AC 2009-878: A REVISITED STUDY ON THE USE OF CLICKER TECHNOLOGY TO EVALUATE SHORT-TERM CONCEPT RETENTION

Adam Czekanski, United States Military Academy

ADAM J. CZEKANSKI is an instructor in the Department of Geography and Environmental Engineering at the United States Military Academy at West Point, New York. He teaches introductory courses in environmental science, environmental engineering, and hydrogeology. Mr. Czekanski's academic and research interests include engineering education and drinking water treatment in developing nations. Mr. Czekanski is a licensed professional engineer in the Commonwealth of Virginia.

David Roux, United States Military Academy

DAVID-MICHAEL P. ROUX is an instructor in the Department of Geography and Environmental Engineering at the United States Military Academy at West Point, New York. He teaches introductory courses in environmental science, environmental engineering, and air pollution control. Mr. Roux's academic and research interests include engineering education and physicochemical/biological treatment in environmental engineering.

A Revisited Study on the Use of Clicker Technology to Evaluate Short-Term Concept Retention

Abstract

The use of electronic student response systems such as "clickers" is a relatively new and innovative means by which instructors can gain real-time feedback on student comprehension. We examined the use of clickers with 3rd-year undergraduates to determine if the anonymity provided by the device influenced student responses, while also assessing the effects of the time gap between exposure to the material and testing. We accomplished the objectives in an introductory environmental science course using multiple-choice questions focused on key course concepts. Two instructors participated in this classroom experiment. One instructor taught four sections of the course (n = 60) and divided his sections by those who answered the review questions using the clickers (n = 31) and those who answered by raising their hands (n = 29). A second instructor was assigned two sections (n = 29) and had students respond to the review questions by using the clickers. Previously, we evaluated short-term concept retention by administering questions to the students at the end of specified lectures. For this study, we modified our experimental protocol by administering the questions at the start of the next lecture in order to determine the extent to which this added time gap would influence student responses. We hypothesized that the lack of anonymity attributed to the hand raising approach can greatly skew results by allowing a few intelligent students to sway the majority and give the instructor a false impression of class comprehension. We further hypothesized that as the duration in time between exposure to the material and testing is increased, the accuracy of results from those not afforded anonymity will decrease. Our study showed that time duration between teaching and testing of material has a significant effect on concept retention and reinforced the value of clickers for accurate short term concept evaluation. Instructors can quickly assess their students' current understanding of the course material and make adjustments to the course based on those results.

Key words: clickers, retention, active learning

Introduction

Educators have always struggled with the challenge of finding innovative ways to accurately assess student comprehension of material in the classroom. Methods historically used by instructors have included calling on a broad range of students to answer questions, having the entire class answer questions through visible means, or using volunteers[1]. Although each of these methods encourages migration toward a more active classroom learning environment, none truly give the instructor an accurate picture of how well or how poorly the class as a whole is grasping recently taught concepts. This downfall is primarily due to issues with the sample size inherent to these methods, namely that the small minority can inherently mislead the instructor into believing that the majority either understands or misunderstands the concept being questioned[1]. Without the use of periodic quizzes or deliberately advertised examinations, the

instructor has historically lacked a means to accurately assess the proficiency of the entire class. Electronic student response systems, or "clickers," are interactive remote answering devices that offer instructors a means to gain this accurate real-time assessment of student comprehension, and students the ability to assess their own grasp of material[2]. Both students and teachers live in a society that is accustomed to instantaneous gratification. Education is not immune to this desire and access to a tool that enables educators to obtain immediate accurate feedback on student proficiency is almost a necessity in today's classroom. Clickers are a vehicle for assessing student understanding of learning objectives, but their growing popularity in the classroom in recent years has also been linked to their role in supporting active learning[3]. The effective integration of clickers into the traditional lesson plan has proven to be an excellent means to capture the attention of students, while also giving them a stake in the material being covered. Students who commit to an answer tend to become invested in the question and develop an increased interest in the discussion surrounding that question[1]. The critical benefit offered by clickers over other methods historically used to query classroom populations is their ability to provide a level of anonymity to the student in answering a question posed in public. They allow students to provide input without fear of embarrassment in front of their peers and without having to be concerned about being lost in discussion surrounding the topic[4]. Without this anonymity, many students are likely to refrain from contributing input due to evaluation apprehension and conformance pressure[5]. This is likely to be true in a classroom of any size, but the effect is amplified in the smaller classroom environment due to the more threatening nature of the intimate setting. This work seeks to investigate the effectiveness of clickers to evaluate short-term concept retention in a small classroom environment, and to assess whether the anonymity they provide plays a significant role in student response to review-based questions. The influence that a time gap between exposure to material and subsequent testing may have on the accuracy of clicker evaluations is also analyzed.

Methodology

The short-term retention of critical course concepts taught in an introductory environmental science course designed for non-environmental science or engineering majors at the junior level was evaluated. To accomplish this, we designed a semester-long experiment in which we could quantitatively capture and evaluate student retention of concepts recently learned. This study involved students divided among six sections, each with an average size of approximately fifteen students. Two instructors participated in this classroom experiment. One instructor taught four sections of the course (n = 60) and divided his sections by those who answered the review questions using the clickers (n = 31) and those who answered by raising their hands (n = 29). A second instructor was assigned two sections (n = 29) and had students respond to the review questions by using the clickers. The course was organized into four blocks of instruction (I through IV), each representing major course themes. The review questions were presented in nine (of 40) scheduled 55-minute lectures and were used to assess the students' grasp of course material presented during the lessons. Three review questions were presented in a multiple-choice format at the beginning of the lesson immediately after the material was initially covered. Each question had four answer choices. Review questions were based on the critical learning objectives and designed to interact with the clickers using Turning Technologies TurningPoint® 2008 plug-in software for Microsoft® Office PowerPoint® 2007. Student responses to the clicker

questions were recorded using a Turning Technologies TurningPoint ResponseCard® radio frequency (RF) wireless response system (Figure 1) or by raising hands.

Figure 1. *Turning Technologies TurningPoint ResponseCard® RF wireless response system. The system includes a universal serial bus (USB) receiver, 25 response cards, installation CD, and an instruction packet (not shown).*

Answering by using the clickers and raising hands were the methods used to capture student responses to the review questions. To assess the impact of response method on answer choice, we fixed the response method for each section. Treatment populations were defined by the aforementioned methods of response. In order to account for the difference in scholastic ability among the populations, we computed the students' average cumulative grade point average (CGPA) (Table 1). Short-term concept retention was measured by using the clicker software to compute the percentage of a section that selected an answer choice or by counting the number of hands raised for each answer choice. These percentages were compiled on a data table that allowed us to gauge quickly how well our students understood concepts presented to them during lectures and to determine if the anonymity provided by the clickers caused a notable difference in the percentage of each section that selected the correct answers. We noted the number of days that elapsed between when the material was covered and when the review questions were asked to determine if time affected how well our students performed on the review questions. We concluded this study by developing a column chart that compares the average review question scores for the clicker and raising hands populations (Figure 2) as well as a two-axis line chart that compares the average performance of each population on the review questions against the amount of time that elapsed between when the course material was covered and review questions answered (Figure 4). A student t-test was used to determine if the response method influenced the percentage of the clicker and raising hands populations that obtained the correct answer on the in-class review questions.

Results

This study sought to determine if the students were able to retain critical course concepts recently presented to them and to see if the condition of anonymity (as provided by the clickers) gave the instructors an accurate assessment of their students' level of comprehension. The average CGPA of the students in each statistical population (Table 1) gave us an initial impression of their scholastic ability.

Table 1. *Average cumulative grade point averages (CGPA) for statistical populations in the introductory environmental science course.*

Population	Cumulative Grade Point Average (CGPA)
Clicker (n = 60)	2.855
Raising Hands (n = 29)	3.073

The percentage of each treatment population that answered the review questions was computed. The population raising hands obtained a higher percentage of students obtaining the correct answer on the review questions. Approximately 91.4% of the population raising hands answered the review questions correctly, compared to 67.1% of the clicker population (Figure 2).

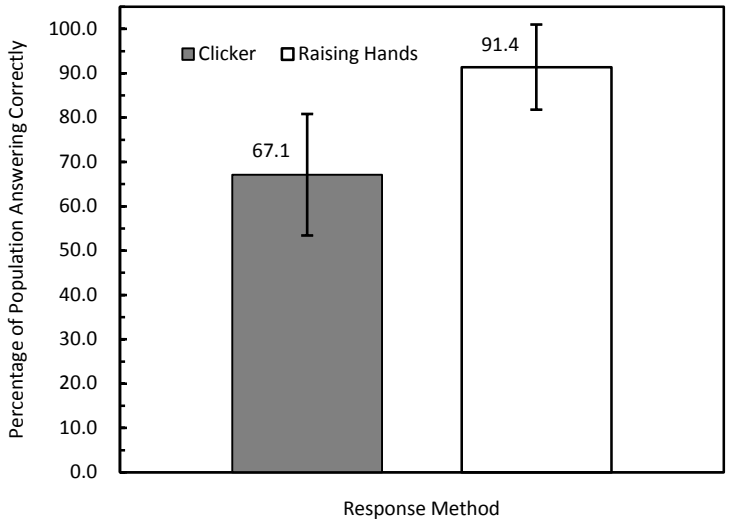

Figure 2. *Percentage of the raising hands and clicker populations that answered the review questions correctly for the introductory environmental science course. Data was averaged over nine lessons. The standard deviations from the mean for the raising hands population (n = 29) was 9.6% and 13.7% for the clicker population (n = 60).*

A two-sample student t-test was performed on the treatment population scores to test our hypothesis that the type of response method influences the percentage of students selecting the correct answer on the in-class questions. We obtained a $t_{statistic}$ of 5.567, which is greater than both the one- and two-tailed $t_{critical}$ of 1.860 and 2.306, respectively (Table 2). This indicates that, at a 95% confidence interval, we may statistically support the assumption that the method of response affects a population's average score on the in-class review questions.

Table 2. *Student t-test results for the raising hands and clicker population in-class review questions from nine lessons in the introductory environmental science course. The means represent the percentage of the statistical population that answered correctly on the in-class review questions.*

	Clickers	Hands
Mean	67.1%	91.4%
Standard Deviation	13.7%	9.6%
Observations	9	9
Degrees of Freedom (df)	8	
$t_{statistic}$ (absolute value)	5.567	
P(T<=t) one-tail	0.000	
$t_{critical}$ one-tail	1.860	
P(T<=t) two-tail	0.001	
$t_{critical}$ two-tail	2.306	

In our previous work, the review questions were asked at the end of the lesson in which the material was covered[9]. For this revisited study, one of the additional parameters we introduced was a time separation between when the material was covered to when we asked our students to answer the review questions. We scheduled our review questions at the beginning of the lesson after the material was covered. The time separation ranged from two to eight days depending on the academic calendar. Although time separations between lessons differed from lesson to lesson, the overall average amount of time separation for each population was approximately 4.7 days over the 40-lesson course.

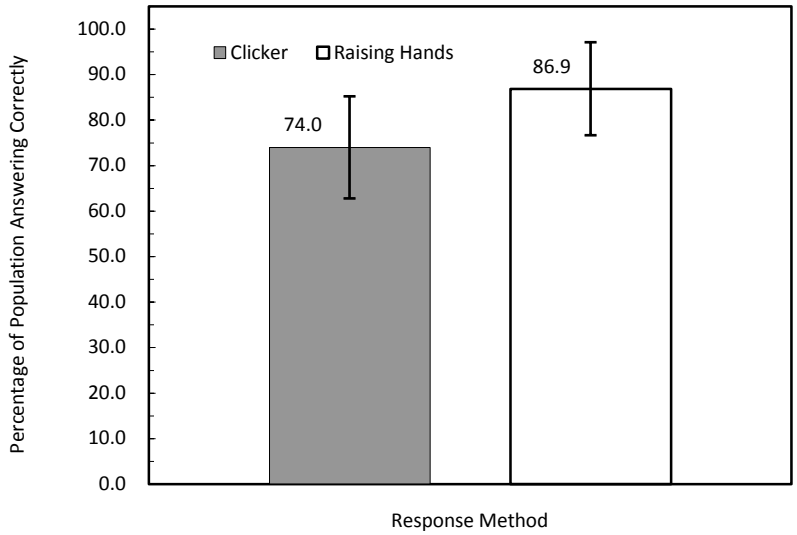

Figure 3. *Percentage of the raising hands and clicker populations that answered the review questions correctly for the introductory environmental science course from our previous work[9]. Data was averaged over ten lessons. The standard deviations from the mean for the raising hands population (n = 48) was 10.2% and 11.2% for the clicker population (n = 48).*

Comparison of Figures 2 and 3 shows that the performance by the clicker population was notably worse (by approximately 6.9%) when the review questions were presented at the start of the following lesson. The hand-raising population, on the other hand, performed better than our previous study by approximately 4.5%. As a result, we developed another chart to visualize and compare the average performance on the review questions to the time separation. Figure 4 represents a plot of the average percentage of questions answered correctly by block of instruction. The time separation (elapsed time) was plotted on a secondary vertical axis to give a visual comparison between the time separation and average performance on the review questions. The hand-raising population consistently scored higher on the review questions regardless of the amount of time that elapsed between when the material was covered in class and when the questions were presented. The performance of the clicker population remained fairly consistent at 67% answering correctly in each block of instruction.

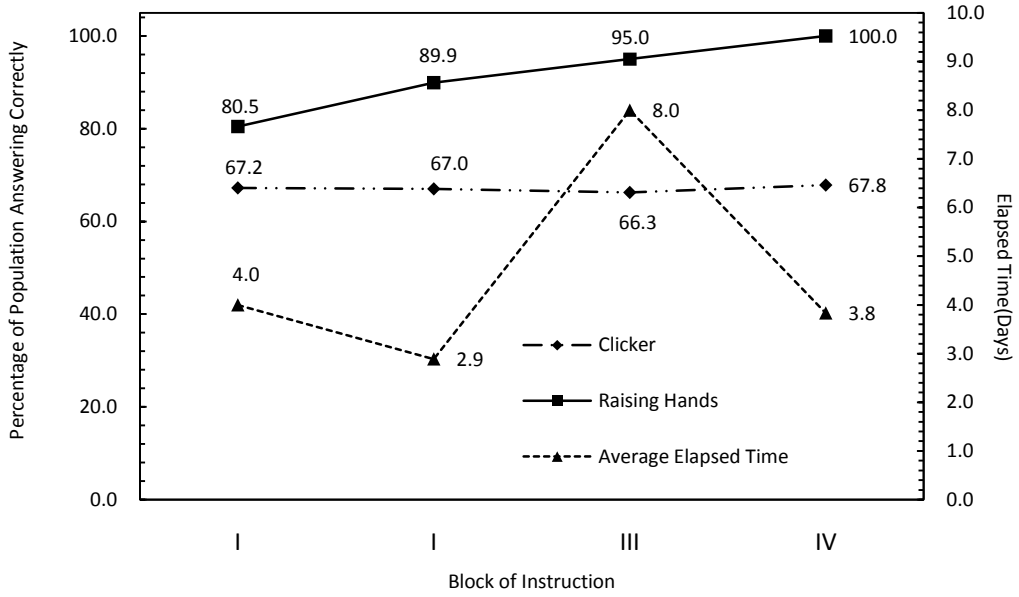

Figure 4. *Comparison of performance on the in-class review questions for the statistical populations in each of the four blocks of instruction (percentage of population answering correctly) to the average elapsed time (days).*

We accounted for the difference in scholastic ability among the statistical populations by comparing the CGPAs listed on Table 1 with the average percentage of correctly answered questions on the review questions (Figure 2). The hand raising population, on average, scored significantly better on the review questions and also had a slightly higher CGPA (3.073) than the clicker population (2.855). This does not, however, indicate any objective link between performance on the review questions and CGPA simply because of the difference in the method of response. Using the average CGPA of our populations only provides us an idea of how well each population would perform on the review questions relative to one another.

Discussion and Conclusions

The results from both semesters of data collection clearly illustrate a correlation between percentage of correct answers on lesson review questions and anonymity provided to the student. The provision of anonymity enabled the student to place self-doubt and fear of embarrassment aside in order to answer according to his/her own understanding[6, 7]. The results from the populations using clickers as their response method were reflective of what one would expect from a class containing students of varying capabilities and levels of comprehension. Conversely, the results of the populations using hand-raising as their response method were excessively inflated. The lack of anonymity amongst these populations inaccurately skewed the data dramatically in favor of the correct response (see Figures 2 and 3). Although success rates in the ranges of those experienced by the hand-raising populations in these studies would be welcome by any teacher, they do not accurately reflect actual student comprehension of the tested concept. Instead, they reflect the preferred action of the average student when the correct answer is in doubt: copy those students in the classroom viewed as being intelligent. The instructor is left with a false sense of teaching accomplishment and no way of knowing the reality of the situation until the next quiz or examination. The success rates of the clicker populations are undoubtedly more sobering, providing the instructor with an accurate real-time picture of student comprehension of recently taught concepts. As valuable as this knowledge may be, the more critical aspect is the *opportunity* given to the instructor to take appropriate action once the gravity of the situation is realized. An educator experiencing success rates hovering near or above 90% on review questions will not likely feel the need to cover the tested topic any further. In his mind, the class has proven their general mastery of the concept and those who did not answer correctly are viewed more as the exception than the norm in the classroom. The accurate breakdown given by the clickers exposes the truth, as ugly as it is at times, and provides the "red flag" of reality that lets the instructor know that maybe he should take a few extra minutes not only to reveal the correct answer to the class, but to refresh everyone on the details of the concept linked to the question[8]. The clicker section success rates of 67.1% and 74.0% (see Figures 2 and 3, respectively) witnessed in our study clearly afford the educator much greater cause to review the tested concept to ensure misunderstandings are cleared up before moving on to more advanced topics.

Although the clicker was clearly proven to be an effective means to evaluate short-term concept retention, we sought to further analyze the effectiveness of the technology by introducing a greater time gap between initial exposure of material and testing. Our original study adhered to the extreme definition of "short-term" concept retention by administering testable review questions immediately after material was taught during the same lesson. The current study maintained questioning in the "short-term" timeframe, but extended the timeframe between exposure and testing to a minimum of 48 hours. The theory that knowledge is perishable over time is not groundbreaking in nature and we did not intend to prove what is already known. Our intent was to illustrate *how much* or *how little* the influence of anonymity is amplified when the tested topics are no longer immediately present in the minds of the student. The decrease in overall success by the clicker populations from 74.0% to 67.1%, and the corresponding *increase* in success rates among the hand-raising populations from 86.9% to 91.4%, validated our initial hypothesis regarding the effects of introducing this measureable lag. As a time gap is introduced between exposure to material and testing on that material, students clearly forget specific

information related to the topic. The clicker populations accurately depict this trend as average student comprehension dropped by nearly seven percentage points between the two studies with the introduction of this minor time lag. The hand-raising population also initially appears to validate the trend with the accuracy in answers increasing by nearly five percentage points. This is in compliance with the findings of our initial study and reflects the trend for students to simply side with those intelligent students in the classroom whose answers they can clearly see when anonymity is not afforded. As classroom material becomes fainter in the memories of students, more are left with the decision to answer according to their own abilities, or answer with "the crowd". The results of both studies clearly show what decision most students make when faced with this dilemma. The impact of natural intelligence of the populations may come into question when analyzing these results, but the average CGPA data in Table 1 clearly shows this parameter to be negligible. Although there is a difference in average CGPA of the two populations in this most recent study, the hand-raising population enjoys a mere advantage in average CGPA of two tenths of a grade point over the clicker population. This is in no way a sufficient discrepancy to explain the 27 percentage point advantage of that population on review questions in this study.

The effect of introducing some time gap between teaching and testing material is apparent, but we also wanted to look at how well-defined this effect was with an increasing time gap. Figure 4 depicts average success by both populations, broken down by block of instruction. This success is also depicted with the average time gap for the questions administered. The results appear to show that increasing time incrementally does not have a significant effect on the success rates experienced by either population. The hand-raising population saw a difference in success rates of nearly 20 percentage points between the first and fourth blocks, even though the time lag in both of these blocks was nearly identical (around four days). The clicker population maintains a steady success rate between 66-68% for time gaps ranging from 2.9 to 8.0 days. Although this is initially discouraging, the data suggests more of a trend than is initially apparent. We have previously established that the clicker data is *accurate real-time* data indicative of student comprehension of tested material. The fact that the success rates of our clicker population remained relatively constant throughout the range of time gaps analyzed, while also being measurably lower than values from our initial study, suggests that there is an *initial* decrease in material retention immediately following initial exposure that is soon followed by a leveling off of comprehension. The material or concepts retained are likely to be those receiving the most attention during the lesson in which they were taught. Some people will fail to retain these critical concepts initially, but a majority of the class will retain these concepts for some period of time as reflected by our data. At some future point not identified in this survey, that comprehension will likely drop off. The fact that the success rates for the hand-raising population were so variable can be explained by the nature of the testing method. The hand-raising method is devoid of anonymity and therefore provides anyone in that particular class with the means to "copy" someone else's answer. This does not, however, mean that an individual must *always* copy the intelligent students and answer as they are answering. Instead, it simply means if students lack the confidence in a particular answer, they are provided with a safety net to avoid potential embarrassment. Students feeling confident enough can answer as freely as they desire. Although having the option to answer freely offers a more accurate depiction of comprehension, it also provides potential variability in overall trends for student responses while using the hand-raising method. This builds a certain degree of error into the hand-raising testing method and produces results as we see in Figure 4 that do not necessarily follow a clean trend.

The clicker itself is neither a tool to directly teach concepts, nor is it meant to replace quality lesson preparation and planning. The clicker is a powerful tool to augment and enhance active learning in the classroom, and is most importantly a means to provide accurate situational awareness to the instructor[1,7]. The accuracy of this feedback is further enhanced as the time between teaching and testing of the material is increased. Although populations not afforded anonymity failed to demonstrate an identifiable trend in success rates when analyzed against an incremental time lag, those populations with anonymity demonstrated consistent success. These consistent success rates lend further credibility to the accuracy and utility of the clickers as an evaluation tool, as well as to their importance in providing the educator with the critical feedback necessary to effectively plan and teach. Future study will focus on more clearly identifying the trend of material retention over a well-defined period of time. The issue of *qualitative* versus *quantitative* questioning will also be explored in more depth. We touched on this topic with our initial study and intended to develop it further, but the topic would have been too broad to have covered in conjunction with time-lag analysis.

References

1. Caldwell, J. (2007). Clickers in the Large Classroom: Current Research and Best-Practice Tips. *Life Sciences Education*, 6, 9-20.

2. Petr, D. (2005). Experience with a Multiple-Choice Audience Response System in an Engineering Classroom. *Proceedings of the 35th ASEE/IEEE Frontiers in Education Conference*, 19-22 October 2005, Indianapolis, Indiana.

3 Auras, R., and Bix, L. (2007). WAKE UP! The Effectiveness of a Student Response System in Large Packaging Classes. *Packaging Technology and Science*, 20, 183-195.

4. Martyn, M. (2007). Clickers in the Classroom: An Active Learning Approach. *EDUCAUSE Quarterly,* 2, 71-74.

5. Lowry, P., Romano, N., and Guthrie, R. (2006). Explaining and Predicting Outcomes of Large Classrooms Using Audience Response Systems. *Proceedings of the 39th IEEE Hawaii International Conference on System Sciences*, 4-7 January 2006, Kauai, Hawaii.

6. Siau, K., Sheng, H., and Nah, F. (2006). Use of a Classroom Response System to Enhance Classroom Interactivity. *IEEE Transactions on Education*, 49(3), 398-403.

7. DeBourgh, G. (2007). Use of Classroom "Clickers" to Promote Acquisition of Advanced Reasoning Skills. *Nurse Education in Practice*, in press.

8. Siegel, J., Schmidt, K., Cone, J. (2004). INTICE – Interactive Technology to Improve the Classroom Experience. *Proceedings of the 2004 American Society for Engineering Education Annual Conference*, 20-23 June 2004, Salt Lake City, Utah.

9. Czekanski, A. and Roux, D. (2008). The Use of Clicker Technology to Enhance Short- and Long-Term Concept Retention. *Proceedings of the 2008 ASEE Zone 1 Regional Conference*, 28 – 29 March 2008, West Point, New York.