Inaugural Conference on Classroom Response Systems:
Innovations and Best Practices, 11/2008, Conference Proceedings

Personal Response Systems in Statistics: Using Clickers to Foster Active Learning and Address Student Understanding of Statistical Inference.[1]

Dr. Jennifer J. Kaplan, Department of Statistics and Probability,
Division of Science and Mathematic Education
Michigan State University
Dr. Mark Urban-Lurain, Division of Science and Mathematics Education
Michigan State University

Overview of the Courses for Which Clickers Were Used: Introductory Statistical Methods

STT 200: Introductory Statistical Methods: Algebra-based introduction to statistics, is a service course for non-majors; it fulfills the University mathematics requirement. It is the "catch-all" course as our department also offers introductory courses for scientists, business majors, etc. The largest major represented is pre-nursing, but there also are a number of criminal justice, journalism, communications, psychology, and packaging majors. There are 120 students in the lecture. The sections I have taught in the last few years consist of one-third freshmen, one-third sophomores, and one-third upper classmen, with only a handful of seniors.
This is a 3-credit course. I meet the students for lectures 3 times a week for 50 minutes. In addition, they meet a teaching assistant (TA) in recitation section 1 day a week for 50 minutes. There are 4 recitation sections of about 30 students each. The clickers are used in lecture, but not in recitation. This course does not include the use of computer technology or computer labs. The students have graphing calculators instead of computers.

There are 3 midterms and a final exam. Homework is collected weekly and graded for correctness. There are 525 total points for the semester: homework 100 points (19%), each midterm 75 points (14%), final exam 150 points (28%), and clickers 50 points (10%).

[1] The bulk of this paper was written as a case study for the i>clicker website. The first author would like to thank the editorial staff at i>clicker for their help with the original manuscript and the management for their support of the work in the scholarship of teaching and learning. Any errors in this manuscript are the responsibility of the authors and not the editorial staff.

Motivation for Using Clickers

In 2006, the Guidelines for Assessment and Instruction in Statistics Education (GAISE) Project, funded and endorsed by the American Statistical Association (ASA), produced a report about the current status of and recommended directions for introductory statistics courses at the undergraduate level. Central to the recommendations for teaching introductory statistics made by the GAISE committee were the following: foster active learning in the classroom, use assessment to improve and evaluate student learning, and use real data (GAISE, 2006). My initial motivation for using clickers was to foster active learning and participation in my large lecture classes, two things I felt were lacking in my previous attempts at teaching in large lecture format.

At Michigan State University, as at most research institutions, the introductory statistics classes are taught in large lecture format with enrollments of 120 – 330 students per lecture. The 2005 – 2006 total enrollment in introductory statistics courses, that is, STT (statistics) courses that do not require an STT prerequisite, was over 4000 students, or more than 11% of the undergraduate student body. Given these large numbers, it is unlikely that the large lecture format will change. This format, however, makes it difficult to foster active learning and to use formative assessment efficiently to improve student learning.

Clickers, however, are becoming ubiquitous on campuses, particularly in large, lecture-based courses. These systems allow instructors to move away from didactic lecture formats towards more active learning strategies that encourage student participation and are consistent with research on active learning. Perhaps one of the best-known advocates for this approach is Eric Mazur, who uses clickers as part of his peer-instruction model in teaching conceptual physics (Mazur, 1996). At Michigan State University, we have found that clickers can help improve student outcomes in large-enrollment engineering (Sticklen & Urban-Lurain, 2006; Urban-Lurain, et al 2006), and biology (Zhang et al, 2005) classes.

In my classes, I have found that clickers help encourage student discussion and attention and provide formative assessment feedback to both the students and the instructor. In addition, the clickers allow me to collect, share, and use student data as part of my classroom presentations and activities.

Implementation

Clickers are a requirement for the course. Students were responsible for purchasing their clickers and registered them online at the i>clicker website. I synched the registrations with my roster using the i>grader software.

In the classroom I use a combination of PowerPoint presentations and writing on the document camera. My classroom has two projectors so I am able to show the students slides and written work simultaneously. I save my PowerPoint presentation on the USB drive that contains the i>clicker software and carry the drive and the base station to my classroom, where I plug it into the computer cart and run everything off the base station and USB drive. I generally have the clicker questions on PowerPoint slides so that the question is captured via screen shot by the i>clicker software. When I run a question off the document camera, I date the page and record the question number so I can match the question to the responses.

The students are encouraged to have a graphing calculator, but it is not required. An informal poll at the beginning of the semester indicated that nearly all of the students already owned one so I have been teaching the course under the assumption that all of the students have access to a graphing calculator and are bringing it with them to class.

Our University uses the ANGEL course management system. I use the course management website to upload my presentations and to communicate with the students and I use the online grade book. I do not upload each session's results to the online grade book or course management website because of the way I am calculating the clicker points, so students have to email me to know their daily clicker counts.

I use a medium stakes scheme for awarding clicker points. Students must answer 75% of the questions for the day and at least 1 must be correct for them to earn the clicker points for the day. Students may miss up to 3 days over the course of the semester without losing any clicker points. While I do not encourage this behavior, students who have forgotten their clicker have begun turning in their answers on paper. I generally assign a half-day of credit to those papers. Points are awarded based on the proportion of days on which the student participated in clicker questions.

I use clickers in every class. The distribution of the number of clicker questions per class is bimodal. On “low clicker” days I tended to ask 2 to 5 questions and on “high clicker” days I tended to ask 9 to 12 questions. There were roughly an equal number of “high clicker” and “low clicker” days.

I use 3 different types of clicker questions:
1) questions designed to highlight common conceptual misunderstandings in statistics,
2) questions designed as review questions for topics already covered, and
3) questions that are part of a class activity that will help students to learn a concept.

The type 1 and 2 questions were generated either by me or by modifying items found in the publicly available database of statistics assessment items on the website of the Assessment Resource Tools for Improving Statistical Thinking (ARTIST): https://app.gen.umn.edu/artist/index.html.

I generated the type 3 questions based on activities I had done previously in a similar course.

<u>Examples of questions designed to highlight common conceptual misunderstandings:</u>

1. In a 95% confidence interval:
   A. 95% of all samples will have proportions that fall within our calculated interval
   B. 95% of all samples will contain the true proportion in the calculated interval
   C. 95% of all samples will produce the same interval that we calculated
   D. 95% of all samples will produce an interval that overlaps the interval that we calculated

2. Students in a statistics class designed a survey about spending habits and gave it to a random sample of 300 students; 282 responded. Please read and evaluate the following statement. There are over 4000 students at the college. Therefore, the results of the survey may not be valid for drawing conclusions about how all students at the college spend money.
   A. Agree, 282 is too small a percentage of 4000 (7%) to allow us to draw conclusions about the population
   B. Agree, you should have a sample that is at least 50% of the population in order to make inferences
   C. Disagree, 282 is a large enough number to use for these purposes if the sample of students is random
   D. Disagree, if the sample is random, the size of the sample does not matter

Example of review questions

In a study to assess whether aspirin reduces the risk of a pregnant woman developing hypertension, 34 pregnant women were randomly assigned to receive a low dosage of aspirin every day while 31 pregnant women received a placebo every day. Of those in the aspirin group, 4 developed hypertension during their pregnancy, compared to 11 of those in the placebo group.

3. This study is

   A. A controlled experiment
   B. An observational study

4. Identify the explanatory variable in this study.

   A. Whether the woman had aspirin or a placebo
   B. Whether the woman developed hypertension during pregnancy
   C. Whether the woman was pregnant
   D. How much aspirin the woman took

5. Identify the response variable in this study.

   A. Whether the woman had aspirin or a placebo
   B. Whether the woman developed hypertension during pregnancy
   C. Whether the woman was pregnant
   D. How much aspirin the woman took

<u>Example of an activity using clickers:</u>

Initial clicker question:

On September 11, 2002, the New York State lottery winning number was 911. How often do you think the lottery number matches the date?

A More than once a year

B. About once a year

C. Once every 2 - 3 years

D Once every 4 - 5 years

E. Less than once every 5 years

Activity:

Even if we can't figure out how to calculate the probability that the lottery number will match the date (either because it is a difficult problem or because we haven't studied probability yet), we can still estimate the frequency with which it happens through SIMULATION.

- Identify the component to be repeated–choosing a 3-digit lottery number
- Explain how the outcome will be modeled–using the 3-digit numbers that could be lottery numbers
- Explain how you will simulate the trial–by having the calculator randomly choose a 3-digit number: randint (0,999)
- State clearly what the response variable is–whether the number generated by the calculator matches the date - 924
- Run several trials–each student will do the simulation 10 times and write down the number of matches to today's date

Clicker question to collect data:

How many times did you get 924?

a. Never b. Once c. Twice d. Three times e. Four Times

The results were that 6 students out of 100 matched the date one time and 1 student matched the date twice. So, out of 1000 trials, we had 8 matches.

Clicker question to wrap up the activity:

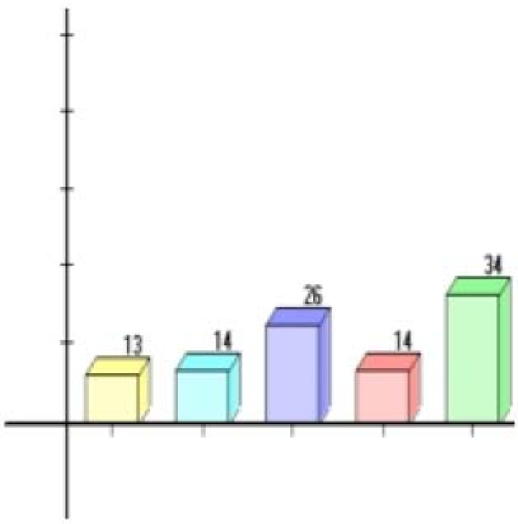
Based on the results of the simulation, how often do you think the lottery number matches the date?

A. More than once a year
B. About once a year
C. Once every 2 - 3 years
D. Once every 4 -5 years
E. Less than once every 5 years

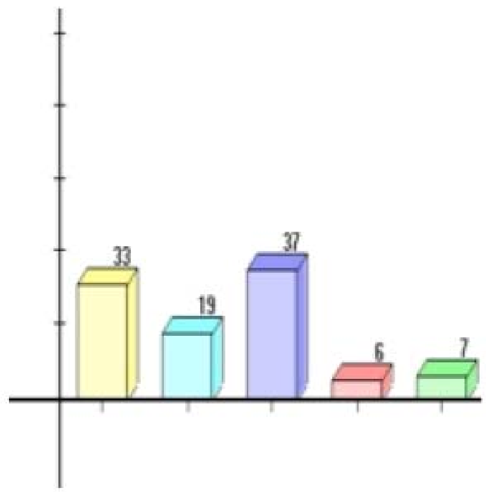
Two examples of student justifications for their choice in the wrap-up question were:

- 8/1000 * 365 = 2.92. so that would be 2 or 3 days a year, which is more than once a year
- Once we did the simulation I realized that we *should* match about one out of every 1000, so that would be about once every three years

How often do you think the lottery number matches the date? Bars for more often are on the left.



Responses before Simulation



Responses after Simulation

In general I show the clicker question, open the voting, and observe how fast the responses are coming in. If the responses are slow to come in, I encourage the students to talk to each other to help in choosing an answer. I usually show the class the distribution of answers. When most students have given a correct response, I generally say something like “good job” and explain the correct answer briefly for the students who were incorrect before moving on. When the responses are more uniformly distributed across the answer choices, I solicit comments from the students about why they chose a particular response and have them explain their reasoning. Sometimes, if I can tell from the base station that many students are responding incorrectly, I will give them hints about the problem to think about so they can change their response before the polling closes.

Simulation Activities from the Session

Slides for enacting these activities in class are available as part of the presentation handout.

The culminating idea of most currently taught introductory statistics courses is inference: drawing conclusions about a population from sample data. Hypothesis testing is one aspect of statistical inference. Students in fields such as pharmacy, nursing, or psychology are encouraged to take statistics early in their academic career so they are prepared to read, interpret, and later produce results based on hypothesis testing (personal conversations with students). There is anecdotal evidence to suggest that students can spontaneously make decisions about a population that is consistent with sample data, hence exhibiting intuitive understanding of the reasoning that underlies a hypothesis test (personal communication, Roxy Peck, 2004). These same students, however, struggle to master the formal reasoning and enactment of a hypothesis test. In fact, Brewer claims that the area of inference is “the most misunderstood, confused, and abused of all … statistics topics” (Brewer, 1985). Noll claims “that an understanding of sampling concepts and processes is necessary for developing a robust understanding of statistical inference” (Noll, 2007). The first simulation activity of the session is designed to help students understand sampling variability, one of the key sampling concepts.

Scenario for Activity 1 – election p: In the 2008 New Hampshire primary, 39% of voters voted for Hillary Clinton and 36% voted for Barak Obama. The polls prior to the primary showed Obama having a large lead on Clinton, possibly more than 10%. Could these results have been due to sampling variability?

Students used their calculators to simulate a random sample of 50 “voters” from a “population” in which 39% of the responses correspond to a vote for Clinton and 36% of the responses correspond to a vote for Obama. Students note the percent of outcomes in their sample that correspond to voters who plan to vote for each candidate. The results, collected via clickers, are used to illustrate the concept of sampling variability. Follow-up questions are used to probe student understanding of sampling variability.

Activity 2 is designed to address understanding of the Central Limit Theorem. Most statistics instructors agree that students have great difficulty in understanding the Central Limit Theorem, which is the basis for classical statistical inference as taught in introduction to statistics courses. Some reasons for this difficulty expressed in the literature are the belief that the sampling distribution should look like the population, especially as the sample size becomes larger, and that the variability of the sampling distribution either remains the same or increases as the sample size increases (Chance, et al, 2004). Both of these beliefs contradict the conclusion of the Central Limit Theorem.

In preparation for this lecture activity, the students would have completed the “penny sampling” activity with their TA in section. In that activity, students sample 5 pennies from a large jar of pennies. They record the age of each penny, the average age of the 5 pennies and the average age of 10 pennies, when they combine their sample with that of another student. These data are collected by the TA for use in the lecture presentation.

In the lecture activity, students simulate sampling from a uniform distribution and a binomial distribution. For the uniform distribution, students simulate rolling 10 dice and find the average of the 10 rolls. For the binomial distribution the students simulate 20 shots taken by a 41% basketball shooter in 1 game and record the number of baskets made. The data from both simulations are collected via the clickers. The data collection is followed by a clicker question about the description of the distribution of the collected data. The Central Limit Theorem is motivated and supported by the following thought question: How did the instructor know 1) what values would occur in the simulations and 2) the correct description of the distribution that would result when preparing the slides the previous day?

The third simulation activity is designed to highlight the correct meaning of a p-value and the basis for making a conclusion when conducting a hypothesis test. In a review of the literature, Lane-Getaz (2005) found 13 types of misconceptions of the meaning of the p-value calculated in classical hypothesis testing. Moreover, she found that these misconceptions are persistent and that many of them are held by doctoral students taking a second graduate level statistics course.

Scenario for Activity 3 - cell phones: A proud legislator claims that your state's new law against talking on a cell phone while driving has reduced cell phone use to less than 12% of all drivers. While waiting for the bus the next morning, you notice that 4 of the 10 people who drive by are using their cell phones. Does this cast doubt on the legislator's figure of 12%? Use a simulation to estimate the likelihood of seeing at least 4 out of 10 randomly selected drivers talking on their cell phones if the actual rate of usage is 12%. (DeVeaux, et al 2006).

This simulation is enacted in class and then used as the basis for introducing classical hypothesis testing. Follow-up clicker questions are used as formative assessment to find out what types of misconceptions of p-values student hold and to help them develop correct conceptions.

Results and Conclusions

I believe that the use of clickers has addressed the issues I previously had in teaching large lecture classes. At this time, the only data I have to support those feelings show a rise in scores on student evaluations. In addition, the mean attendance (as measured in the number of clickers recorded during class) over the course of the semester was 97 students out of 118. The standard deviation was 5.7 students. The minimum attendance was 82 students and the maximum was 106 students.

The mid-semester evaluation form included the statement: "Clickers are a valuable learning tool."; 39 students strongly agreed, 47 students agreed, 17 students were neutral, 6 students disagreed, and 2 students strongly disagreed. Among the free responses there were 36 positive comments about clickers and 12 negative comments. The most common positive comments were: it helps with understanding (11), it increases participation and interactivity (11), the clicker questions are a good self check of progress (8), and the clicker questions are good practice for tests (4). The most common negative responses were that the clicker questions take too much time (3) and that they aren't enough like the homework and test questions (3). One student requested that the correct answers to the clicker questions be available online after class.

Final Conclusion/Discussion

I would not consider teaching a large lecture class in the future without clickers. I would like to explore the possibility of using clickers that allow students to enter numeric responses, because it would make the types of activities and data collection that I could do more general. On the other hand, I am reluctant to give up the ease of use of the software that comes with the clicker.

One issue I would address with faculty new to clickers is the timing factor. I find that each clicker question takes 1 to 2 minutes for students to complete and then 1 or 2 minutes for discussion. This takes longer than showing a slide and reading information from it. While a clicker question takes more time, it requires the students to engage with the material and presentation in a way that copying a slide or listening to the instructor talk does not. On the other hand, the change in timing does require attention on the part of the instructor. One of the things I did to compensate for the timing was to have my TA do many of the problems I had previously used as class examples in the recitation sections In addition, I had him do some of the activities that were better suited to smaller groups or from which I did not have a good method for collecting responses via the clickers in lecture.

One issue that surfaced in my first semester of using clickers was the number of students who did not engage with the question or activity and, instead, chose a random answer. These students are evident because they will choose, for example, option E when there are only 3 or 4 possible responses. In most cases, these students do not affect the overall learning in the class, but there was one day in particular that an activity did not have the intended outcome and it affected student learning. I had asked the students to select a random sample using a calculator simulation and then create a 95% confidence interval from their sample. The clicker question was “Did your interval cover the true populations proportion?” with responses “Yes” or “No.” It is expected that 95% of the students would respond “Yes,” but only about 80% responded “Yes.” After class one of the students told me that many students had not actually participated in the activity and then had randomly chosen an answer to the covering question. In the future I might be able to solve the problem by adding a third response option, “I haven’t completed a confidence interval yet.” This would provide an opportunity for students to feel they have earned the clicker point without a judgment about why they did not have a response.

References

Brewer, J.K. (1985). Behavioral statistics textbooks: Sources of myths and misconceptions? *Journal of Educational Statistics*, 10 (3), pp. 252-268.

Chance, B., delMas R., & Garfield, J. (2004). Reasoning about sampling distributions. In Dani Ben-Zvi & Joan Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 295-323). Dordrecht, The Netherlands: Kluwer.

Caldwell, J. E. (2007). Clickers in the large classroom: Current research and best practice tips. *CBE Life Sciences Education*, (6), pp. 9-20.

delMas, R., Garfield, J., & Chance, B. (1999). A model of classroom research in action: Developing simulation activities to improve students' statistical reasoning. *Journal of Statistics Education*, 7 (3), Online: www.amstat.org/publications/jse/v7n3/delmas.cfm

DeVeaux, R., Velleman, P., & Bock, D. (2006). *Intro Stats* (second edition). Boston, MA: Pearson Education, Inc.

Lane-Getaz, S. J. (2005). Summary of p-value research. Presentation at the U.S. Conference on Teaching Statistics. Available online: www.causeweb.org/uscots/spotlight/files/R6.pdf

Lane-Getaz, S. J., & Zieffler, A. S. (2006). Using simulation to introduce inference: An active learning approach. Proceedings of the American Statistical Association, Statistical Education Section, Alexandria, VA: American Statistical Association.

Mazur, E. (1996). *Peer instruction: A users' manual* (first edition.). Upper Saddle River, NJ: Prentice Hall.

Noll, J.A. (2007). Graduate teaching assistants' statistical knowledge for teaching. *Unpublished Doctoral Dissertation.* Downloaded March 17, 2008 from: http://www.stat.auckland.ac.nz/~iase/publications/dissertations/dissertations.php

Pfannkuch, M. (2005). Probability and Statistical Inference: How Can Teachers Enable Learners to Make the Connection? In *Exploring Probability in School*: *Challenges for Teaching and Learning* , Jones, E (ed) (pp. 267-294). Retrieved July 30, 2008, from http://dx.doi.org/10.1007/0-387-24530-8_12.

Saldanha, L., & Thompson, P. (2003). Conceptions of sample and their relationship to statistical inference. *Educational Studies in Mathematics*, 51, pp. 257 - 270.

Sticklen, J., & Urban-Lurain, M. (2006). *Enhancing learning of low ability students in multi-section freshman lecture/laboratory classes.* Paper presented at the American Society of Electrical Engineers Annual Meeting, Chicago, IL.

Urban-Lurain, M., Sticklen, J., & Buch, N. (2006). *High enrollment, early engineering courses and the personal response system.* Paper presented at the American Society of Electrical Engineers Annual Meeting, Chicago, IL.

Zhang, B., Patterson, R., Richmond, G., Parker, J., Merrill, J., Urban-Lurain, M., et al. (2005). *Using self-response system and online learning environment in large college science classes-the technologies, instructional design, and implications.* Paper presented at the International Conference on Computers in Education, Singapore.