Eison (1991) add: "To be actively involved, students must engage in such higher-order thinking tasks as analysis, synthesis, and evaluation" (p. 1). Chickering and Gamson (1987) claim that "students do not learn much just sitting in classes listening to teachers. . . . They must make what they learn part of themselves" (p. 1). Modifying large-classroom environments to allow active learning is central to the effective application of clicker technology and pedagogy and will be discussed further in the "Selected Clicker Research" section below.

Another core principle of student-centered learning is *giving prompt feedback*. Chickering and Gamson (1987) write: "Knowing what you do and don't know focuses learning. . . In classes, students need frequent opportunities to perform and receive suggestions for improvement" (p. 1). Such regular feedback helps both students and instructors to actively adapt the learning environment. Students who discover that they have sufficiently understood lecture material may be motivated to continue engaging in the course, while students who receive feedback that their understanding is incomplete may be motivated to seek additional contact with their peers or with the instructor. Clickers are especially well-suited to give the kind of regular and prompt feedback that facilitates adaptive teaching and learning in large classrooms.

Emphasizing time on task and *communicating high expectations* are linked to each other in that many instructors expect students to remain on task throughout course instruction. Chickering and Gamson (1987) argue that "how an institution defines time expectations for students, faculty, administrators, and other professional staff can establish the basis for high performance for all" (p. 2). In a roundtable discussion about clicker use at UT, James Bryant, a biostatistics lecturer at the UT School of Biological Sciences, explained how time on task and high expectations are related in his classrooms: “At first, when I started using clickers, I was worried that I run out of time to cover everything I wanted. But now, I find I can actually cover ever more material than before, because students know that we will be moving fast and they want to keep up” (forthcoming MERLOT grant project, July 8, 2009). Bryant often asks twelve to fifteen questions in a three-hour lecture and expects his students to respond to and understand each of them. As Chickering and Gamson (1987) write, “Expecting students to perform well becomes a self-fulfilling prophecy when teachers and institutions hold high expectations of themselves” (p. 2).

Finally, student-centered learning environments should *respect diverse talents and ways of learning*. The good practices described above “work for many different kinds of students—white, black, Hispanic, Asian, rich, poor, older, younger, male, female, well-prepared, under prepared. But the ways different institutions implement good practice depends very much on their students and their circumstances” (Chickering & Gamson, 1987, p. 3). What constitutes active learning for one student or one population of learners may not always be reflected by other students. Although clicker use may at first appear to be fairly monolithic, especially among instructors who require all students to respond to each question (typically through grade incentives), it can be used to promote diverse thinking and discussion in large-scale classrooms. For example, anonymous polls can allow students who hold minority viewpoints to determine how many peers share their position before they begin to overtly express themselves. Students who can see that their opinions are not completely unique may feel more comfortable contributing to class-wide discussions. Also, techniques such as peer instruction or think-pair-share expose students to viewpoints that may differ substantially from the instructor's perspective.

Selected Clicker Research

As demonstrated above, clicker technology can be used to facilitate the application of Chickering and Gamson's principles to classes of various sizes. This flexibility has made clicker technology the focus of numerous studies over the past fifteen years. Most investigators acknowledge that understanding clickers as an instructional tool requires a deep and sound understanding of their impact on students. As Wilbert McKeachie writes, "What is important [in teaching] is learning, not teaching. Teaching effectiveness depends not just on what the teacher does, but rather on what the student does" (McKeachie & Svinicki, 2006, p. 6). Further, in a literature review of various forms of computer-based or technology-supplemented instruction, Fitch (2004) concluded that "there is convincing evidence that interactivity is a critical part of any form of technology-based learning." Appropriately then, the majority of clicker research has focused on the interactivity of clickers and its effect on student learning, student recall, student affect, student motivation, and student engagement.

Caldwell and others have performed detailed literature reviews of the research on clickers (Caldwell, 2007). Caldwell describes numerous potential uses for clickers, including “to increase or manage interaction. . . , to assess student preparation and ensure accountability. . . , to find out more about students. . . , for formative (i.e., diagnostic assessment). . . , for quizzes or tests. . . , to do practice problems. . . , to guide thinking, review, or teach. . . , to conduct experiments on or illustrate human responses. . . , [and] to make lecture fun” (Clickers in the large classroom: Current research and best-practice tips, 2007, p. 11). It has been repeatedly suggested that clicker use is no more useful than the technique with which it is associated, and one clear benefit of clicker use is that it allows techniques such as peer instruction, think-pair-share, and agile teaching to be scaled up for use in large classrooms (see Bruff, 2009). However, simply adding clicker use to a traditional lecture without further modification may add some value; Wood (2004) writes that the “give-and-take atmosphere encouraged by use of clickers. . . makes the students more responsive in general, so that questions posed to the class as a whole during lecture are much more likely to elicit responses and discussion” (p. 798).

Some studies (Mayer, et al., 2008; Stowell & Nelson, 2007) have been experimental, classroom investigations of students in various clicker conditions, with controlled learning outcomes. Mayer described this approach as balancing “ecological validity” and “experimental control” (Mayer remarks, 10-2-2009). Mayer’s research was inspired, at least in part, by discovering the relative paucity of controlled, empirical studies regarding clickers. As he and his colleagues reviewed the existing literature, they determined that most, if not all, of the published studies available focused on student evaluation of clicker technology (Connor, 2009; Crossgrove & Curran, 2008; Duncan, 2005; Hatch, Jensen, & Moore, 2005; Hoekstra, 2008; Latessa & Mouw, 2005). As a result, Mayer and his colleagues conducted the longest, and possibly largest, controlled clicker experiment to date.

Mayer et al. (2008) performed a three-year long study of classroom learning in an educational psychology class. Each year had a distinct learning condition, which was used continuously throughout the entire class (Mayer, et al., 2008). One group (control group) completed the course without clicker technology or group questioning, a second (clicker group) completed the course with clickers, and the third used group questioning without clickers (no-clicker group). All students completed a pre- and post-questionnaires and the groups were roughly an equivalent proportion of upperclassmen, and proportion of women. The no-clicker group had a significantly higher mean SAT score that the others, but Mayer points out that “the group that we predicted would show the greatest learning (i.e., clicker group) did not have significantly higher SAT scores than the other groups” (Cllickers in college classrooms: Fostering learning with questioning methods in large lecture classes, p. 55). Students in both the clicker and no-clicker groups could earn up to 40 points for answering in-class questions (Mayer, et al., 2008).

The mid-term and final exam results for each group revealed that the clicker group had a mean score of 83.4%, compared with 80.3% for the control group and 80.2% for the no-clicker group. Further analysis “revealed that the clicker group outperformed the control and no-clicker groups, which did not differ from each other", which supports "the main prediction that the clicker-supported questioning method would improve academic achievement" (Mayer, et al., 2008, p. 55). These results are instrumental in establishing an empirical foundation for further research regarding the effect of clickers on learning outcomes.

Other researchers have conducted controlled experiments using clickers, but the majority of these have taken place in labs or other simulated conditions, unlike Mayer et al.'s classroom study. Stowell and Nelson (2007) conducted one such experiment using multiple classroom conditions in simulated psychology courses. Stowell and Nelson's study consisted of four groups: standard, hand-raising, flashcards, and clickers. Each group was given the same 30-minute psychology lecture and compared on both in-course participation rates and post-lecture self-reports. The researchers also compared student scores on in-course review questions and on a postlecture quiz. Stowell and Nelson were also interested in the academic emotions of their participants, as well as evidence of honesty or dishonesty.

On the post-lecture quiz scores, the researchers found no significant difference among the four groups. However, they did find that the clicker group's scores were significantly more consistent with their in-course review question scores (Stowell & Nelson, 2007, p. 255). For each of the other groups, the review question scores were somewhat or considerably higher than the post-lecture scores. Stowell and Nelson argued "that certain classroom feedback techniques have moderately large effects on honesty of student feedback and participation rates and small effects on academic emotions" (Stowell & Nelson, 2007). This finding suggests that clickers can be very useful in gathering accurate and thorough feedback, even if the learning gains are minimal.

Some researchers have questioned the need for clicker systems at all. Nick Lasry (2008) conducted a study comparing the effects of clicker use to that of a flashcard response system. Both groups of students used the Peer Instruction (PI) technique, briefly mentioned above. In essence, PI (Mazur, 1997) is a socio-constructivist approach that depends upon students to teach each other core course concepts. During PI, lectures are divided into a series of short presentations, followed by a conceptual question, termed a ConcepTest by Mazur (Crouch & Mazur, 2001). Typically, students are asked to consider the question, formulate an answer, and report their answer to the instructor, often using clickers. The instructor may choose to reveal these results to the class, and then always asks students to discuss their choices with each other. Crouch and Mazur (2001) write that "the instructor urges students to try to convince each other of the correctness of their own answer by explaining the underlying reasoning" (p. 970). After a few minutes of peer instruction, the student report revised responses to the instructor, who offers any final comments or explanation before moving to the next presentation.

For Lasry's study (2008), students were lectured briefly and then asked a multiple choice conceptual question. If fewer than 30% of students answered correctly, the instructor would revisit the concept and then allow students to revote. If 30%-80% of students were correct, students were asked to find a partner, discuss their choices, and then revote. Once at least 80% of students answered correctly, the lecture could proceed.

Students were randomly assigned to use either clickers or flashcards during the PI portions of each lecture. All students were given a pre- and post-test of course concepts. Both sets of students demonstrated significant learning gains during the course, but the two groups did not differ significantly from each other on either pre- or post-test performance. According to Lasry (2008), “this implies that PI is an effective instructional approach that is independent of the use of technology such as clickers” (Clickers or flashcards: Is there really a difference?, p. 243). He adds:

> From a *teaching* perspective, clickers have a number of very practical advantages: they allow instructors to get precise real-time feedback and store students’ responses. . . . From a *learning* perspective, using PI with clickers does not provide any significant learning advantage over low-tech flashcards. (Lasry, 2008, p. 244)

Lasry concludes that instructors should carefully consider the costs of employing clicker systems and whether or not similar instructional goals can be achieved through cheaper and simpler means.

Clearly, there is room to debate whether clickers themselves have any inherent instructional value and this is precisely the reason why controlled experiments to investigate the specific claims of clicker pedagogy are needed. One of the most basic—and universal—of these claims regards student engagement. As Stowell and Nelson (2007) wrote, "It might not be the experience of enjoyment (or any other emotion) that mediates the benefits of clickers, but rather the enhanced cognitive processing (attention) associated with it" (p. 256).

Clicker use at UT-Austin

Before investigating the different effects of clicker use on students, I wanted to understand how they were typically used in a classroom. My co-researchers and I obtained the names and course schedules of 48 instructors at the University of Texas at Austin who requested that clickers be available for their students through the University Coop. Eighteen of these instructors agreed to participate by completing a survey and allowing us to observe one classroom lecture. The survey was conducted at the University of Texas at Austin from September 26, 2009 through November 16, 2009.

The instructors who responded were from ten different departments, including biology, chemistry, classical civilization, computer science, economics, management, and physics. Eleven of the eighteen respondents were lecturers, one was an assistant professor and the remaining six were full professors. Altogether, the courses included represented 3,595 enrollees. However, some of these students may have been enrolled in more than one of these courses and the scope of our study did not allow us to obtain information about how many individual students were represented.

In order to understand faculty expectations of and experiences with clickers, we administered a survey to these faculty members and asked why they decided to use clickers and what impact clickers had on their classrooms. Their responses to the first question can be found in Figure 1. Eighteen total faculty members responded to the survey and each was permitted to select as many options as applied. Note that nearly 100% of the respondents indicated using clickers in order to increase both student learning and student engagement. Also, thirteen of the eighteen wanted to increase feedback for both the students and for themselves. Finally, it should be noted that "Increase Attendance" was not among the original response options provided for this question, but seven faculty members wrote it in. This suggests that a much higher percentage of respondents may have chosen this option if it had been initially available.

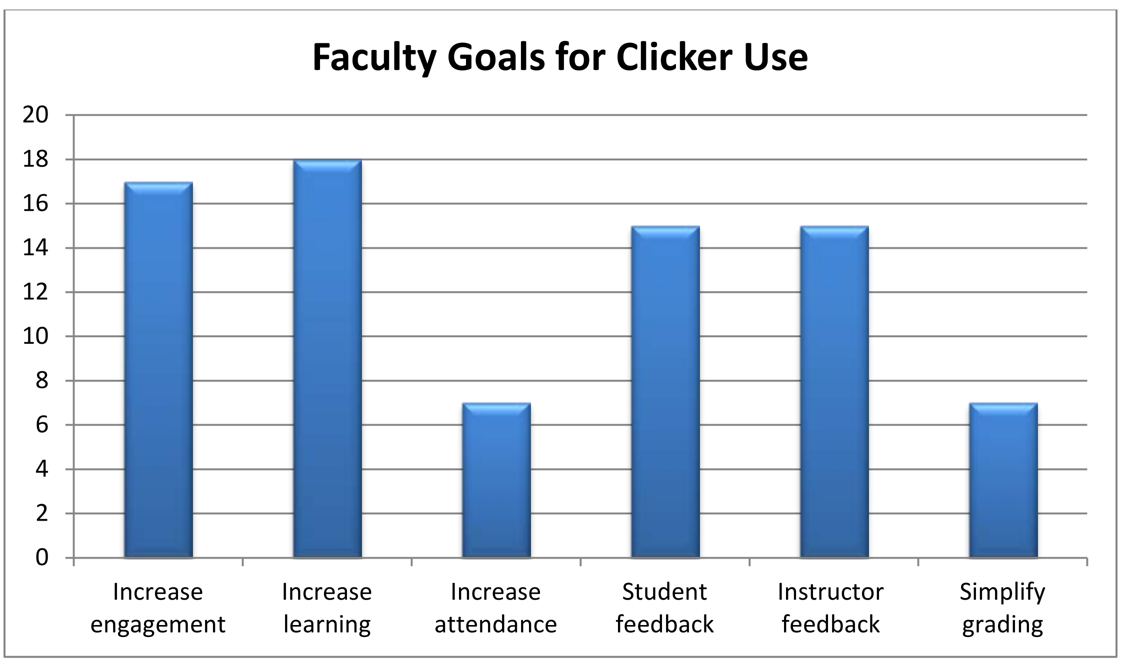

Figure 1. UT faculty goals for clicker use.

In addition to asking why faculty had decided to use clickers, we also asked how their classrooms had changed since they introduced clickers. The responses to this question can be seen in Figure 2. The respondents were asked to indicate how six classroom variables had changed since they began using clickers in their lectures. These variables were:

- Level of classroom community
- The instructor's ability to monitor student learning
- The student's ability to monitor their own learning
- Student learning
- Student motivation
- Student engagement

Two other variables—time spent preparing lectures and time spent grading—were included in the survey but are not displayed in Figure 2 because those variables refer to activities that occur outside of the classroom.

Respondents were asked to rate changes in classrooms they had taught without clickers and with clickers. Each respondent was required to rate the change in each category and only one response per question was permitted. Each of the variables was scored on a five-point scale from Decreased to Increased (see the legend in Figure 2 for all five points) and mean total response scores were calculated for each item. Out of the 108 total responses (six categories and eighteen respondents), zero were Decreased or Slightly Decreased, indicating that the instructors believed that each of these had remained the same or improved. Furthermore, only 7% of responses were Neutral and 33% were Increased, the highest response possible in our survey. Overall, 93% of possible responses indicated an increase in a desirable classroom condition.

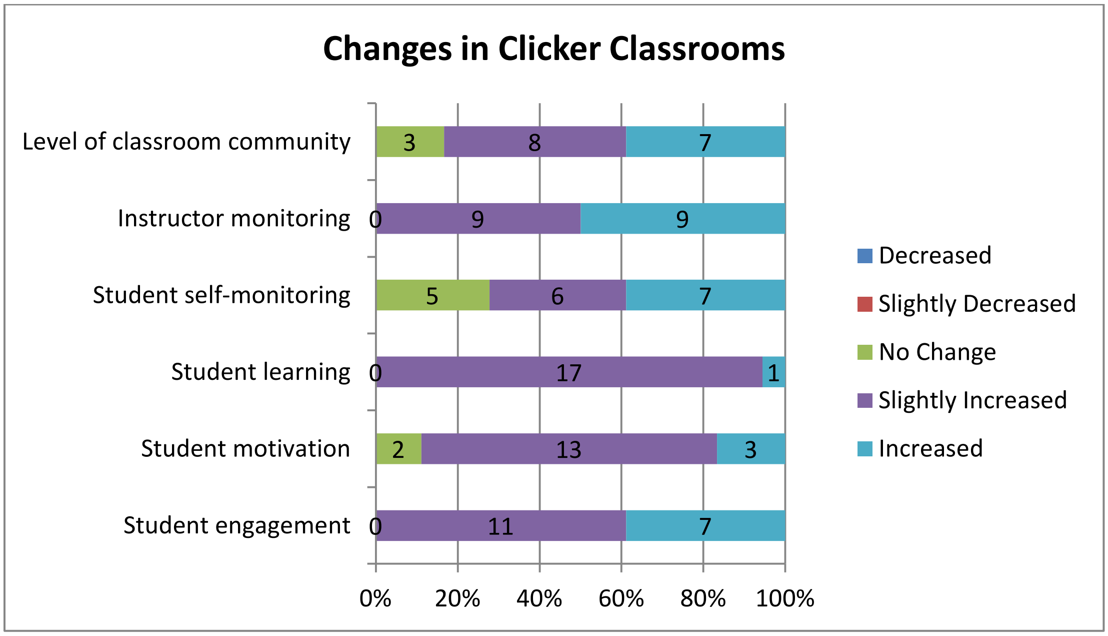

Figure 2. Changes in clicker classrooms at UT.

However, our survey did not reflect un-tempered satisfaction with clickers. A number of faculty members expressed frustration at the technical limitations of the devices and at issues relating to lost or malfunctioning clickers. Three of the respondents rated their clicker system as “Difficult” or “Somewhat difficult” to use. When asked how likely they were to recommend clicker use to a colleague, fourteen respondents said “Likely” or “Very likely”, but two said “Neutral” and one said “Very unlikely”.

In response to our open-ended questions, one instructor noted, “CPS [a clicker brand] has great promise but terrible implementation.” Another added, “I'm neutral as to the use of clickers. I like them as they enforce student attendance, but I get many complaints about the inability to make up missed quizzes." Still, the majority of the comments were positive, such as, "I should mention that I am hopelessly addicted to using clickers! They add so much to my classroom approach. I use the system to diagnose, to teach, to entertain, occasionally to test, you name it. It keeps the students as engaged as possible. I hope to never teach without a wireless system." Several other instructors mirrored this instructor's belief that clickers can "keep the students as engaged as possible", which has been a consistent theme of clicker research in general. It is precisely this sentiment that my study is intended to investigate.

Classroom Engagement

As Skinner et al. (2009) indicate, researchers from numerous fields, including assessment, teacher development, educational technology, and even neuroscience, have become increasingly interested in classroom engagement. One reason that engagement is viewed as a fertile field for further research is that it "represents a potentially malleable proximal influence shaping children's [and other students'] academic retention, achievement, and resilience" (Skinner, Kindermann, & Furrer, 2009, p. 494). Such malleability encourages intervention research from several different branches of educational psychology. Additionally, Svanum and Bigatti (2009), who observed 235 students for a five-year period, noted that "engaged students were more likely to succeed than their less engaged counterparts; succeeded more rapidly; and performed better than expectations based upon" GPA or admissions exam scores (p.

128). The interest in investigating engagement can be easily understood in terms of its openness to intervention and potential for meaningful impact.

Unfortunately, this surging interest has thus far failed to yield a singular understanding of engagement. While all researchers agree that being engaged consists of more than being physically present, there are divergent opinions regarding what else constitutes engagement. Is emotional involvement necessary? What about behavior? Must students commit participatory acts? In how far do attentiveness, participation, and engagement differ from or overlap with one another? Although it is beyond the scope of this study to attempt a definitive answer to any of these questions, I do believe that both emotional involvement and physical behavior are valuable components of engagement. I agree with Astin (1984) that engagement "refers to the amount of physical and psychological energy that the student devotes to the academic experience" (p. 297). While it may be theoretically possible for a student to be engaged without one or the other of these components, it is nonetheless my belief that typical student engagement results in or derives from both physical and emotional involvement.

Accordingly, Skinner, Kindermann, and Furrer (2009) envision engagement as having both an emotional and a behavioral axis. For each of the two axes, the positive affect is denoted as engagement and the negative aspect as disaffection, yielding four separate but highly interrelated measures of engagement: (a) behavioral engagement, (b) behavioral disaffection, (c) emotional engagement, and (d) emotional disaffection (Skinner et al., 2009, p. 500). Their self-report data includes items intended to measure each of these components. In addition, teachers can observe many of the behavioral components among students, such as focusing on the teacher, remaining on task during class-work, attending classes, and limiting conversation to course-related topics. These observations should generally correlate with the self-report.

Still, Skinner et al. (2009) acknowledge that it remains difficult to distinguish between on-task passivity and off-task passivity. Students who don't appear to be doing anything may be intensely considering the lecture or ignoring it altogether. Furthermore, increased awareness of cultural norms and differences has led to divergent interpretations of student engagement. Students from some cultures may hesitate to offer observations during class, not because they have failed to process the lecture, but because it is considered rude in these cultures to interfere with or interject during an instructor's presentation. And, as will be described in further detail below, there is some likelihood that students will report a greater level of engagement than they actually felt during the course. Thus, no single method of measuring student engagement is likely to be fully reliable.

Numerous studies and observations of standard lecture classrooms have reported lack of engagement among students, particularly in large-scale auditorium-style settings (McKeachie & Svinicki, 2006; Trees & Jackson, 2007). Cooper and Robinson (2000) write, "The large-class-lecture-centered approach seems to be inviting increasing degrees of student disengagement" (p. 6). This is particularly problematic in view of research suggesting that active learning, and even general success in college (Robbins, Lauver, Le, David, Langley, & Carlstom, 2004), are related to student engagement.

In addition, clicker use is almost always accompanied by some form of in-class review question, which encourages active learning. Stowell & Nelson's (2007) study suggested the following about in-class review questions:

> Regardless of the classroom feedback technique, in-class review questions will likely increase student participation and reduce boredom. If technologically and financially feasible, a good choice for getting honest feedback, increased participation, and possibly greater student enjoyment is an audience response system with clickers. (p. 257)

In fact, regardless of subject matter or pedagogy, almost all studies of clicker use have reported increases in student engagement, primarily as indicated by either student self-report, teacher observations, or attendance measures (Hatch, Jensen, & Moore, 2005; Draper & Brown, 2004; Hoffman & Goodwin, 2006; Beekes, 2006; Fitch, 2004). In most cases, the increased engagement has been attributed to the use of clickers rather than the inclusion of in-class review questions. Beatty (2004) reports the following:

> CCS [Classroom Communication Systems] classes are popular with students, and they can usually articulate why. They appreciate the system's value for engaging them in the material. They acknowledge that hearing other students' reasoning helps to clarify their own. They particularly like class-wide histograms: they like the reassurance that they're not alone when they're wrong, as well as the

> perception that they're part of a "community of learners" all struggling with the same ideas. (p. 6)

Indeed, Beatty's claims are consistent with the self-report data found by other researchers and a broad consensus has emerged regarding the feelings of students about their own engagement in clicker classrooms. However, relying solely on self-report data may obscure the effect of clickers on classroom engagement.

For example, many of the self-report questions have an obvious bias (after all, most students are unlikely to report that they were not paying attention). Certainly, the anonymity of the self-reports diminishes the social pressure to misreport, but it remains likely that students who wish to think of themselves as conscientious or diligent students will tend to over-report their own engagement. Additionally, students may not always be aware of their own cognitive processes. Pintrich (2003)writes that:

> There are many occasions when motivation and learning, in the classroom and in life in general, are not so conscious, intentional, and self-regulating. In research on cognition, there has been a great deal of research on implicit cognition where cognitive processing occurs outside conscious awareness and control. (p. 678)

Students may not be able to accurately recall or articulate how engaged they were in a lecture if some of their cognition "occurs outside conscious awareness".

Although descriptions of increased student engagement based solely on self-report data are inherently dubious and should be confirmed with additional, independent measures, this limitation should not be viewed as sufficient reason to dismiss or disregard self-report measures. Unconscious cognitive processes notwithstanding, this study agrees with Skinner et al. (2009) that “we do not assume that students know why they are motivated, but we do assume that students know whether they are motivated; that is, students are excellent reporters of their own engagement” (p. 496). Self-reports remain a necessary component of investigating the effects of technology on the classroom environment and on student engagement.

Thus, it is important to identify a reliable and valid measure of short-term student engagement to be adapted for use in a single lecture. Many useful measures are aimed at understanding engagement on a “macro” level. The National Survey of Student Engagement (NSSE), conducted by researchers at Indiana University, measures “whether an institution’s programs and practices are having the desired effect on students’ activities, experiences, and outcomes” (National Survey of Student Engagement, 2000, p. 1). However, as Handelsman et al. (2005) write, “the NSSE focuses on active learning and other educational experiences but does not focus on individual courses; rather, it assesses students’ overall perceptions” (p. 184).

In order to better assess short-term engagement, Handelsman et al. (2005) developed a scale for use during a single course. The Student Course Engagement Questionnaire (SCEQ) consists of 24 items, which load onto four primary factors. These factors have been deemed Skills Engagement, Emotional Engagement, Participation/Interaction Engagement, and Performance Engagement. Handelsman et al. validated the SCEQ by comparing to constructs previously associated with student learning and motivation.

They found that “all four of the SCEQ factors were associated with at least one other measure; the different patterns among the variables supported the distinctiveness of the student engagement factors” (Handelsman, Briggs, Sullivan, & Towler, 2005, p. 189). Additional reliability and validity measures confirm the usefulness of the SCEQ as a tool for measuring short-term student engagement. An adapted version of the SCEQ, intended for use after a single lecture, will be used in this study.

Furthermore, a number of engagement results are heavily dependent on student behaviors, such as class attendance or participation in discussions (Miller, Greene, Montalvo, Ravindran, & Nichols, 1996; Caldwell, 2007; Duncan, 2005; Bruff, 2009). Another study reported that students who incorrectly answered clicker questions paid more attention to succeeding questions and explanations (Rice & Bunz, 2006). While there is no question that clicker use positively affects the number of instances of participation and number of classes attended, it remains unclear how fully participation and engagement overlap. Thus, although clicker use is associated with outwardly visible signs of participation, more research is needed to determine the connection between these signs and the type of engagement needed for learning. During this study, investigators will also compare visible signs of student participation to self-report data gathered from the modified SCEQ.

This study will use the following definition of classroom engagement: Classroom engagement is the application of available cognitive, affective, and physical resources toward learning-related goals and activities. Engagement can be further divided into its cognitive, affective, and physical components, referred to here as attentiveness, interest, and participation. Attentiveness is typically a fixed quantity, reflected by neurological processes, and can be divided in several ways. Students are not typically fully aware of their own attentiveness and teachers often even less aware of what their students are paying attention to. Indirect measures appear to best suited for capturing attentiveness. Interest is an emotional quality and can be well-captured by self-report data such as surveys and questionnaires. Students are fully aware of their own interest, while teacher observation is only partially effective at revealing interest. Finally, participation includes physical activities such as engaging in classroom discussion, asking questions, completing classroom activities, and remaining on-task. Participation is transparent to both teachers and students. A full measure of engagement should account for all three components.

Statement of Purpose

This study is intended to demonstrate the effect of a technological intervention and the accompanying pedagogy on student attentiveness and engagement. It is also intended to support previously demonstrated associations between engagement and learning.

In order to better understand the connection between clicker use and engagement, this study will address the following questions:

RQ1: What is the effect on student engagement of asking comprehension questions during a lecture?

RQ2: What is the effect on student engagement of using clickers during a lecture?

RQ3: To the extent that either comprehension questions or clicker use during lectures increase student engagement, is there an associated effect on student learning?

Method

Subjects

The subjects were 142 college subjects (including 4 graduate subjects) who were members of the Educational Psychology subject pool at a large public research university in the Southwestern United States during the 2008-2009 school year. The subjects had a variety of different majors including biology, kinesiology, speech pathology and journalism. All subjects took both a pre-test and post-test administered in a lab setting. As shown in Table 1, 92% of the subjects were upper division or graduate subjects, and 59% were female.

Table 1

Comparison of three groups on demographic characteristics

Characteristics	Control group	Question group	Clicker group
N	50	46	46
Proportion of females	.55	.59	.61
Proportion of upperclass	.94	.93	.91

The experiment was conducted over two weeks in a computer lab in the university's education building. There were fourteen sessions altogether, with three to twelve subjects in each session. The subjects used an online scheduling system to sign up for sessions and the maximum session size was capped by the number of computers in the lab that faced the projector screen. The fourteen sessions were matched to achieve nearly equal overall group size, and the conditions were randomly assigned to each group. Here are the final group sizes and assignments: control group (5 sessions, n=50), question group (5 sessions, n=46), and clicker group (4 sessions, n=46). Sessions times were scheduled at various times of day to accommodate student schedules, with sessions as early as 9 a.m. and as late as 6 p.m.

Materials and Measures

All participants were given a pre-recorded lecture on Cisco® networking systems and seated at a computer workstation. Other materials included in-class questions, PowerPoint presentation slides, clickers, and a mock classroom. The measures were a pretest and posttest for the Cisco® material, a survey of student engagement, and the time elapsed for a secondary attention task. The materials and measures are discussed in greater detail below.

Materials

Cisco® pre-recorded lecture. Nihalani and Mayrath (2009) developed a lecture for a study on modal instruction. The lecture covers introductory material from Cisco®'s server and network installation training program. It consists of a twelve-minute video file, with narration and animation. Using highly technical material reduced, but did not eliminate, the likelihood that any of the subjects had high prior knowledge and limited the possibility of expertise as a confounding variable for the pre- and post-test learning measure.

Computer workstations. Each computer workstation had a monitor, keyboard, and mouse. The lab was equipped with 24 Core 2 Duo iMacs, with 24" screens, running in Windows. Since students whose monitors are perpendicular to the lecture presentation will likely respond more slowly to the screensaver program than those whose monitors are parallel, only the 12 monitors whose screens are easily visible when students are looking toward the front of the lab were used. The front of the lab had a projection screen, displaying the lecture and other experimental content from the experimenter's computer, as seen below.

Figure 3. Computer lab with workstations oriented toward front of room.

Clickers. iClicker devices were used for this study. iClickers are generally considered the simplest and most robust clicker system, although they do not offer as many input options as some other devices. The iClicker system consists of individual devices, instructor management software, and a receiver. The software and receiver were both preloaded onto the investigator's computer before the subjects arrived. Each computer workstation had one clicker transmitter and subjects were given brief instructions about turning on the device and inputting responses.

Lecture questions. The lecture questions were primarily knowledge and comprehension questions (Bloom, Engelhardt, Furst, Hill, & Krathwohl, 1956). Due to the relative brevity of the lecture, there were only three such questions, spaced throughout the presentation.

Both the clicker and question groups were shown the same three questions at the same intervals during the video lecture. The experimenter paused the video lecture and used PowerPoint to display the questions. Here is a sample question:

> Which of the following must be attached to executable software? (Correct answer bolded).
>
> a) **Virus**
> b) Zombie
> c) Trojan
> d) Worm

Engagement self-report instrument. The self-report instrument consisted of eight items adapted from Handelsman et al. (2005). Students in all three groups were asked to complete the survey. All items were Likert-type items, with five response options. Students were asked to read the items and select how strongly they agreed or disagreed with each.

The engagement self-report instrument was administered after the lecture, in conjunction with the post-test items. Here is a sample item:

	Strongly Disagree				Strongly Agree
1. I put forth effort during this presentation.	1	2	3	4	5

Additional details regarding engagement scale selection and construction are given below.

Post-test items. The post-test was administered through SurveyMonkey. Students were given a total of ten items regarding the video lecture. Here is a sample item:

> An e-mail bomb is a large quantity of bulk e-mail that overwhelms the e-mail server preventing users from accessing it. This is an example of what type of an attack? (Correct answer bolded).
>
> a) DNS poisoning
> b) Replay attack
> c) DDoS attack
> d) **DoS attack**

Measures

This study used two different measures of engagement: 1) self-report data from students, and 2) time elapsed in completing a secondary task, disguised as a computer screensaver. It will also include a pre- and post-test measure of learning based on the Cisco® lecture material.

The first measure has been has been used extensively in other research on clickers (Hoekstra, 2008; Armstrong, 2008; Deal, 2007; Kenwright, 2009; Herreid, 2006; Martyn, 2007). Although the limitations of self-report data have been previously noted, such measures remain an important source of information about student learning processes. In a book based on the proceedings of the 1996 National Institute of Health conference entitled *The Science of Self-Report: Implications for Research and Practice*, researchers in law, sociology, medicine, and psychology discuss both the limitations and necessity of self-report data (Stone & Turrkan, 2000). In a chapter entitled "Information No One Else Knows: The Value of Self-Report", Wendy Baldwin wrote: "Validating that certain information can be obtained through self-report is important because it opens up avenues of research where self-report data are the only data available" (Baldwin, 2000, p. 5). Thus, by comparing two independent measures of student engagement, this study may not only support previous research claims but also simplify and legitimize future lines of inquiry into clicker technology.

As mentioned previously, Handelsman et al. (2005) created the Student Classroom Engagement Questionnaire (SCEQ) to measure engagement for the duration of a single course. This 24-item scale showed adequate reliability and validity when compared with other measures, such as GPA, goal-orientation, and incremental vs. entity self-theories (Handelsman et al., 2005). Handelsman, et al. found that these items loaded onto four factors: Skills, Emotional, Interaction/Participation, and Performance (coefficient alpha values and factors loadings are reported in Table 2).

Although the SCEQ was deemed a useful measure of affective engagement in this study, a number of the specific items it contains refer to behaviors or attitudes that are unlikely to be present in a single lecture lab setting. For example, the Skills factor includes items such as “Doing all the homework problems” and “Looking over notes between classes to make sure I understand the material”. The Emotional factor includes “Thinking about the course between class meetings”, while the Participation/Interaction factor refers to getting good grades, helping classmates, and going to the professor’s office hours.

In order to adapt the SCEQ for this experiment, only items that could describe attitudes and behaviors during a single lecture were retained. Table 2 (on following page) shows the items that were retained, as well as which factor they originally loaded onto:

Table 2

Modified SCEQ with Factor Structure and Unmodified Factor Loadings.

Items (coefficient alpha)	Skills Factor (.82)	Emotional Factor (.82)	Part/Int Factor (.79)	Performance Factor (.76)
I put forth effort during this presentation.	.59			
I tried to answer each question fully.	.57			
I listened carefully to the presentation.	.51			
I tried to make the presentation material relevant to my life.		.86		
I found the presentation material interesting.		.54		
I had a strong desire to learn the material.		.43		
I enjoyed the presentation.			.57	
I believe that I under-stood the material.				.64

Note. Part/Int = participation/interaction.

Although efforts were made to mirror the original wording of these items, some changes were necessary (item 2, for example, originally referred to homework rather than questions and *presentation* was substituted for *course* throughout).

For the second measure, it was important that the task be plausible, worthy of completion, and not highly prioritized. Previous studies of secondary task attentiveness have employed measures such as responding to pre-determined instant messages, clicking on a screen prompt, or following a series of dots on a computer display while trying to memorize a list of words (Lavie, 2005). However, such tasks would be both too artificial and too distracting to serve as useful measures of engagement in the present study.

Therefore, this study used the following process: each student was seated at an individual workstation with a desktop computer and monitor. Three different conditions were present, as described in Table 3 below. All three groups viewed the videotaped Cisco® lecture. Those in the clicker condition were also given a clicker response device and asked to answer several questions during the lecture. The students in the control group and question group did not receive the clickers, and the control group did not have any in-course questions. With the exception of the in-course questions, the lecture used for all three conditions was identical.

Table 3

*Description of Experimental Conditions (Group differences in **bold**)*

Condition	Control	Question	Clicker
Secondary measure	Yes	Yes	Yes
Pre- and post-test	Yes	Yes	Yes
Survey	Yes	Yes	Yes
Clicker	No	No	**Yes**
In-course questions	No	**Unanswered**	**Answered**

At predetermined times during the lecture for the question group and the control group, the experimenter paused the presentation and displayed an in-course question. The experimenter read the question and multiple choice options aloud and asked the subjects to silently consider the correct answer. Subjects in the question group were not required to provide their answers to the experimenter or each other. The experimenter then asked if any of the subjects wanted to answer the question and, whether a subject responded or not, explained the correct answer to each question as well as each of the incorrect options.

In the clicker condition, subjects were each given a personal response device. Because the analysis did not depend upon subject-level data, the clickers were not registered to individual subjects. In addition, only the clicker group generated responses to these in-class questions, so response data for these questions were not gathered for comparison with the other subject groups. However, the questions were posed immediately following the relevant concept, and the overwhelming majority of subjects were able to answer them correctly. For each question, the experimenter paused the lecture video, and said, “Here’s a clicker question. Please use your clickers to select the best answer.” The experimenter then presented a PowerPoint slide with a question and four multiple choice answer options (a, b, c, d) to the subjects and read the slide aloud before activating the clicker response software.

Because the clicker sessions had a maximum of twelve students each, the experimenter waited until all subjects had responded and ensured that each subject participated in the intervention for each question. After collecting all of the responses, the experimenter deactivated the clicker software and explained the correct answer and incorrect options to the subjects prior to resuming the lecture video.

In addition, all three groups were asked to complete the secondary task and were given the following instruction: “The computer in front of you will be recording data once you log in. During the lecture, you will not need to use the computer, so the screensaver may appear. If you see that the screensaver comes on, please hit a key so that the system will resume.” A screenshot of the screensaver is provided below (see Figure 4 below).

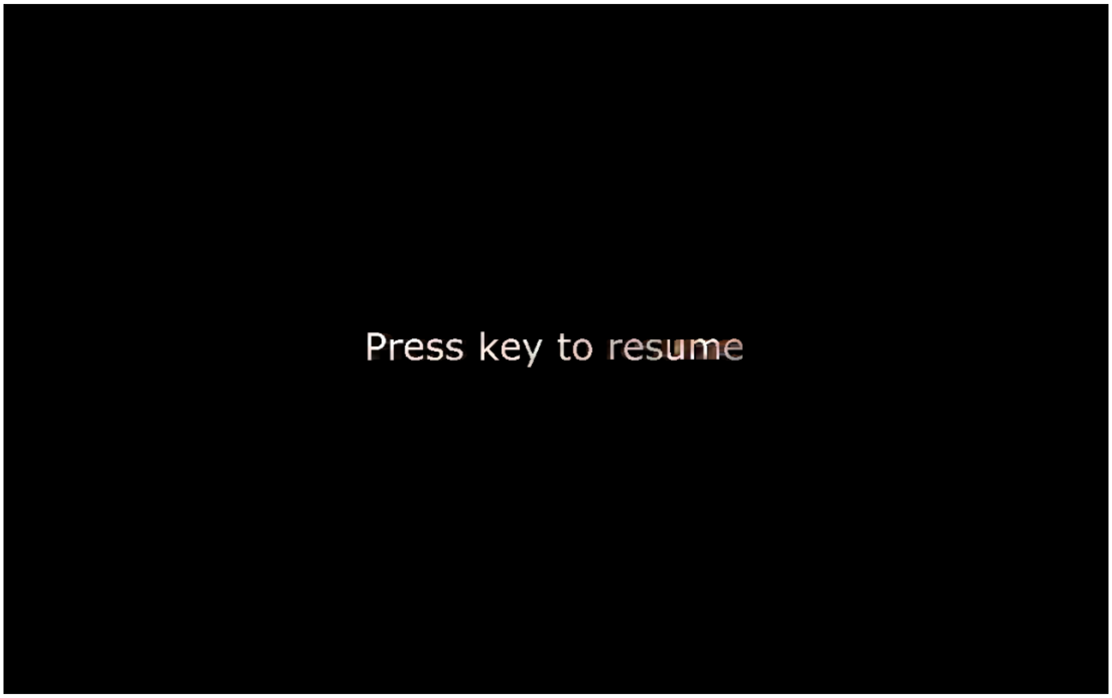

Figure 4. Screensaver.

The screen shown in Figure 4 was displayed three times throughout the lecture. When the screensaver appeared, a timing program began and recorded the amount of time needed for each student to press a key. The screensaver intervals were timed so as to not appear during the in-course question portions of the experiment for the question and clicker groups. The mean response time for all students in each of the three conditions was recorded as a secondary measure of attentiveness.

Previous research on dual task studies has reinforced the importance of attention for learning (Lavie, 2005; Rees, Frith, & Lavie, 1997; Yi, Woodman, Widders, Marois, & Chun, 2004; Nissen & Bullemer, 1987). Such research supports the use of simple tasks and response time measures as a proxy for the cognitive processes of engagement and attentiveness. However, in these cases, no attempt was made to disguise the nature of the tasks. Subjects were simply asked to give full effort to each task, and measures of response time were analyzed to determine how multiple processes interfered with one another.

Because this research does not concern whether or not attention is important to learning but whether or not classroom techniques involving clickers can help increase purposeful attention, disguising the secondary task was deemed appropriate. In this instance, the screensaver can be considered a measure of distraction rather than attention. It was hypothesized that students who are highly engaged in the primary task (i.e., listening to and answering or considering questions about the lecture) would be slower to attend to the distraction. This is consistent with a cognitive science hypothesis that tasks with high perceptual load (i.e., heavy demands on a person's attention processes) "leave no capacity for perception of task-irrelevant stimuli" (Lavie, 2005, p. 65). This hypothesis has been supported by numerous studies of distractor effects and perceptual load (Rees, Frith, & Lavie, 1997; Yi, Woodman, Widders, Marois, & Chun, 2004)

Thus, where $\mu_{control}$ = mean response time (in sec) for the control group with no in-course questions and no clickers, $\mu_{question}$ = mean response time (in sec) for the question group with in-course questions and no clickers, and $\mu_{clicker}$ = mean response time (in sec) for the clicker group with in-course questions and clickers, the hypothesized response rates across all three conditions were as follows:

$$H_0: \mu_{control} = \mu_{question} = \mu_{clicker}$$

$$H_1: \mu_{control} < \mu_{question} < \mu_{clicker}$$

That is, a true null result would indicate no difference in response time among the three groups, or one that is contrary to the hypothesis that clicker use increases student engagement. Such results would include any inequality in which the control group takes the same amount or longer to respond to the secondary task than the question or control groups.

Alternately, it is believed that increasing levels of participation in the main lecture (physically responding to questions with clickers compared to considering questions without clickers compared to passively receiving lecture instruction) should demand increasing levels of student attentional resources. In turn, these demands should reduce students' ability to respond to distractions such as the screensaver. Thus, it is anticipated that the control group will respond most quickly, followed by the question group, and finally the clicker group, and that those results will correlate with the self-report data gathered after the lecture.

However, there is a sentiment among some researchers that clicker use itself has no educational benefits compared to previously existing instructional methods (Lasry, 2008; Hatch, Jensen, & Moore, 2005; Stowell & Nelson, 2007). They suggest that most gains are due to the teaching pedagogies associated with or enabled by clicker use, and not clickers themselves. Recent research by Jones, Guerrero, Crandall and Robinson supports this conclusion (unpublished observation, 2009). This study compared unit test scores for undergraduate students in an upper-division Educational Psychology course.

For each of two units, different halves of the class were given clickers and offered the opportunity to earn bonus points by answering in-course questions. The other half of students who were present heard the same lecture and questions, but did not actively answer the questions. Not surprisingly, there were no significant differences in text outcomes between clicker and non-clicker students.

Because it is generally difficult to separate the impact of the technology from the pedagogy, the current study limits differences between the question and clicker conditions. The vast majority of clicker studies have reported increases in student engagement, regardless of instructional technique, subject matter, or student population (Rice & Bunz, 2006). Since it is unlikely that each of the various techniques employed account for all of the increased engagement independent of clicker use, these results suggest that clicker use itself is associated with increased student engagement. To help test this hypothesis, both the clicker and question groups reviewed the questions for the same amount of time and both groups were given a brief explanation of the correct answer. Thus, any increase in engagement among the clicker students should not be the result of a generally more captivating classroom environment.

Conversely, the hypothesized results of the post-test questions are as follows:

Post-test score (difference in # questions correct):

$$H_0: \mu^*_{control} = \mu^*_{question} = \mu^*_{clicker}$$

$$H_1: \mu^*_{control} < \mu^*_{question} < \mu^*_{clicker}$$

where $\mu^*_{control}$ = the adjusted mean post-test score for the control group, $\mu^*_{question}$ = the adjusted mean post-test score for the question group, and $\mu^*_{clicker}$ = the adjusted mean post-test score for the clicker group. Of course, just as in the response time analysis, this null hypothesis yields other alternate hypotheses. However, this study's primary focus is on student engagement and the simplified implementation of the clicker technology makes learning gains unlikely. Therefore, any outcome other than the proposed alternate hypothesis will be considered a failure to reject the null hypothesis and will support the hypothesis that clicker use alone does increase student learning (Bugeja, 2008; Lasry, 2008; Duncan, 2005).

The primary research question here is whether clicker use alone compels students to pay more attention and, ultimately, be more engaged. If the use of in-course questions and clickers is positively associated with attentiveness, engagement, and active learning, then students using clickers to answer in-course questions should be slower to respond to distractions and have higher post-test scores than students who either hear in-course questions without responding to them or observe a standard lecture with no questions and no clickers.

Results

The researcher conducted a single factor ANOVA to analyze differences in learning between groups. The F-test of between group differences on the post-test was non-significant. The average post-test scores for the control and question groups were 5.96 and 5.93 questions correct out of ten, respectively. The clicker group scored 6.28 on the post-test. Pre- and post-test difference scores were also non-significantly different between all three groups (Control – 2.6, Question – 3.28, Clicker – 3.26).

However, a single factor ANOVA test of response time to the attentiveness measure, averaged across all three trials for each subject, did show a main effect of group membership. The mean response time for the Control group was 17.07 s (*SD* = 56.53) with a 95% CI [7.91, 23.81]. The Question group had a mean response time of 16.41 s (*SD* = 53.84) with a 95% CI [7.12, 25.7], while the Clicker group's mean response time was 82.76 s (*SD* = 108.86) with a 95% CI [63.97, 101.54]. The omnibus test of the main effect of group type on response time means was statistically significant at the .05 level, $F(2, 398) = 32.39$, $p < .001$. A Tukey post-hoc contrast indicated that those in the clicker group took significantly longer to respond than those in each of the other groups, neither of which differed from one another.

The size of the difference between these groups indicated potential data problems that required further analysis. It was determined that these results were heavily influenced by a number of trials in which subjects completely failed to respond to the screensaver. These trials were recorded as a maximum possible value of just over four minutes (250 s). Of the 401 total recorded trials (excluding missing values), forty-eight recorded a value of at least 230 s. Due to some variations in the timing program, it is likely that all forty-eight of these trials indicate no response and the timer simply reset when the subsequent screensaver attempt began. Another nine trials elapsed at least 120 seconds before the subject disabled the screensaver. Because the vast majority (80%) of the trials concluded in 10 seconds or less, these fifty-seven trials were considered extreme outliers ($Z > 2.5$) and exercised disproportionate influence on the mean response times for each group.

In order to mediate the impact of those outliers on the group comparison, the data were reanalyzed with these values set to 120906 ms (Z =1.005), the least of the moderately extreme response times. The outlier adjusted response times and standard deviations are reported in table 4 below.

Table 4

Outlier Adjusted Means and Standard Deviations for Response Time across Groups

	Pooled	Control group	Question group	Clicker group
Mean (s)	21.79	9.98	9.98	45.97
Standard deviation (s)	42.19	28.96	26.91	55.11

Note. Outliers greater than 230 s were adjusted to 120 s before recalculating descriptive statistics.

The omnibus test of the main effect of group type on outlier adjusted means remained statistically significant at the .05 level, $F(2, 398) = 38.31$, $p < .001$. Tukey post-hoc contrasts indicate that the clicker group differed from each of the other groups, which did not differ from one another. Although adjusting the outliers did reduce the disparity between the groups, the clicker subjects still took, on average, significantly longer to respond to the screensaver program than the subjects in either the Control or Question groups.

One further adjustment was attempted based on the experimenter's observation of a certain behavior in two sessions (one control session and one clicker session). In each of these sessions, a cluster of subjects left their computers unattended throughout the entire lecture. Because subjects could observe computers adjacent to and in front of their own, it is hypothesized that a subject in the front row who did not respond to the screensaver may have inadvertently influenced subjects in the immediate vicinity to ignore the screensaver, as well. The experimenter removed two sessions where several adjacent students (at least 40% of the total subjects in the session) timed out on all three screensaver trials. The session adjusted statistics are reported in Table 5 (on following page).

Table 5

Session Adjusted Means and Standard Deviations for Response Time across Groups.

	Pooled (n=120)	Control group (n=43)	Question group (n=43)	Clicker group (n=34)
Mean (s)	24.49	7.3	16.41	52.77
Standard deviation (s)	64.52	30.55	53.84	93

Note. Sessions in which at least 40% of trials yielded no response were removed due to concerns about potential social influences. New group sizes reported in column headings.

Even with these sessions removed, the main effect of group type on session adjusted means was still statistically significant at the .05 level, $F(2, 356) = 16.64$, $p < .001$. The Tukey post-hoc test of contrasts indicated that all three groups differ significantly from one another. The control group's mean response time was now significantly faster than the Question group.

Due to the unexpectedly high number of subjects who failed to respond to the screensaver, a Chi-Square analysis was conducted to determine if group membership had a significant impact on likelihood of failing to respond. Six subjects ignored the screensaver once, five subjects did so twice, and fourteen more subjects ignored it for all three trials. Of the twenty-five subjects who timed out at least once, 3 were in the Control group, 5 in the Question group, and 17 in the Clicker group. The percentage of subjects who timed out did vary by group membership, $\chi^2 (2, N = 25) = 14.79$, $p < .001$.

Discussion

Limitations

Two significant limitations of this study are its lack of ecological validity and its brevity. As mentioned previously, clicker technology allows instructors a number of affordances in lecture classrooms, including instant feedback and anonymity. In addition, it allows instructors to scale up techniques such as peer instruction and team-based learning for use in large classrooms. However, this study took place in a mock classroom and the investigators will not use any of the pedagogical techniques such as peer instruction or think-pair-share. Such limitations may decrease the effect size of group placement and learning condition.

Additionally, previous studies of student engagement have lasted a minimum of one semester. Many studies have been two or more years. This study consisted of one half-hour lab session per subject. Accordingly, the SCEQ has been adapted and may require a new item analysis to determine if the item loadings remain sufficient. However, in spite of these limitations, this study is positioned to make valuable contributions to our current understanding of the relationship between clicker systems and student engagement. If there are significant findings under these limited conditions, it is likely that the findings would be more pronounced in more ecologically valid conditions or over longer periods.

Interpretation

The experimenter expended considerable effort interpreting the possible causes of the non-responsive subjects because of their impact on the data. It is possible that these subjects either misunderstood or disregarded the study instructions, although 83% of subjects (121 out of 146) responded as expected to all three trials. Another explanation is that some of the subjects determined that the attentiveness program was not a genuine screensaver. Although the image used resembled typical screensaver text and the program's initial appearance mimicked a genuine screensaver, the attentiveness program used a static image rather than a screensaver animation. Also, the program could not be disabled by mouse movement because the associated timer did not respond properly to dynamic input. While these features may have caused some subjects to question the authenticity of the screensaver and to alter or even cease their responses, such responses are likely to have been randomly distributed across experimental conditions. Further, pre-test responses to the computer familiarity item indicated no significant prior differences between groups in computer knowledge, so it is unlikely that students in one group were significantly more apt to conclude that the screensaver was not genuine [control = 3.56 (n=50), question = 3.43 (n=46), clicker = 3.23 (n=46)]. In fact, the clicker group had the lowest reported familiarity score and was the most likely to not respond to the screensaver.

It is also possible that the Clicker group was delayed in responding to the screensaver stimulus because these subjects were holding the clicker device when the screensaver appeared. Indeed, the experimenter observed many clicker subjects who had to first put down the clickers before either clicking the mouse or a button on the keyboard. However, this process lasted at most a few seconds and cannot account for the minimum thirty-second difference in mean response time between the Clicker group and the remaining groups. Absent any other apparent group difference, at least some of disparity in group response times must be attributed to the use of clickers itself.

Thus, the use of clickers, even during a single session and with a very simplistic pedagogical implementation, does affect student engagement. In addition to creating additional opportunities for student participation, clicker use also appears to be associated with a decreased ability to attend to secondary, non-educational tasks. However, the educational value of this finding remains unclear. First, it remains to be determined if the effect observed here can be extended to true diversions and whether or not clicker use decreases student susceptibility to classroom distractions. Further, even if clicker use is shown to decrease the effect of classroom distractions, perhaps students simply begin to pay more attention to the device or the physical process of answering questions, without processing lecture material more deeply or thoroughly. In this study, clicker students seemed to be much less engaged in the screensaver activity, but it is not obvious that these students were more engaged in the lecture. Although clicker students had both the highest post-test scores and self-reported engagement scores, neither difference was statistically significant.

Future Research

Future research should include a calibration of the dual-task measure and the intervention itself. One way to help verify that the screensaver measure does, in fact, measure attentiveness would be to perform a similar experiment with a control group and a single intervention independently demonstrated to increase student attentiveness during a single class lecture. Alternately, the screensaver measure could be used together with another attentiveness measure—such as eye-movement tracking—to confirm that both measures correlate. Use of such procedures could provide additional empirical foundation for interpreting this study's results as support for a positive association between clicker use and student attentiveness.

Another possible extension of this study's method would be to present the screensaver measure as a performance goal rather than an avoidance goal. Rather than asking subjects to respond to the screensaver in order to avoid a negative outcome for the experimenter (i.e., loss of data), the experimenter could tell subjects that they are being measured on how quickly they are able to respond to the screensaver, as well as how well they answer questions about the lecture. This procedure may indicate if using clickers inhibits multi-tasking in general or simply deters subjects from attending to relatively non-salient tasks.

Conclusion

Instructors in large classrooms must adapt their teaching to various constraints, particularly the challenge of assessing student learning with limited student-faculty interaction. Various instructional methods and technologies have been developed to help teachers alleviate these challenges. One promising technology is the clicker, which allows teachers to receive anonymous, instantaneous feedback about student comprehension and allows students to answer questions, express opinions, and practice problem-solving. Many observational and experimental studies of clicker use have found little or no gain in student learning but moderate gains in student engagement.

This study operationalized a definition of student engagement and comparing several independent measures, finding significant differences in student attentiveness among students who used clickers during a lab presentation. Investigator observations of student behavior also provided anecdotal support for experimental findings. As these measures reflected higher student engagement among students using clickers, similar measures may be used for long-term studies or for investigations of more sophisticated pedagogical techniques. Ultimately, such studies will help us better evaluate the impact of clickers or similar technologies on teaching and learning in large classrooms. However, in the absences of further research to interpret clickers’ effect on student engagement and learning, this study’s findings cannot be used to offer clear recommendations for the academic use of clickers.

References

Armstrong, D. (2008). Clickers in the classroom. *CDHA Journal*, 32.

Astin, A. W. (1984). Student involvement: A developmental theory for higher education. *Journal of College Student Personnel, 25*, 297-308.

Baldwin, W. (2000). Information no one else knows: The value of self-report. In A. A. Stone, J. S. Turkkan, C. A. Bachrach, J. B. Jobe, H. S. Kurtzman, & V. S. Cain (Eds.), *The Science of Self-report: Implications for Research and Practice* (pp. 3-8). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Beatty, I. D. (2004). Transforming student learning with classroom communication systems. *EDUCAUSE Research Bulletin*, pp. 2-13.

Beatty, I. D., Gerace, W. J., Leonard, W. J., & Dufresne, R. J. (2006). Designing effective questions for classroom response system teaching. *American Journal of Physics*, 31-19.

Beekes, W. (2006). The 'Millionaire' method for encouraging participation. *Active Learning in Higher Education*, 25-36.

Bloom, B. S., Engelhardt, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). *Taxonomy of educational objectives: The classification of educational goals. Handbook 1: Cognitive domain.* (B. S. Bloom, Ed.) New York: David McKay.

Bonwell, C. C., & Eison, J. A. (1991). Active learning: Creating excitement in the classroom. *ERIC Digest*, 1-4.

Bruff, D. (2009). *Teaching with Classroom Response Systems: Creating Active Learning Environments.* San Fransisco: Jossey-Bass.

Bugeja, M. (2008). Classroom clickers and the cost of technology. *Chronicle of Higher Education*.

Caldwell, J. E. (2007). Clickers in the large classroom: Current research and best-practice tips. *CBE-Life Sciences Education*, 9-20.

Chickering, A. W., & Gamson, Z. F. (1987). Seven principles for good practice in undergraduate education. *AAHE Bulletin*, 3-7.

Cleary, A. M. (2008). Using wireless response systems to replicate behavioral research findings in the classroom. *Teaching of Psychology*, 42-44.

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences.* New York: Academic Press.

Connor, E. (2009). Perceptions and uses of clicker technology. *Journal of Electronic Resources in Medical Libraries, 6*(1), 19-32.

Cooper, J. L., & Robinson, P. (2000). The argument for making large classes seem small. In J. L. Cooper, & P. Robinson (Eds.), *New Directions for Teaching and Learning* (pp. 5-16). San Fransisco, CA: Jossey-Bass.

Corcos, E., & Monty, V. (2008). Interactivity in library presentations using a personal response system. *Educause Quarterly*(2), 53-60.

Crossgrove, K., & Curran, K. L. (2008). Using clickers in nonmajors- and majors-level biology courses: Student opinions, leaninrg, and long-term retention of course material. *CBE Life Sciences Education, 7*(1), 146-154.

Crouch, C. H., & Mazur, E. (2001). Peer Instruction: Ten years of experience and results. *American Journal of Physics, 69*(9), 970-977.

Deal, A. (2007, November 30). *Researching on Teaching with Technology - Enhancing Education - Carnegie Mellon University.* Retrieved March 19, 2009, from Carnegie Mellon University Teaching with Technology: http://www.cmu.edu/teaching/technology/research/index.html#crs

Draper, S. W., & Brown, M. I. (2004). Increasing interactivity in a lectures using an electronic voting system. *Journal of Computer Assisted Learning, 20*, 81-94.

Duncan, D. (2005). *Clickers in the classroom: How to enhance science teaching using classroom response systems.* San Fransisco: Pearson/Addison-Wesley.

Edmonds, C. T., & Edmonds, T. P. (2008). An empirical investigation of the effects of SRS technology on Introductory Managerial Accounting students. *Issues in Accounting Education, 23*(3), 412-434.

Fitch, J. L. (2004). Student feedback in the college classroom: A technology solution. *Education Technology Research and Development*, 71-81.

Furrer, C., & Skinner, E. (2003). Sense of relatedness as a factor in children's academic engagement and performance. *Journal of Educational Psychology*, 148-162.

Guerrero, C. (2009). *Classroom response systems: Using clickers and collaboration to foster learning in a postsecondary setting.* UT Austin: Unpublished prospectus.

Handelsman, M. M., Briggs, W. L., Sullivan, N., & Towler, A. (2005). A measure of college student course engagement. *Journal of Educational Research, 98*(3), 184-191.

Hatch, J., Jensen, M., & Moore, R. (2005). Manna from heaven or clickers from hell: Experience with an electronic response system. *Journal of College Science Teaching, 34*, 36-39.

Herreid, C. F. (2006). "Clicker" cases: Introducing case study teaching into large classrooms. *Journal of College Science Teaching, 36*(2), 43-47.

Hoekstra, A. (2008). Vibrant student voices: Exploring effects of the use of clickers in large college courses. *Learning, Media, & Technology, 33*(4), 329-341.

Hoffman, C., & Goodwin, S. (2006). A clicker for your thoughts: Technology for active learning. *New Library World, 107*(1228), 422-433.

Jackson, M. H., & Trees, A. R. (2007, March). The learning environment in clicker classrooms: Student processes of learning and involvement in large university-level courses using student response systems. *Learning, Media, and Technology, 32*(1), 21-40.

Kenwright, K. (2009). Clickers in the classroom. *TechTrends*, 74-77.

Lasry, N. (2008). Clickers or flashcards: Is there really a difference? *The Physics Teacher*, 242-44.

Latessa, R., & Mouw, D. (2005). Use of audience response system to augment interactive learning. *Family Medicine, 37*, 12-14.

Lavie, N. (2005). Distracted and confused?: Selective attention under load. *Trends in Cognitive Science, 9*(2), 75-82.

Lenth, R. V. (2008, August 27). *Russ Lenth's power and sample size page*. Retrieved December 16, 2009, from http://www.cs.uiowa.edu/~rlenth/Power/

Lyman, F. (1981). The responsive classroom discussion: The inclusion of all students. In A. S. Anderson (Ed.), *Mainstreaming Digest* (pp. 109-113). College Park: University of Maryland.

Martyn, M. (2007). Clickers in the classroom: An active learning approach. *Educause Quarterly*, 71-74.

Mayer, R. E., Stull, A., DeLeeuw, K., Almeroth, K., Bimber, B., Chun, D., et al. (2008). Cllickers in college classrooms: Fostering learning with questioning methods in large lecture classes. *Contemporary Educational Psychology, 34*, 51-57.

Mayrath, M., Nihalani, P. K., & Robinson, D. H. (n.d.). *Instructional technology delivery methods: Evidence of learner's Preference Paradox.* unpublished manuscript.

Mazur, E. (1997). *Peer Instruction: A Users' Manual.* Upper Saddle River, NJ: Prentice Hall.

McKeachie, W. J., & Svinicki, M. (2006). *McKeachie's Teaching Tips.* Boston: Houghton Mifflin Company.

Michaelsen, L. K., Knight, A. B., & Fink, L. D. (Eds.). (2004). *Team-based learning: A transformative use of small groups in college teaching.* Sterling, VA: Stylus.

Miller, R. B., Greene, B. A., Montalvo, G. P., Ravindran, B., & Nichols, J. D. (1996). Engagement in academic work: The role of learning goals, future consequences, pleasing others, and perceived ability. *Comtemporary Educational Psychology*, 388-422.

National Survey of Student Engagement. (2000). *The NSSE report: National benchmarks of effective educational practice.* Bloomington: Indiana University Center for Postsecondary Research and Planning.

Nelson, M. L., & Hauck, R. V. (2008). Clicking to learn: A case study of embedding radio-frequency based clickers in an introductory management information systems course. *Journal of Information Systems Education, 19*(1), 55-64.

Nissen, M. J., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology*, 1-32.

Nystrand, M. (1997). *Opening dialogue: Understanding the dynamics of language and learning in the English Classroom.* New York, NY: Teachers College Press.

Pintrich, P. R. (2003). A motivational science perspective on the role of student motivation in learning and teaching contexts. *Journal of Educational Psychology, 95*(4), 667-686.

Rees, G., Frith, C. D., & Lavie, N. (1997). Modulating irrelevant motion perception by varying attentional load in an unrelated task. *Science*, 1616-1619.

Rice, R. E., & Bunz, U. (2006). Evaluating a wireless course feedback system: The role of demographics, expertise, fluency, competency, and usage. *Studie in Media and Information Literacy Education*.

Robbins, S. B., Lauver, K., Le, H., David, D., Langley, R., & Carlstom, A. (2004). Do psychosocial and study skills factors predict college outcomes? A meta-analysis. *Psychological Bulletin, 130*(2), 261-288.

Skinner, E. A., Kindermann, T. A., & Furrer, C. J. (2009). A motivational perspective on engagement and disaffection: Conceptualization and assessment of children's behavioral and emotional paricipation in academic activities in the classroom. *Educational and Psychological Measurement, 69*(3), 493-525.

Stevens, J. P. (2007). *Intermediate Statistics: A Modern Approach.* New York, NY: Lawrence Erlbaum Associates.

Stone, A. A., & Turrkan, J. S. (2000). Preface. In A. A. Stone, J. S. Turkkan, C. A. Bachrach, J. B. Jobe, H. S. Kurtzman, & V. S. Cain (Eds.), *The Science of Self-report: Implications for Research and Practice* (pp. ix-xi). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Stowell, J. R., & Nelson, J. M. (2007). Benefits of electronic audience response systems on student participation, learning, and emotion. *Teaching of Psychology, 34*(4), 253-258.

Svanum, S., & Bigatti, S. M. (2009). Academic course engagement during one semester forecasts college success: Engaged students are more likely to earn a degree, do it faster, and do it better. *Journal of College Student Development*, 120-132.

Trees, A., & Jackson, M. (2007). The learning environment in clicker classrooms: Student processes of learning and involvement in large university-level courses using student response systems. *Learning, Media and Technology, 32*, 21-40.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes.* Cambridge, Massachusetts: Harvard University Press.

Wood, W. B. (2004). Clickers: A gimmick that works. *Developmental Cell*, 796-798.

Yi, D.-J., Woodman, G. F., Widders, D., Marois, R., & Chun, M. M. (2004). Neural fate of ignored stimuli: dissociable effects of perceptual and working memory load. *Nature Neuroscience*, 992-996.