

Teaching and Assessing Higher Order Thinking in the Mathematics Classroom with Clickers

Jim Rubin & Manikya Rajakaruna
Union College, USA

•Received 17 September 2014•Revised 13 Marc 2015•Accepted 11 April 2015

Many schools have invested in clicker technology, due to the capacity of the software to track formative assessment and the increased motivation that students show for incorporating technology in the classroom. As with any adoption of new software that demands amending pedagogy and learning applications, the extent to which clickers are living up to expectations has not yet become apparent. The present study sought to explore the potential of using clickers to teach the reasoning processes behind solving higher order thinking word problems in a mathematics class. A pilot study was conducted with a college algebra class to refine questions used in the coursework and field test a survey to measure student attitudes towards the teaching methodology. The main study took place over the fall semester with a college algebra class (N=21). Results showed increased student motivation and acumen for using the technology and higher test scores, but frustration on the part of both the teacher and students when trying to apply the pedagogy for the purpose of learning higher order thinking reasoning processes. The potential for the technology to offer an alternative for formative assessment was a strong positive element.

Keywords: clickers, college algebra, higher order thinking, mathematics

INTRODUCTION

Recent trends have re-targeted mathematics as a subject that is need of an overhaul in terms of the emphasis in the curriculum, the degree to which it should be assessed, and the focus on developing higher order thinking skills. The recent change in educational standards that have been adopted by a vast majority of the states has put renewed demands on teachers to place an emphasis on teaching higher order thinking skills (Common Core State Standards Initiative, 2015) amidst an ongoing need to track student outcomes and a technological revolution that offers the potential to support both of those seemingly disparate goals. The Federal common core standards reflect a perceived need to educate students with the ability to analyze and problem solve, attributes which align strongly with the skills most sought after by many 21st century employers (National Education Association,

Correspondence: Jim Rubin,
Department of Education, Union College, 310 College St., Barbourville,
KY 40906, USA.
E-mail: jrubin@unionky.edu
doi: 10.12973/mathedu.2015.103a

2015). Identifying, teaching, and assessing higher order thinking skills has been a challenge for teacher pedagogy for many years, as well as for the design of an assessment system that can track progress in a systematic way for each student, individual classes, and school systems. Using clickers with well-planned questions that target higher order thinking reasoning processes seemed to offer a viable solution.

With these concerns in mind, the purpose of the current study was to explore the use of clicker technology in a college algebra class, with the expressed goal of utilizing formative instruction to enhance the support of learning higher-order problem solving skills in mathematics. . While clickers have been shown to enhance student participation, there have been no previous studies that sought to single out the potential for targeting the teaching of solving math word problems. The criteria for characterizing higher order thinking problems was based on the need to apply multiple steps for solving a problem for which no known formula or pattern could be directly applied (Resnick, 1987). Related research questions included: would there be an increase in student satisfaction concerning pedagogy with the use of clicker technology; would students be more motivated to participate in class using clicker technology; would students feel more confident about solving math problems after being taught with clicker technology; how would the use of clickers to support higher order thinking fit into the demands of a preset curriculum; and what issues would be important for teachers when adopting clicker technology in the classroom?

LITERATURE REVIEW

What appears as a shift in priorities for educators as a result of new standards has actually been a facet of the literature for over 20 years. In the mid-seventies Eisenberg (1975) recognized the need to broaden the goals of mathematics education to include higher order thinking skills, problem solving, formulating hypotheses, and investigation, with a focus on cognitive mechanisms. Romberg, Zarinnia, and Collis (1990) studied the relationship that assessment should have on directing curriculum and instruction and realized the importance of developing evaluation procedures that are sensitive to the thinking processes that produced answers as well as the answer itself.

The higher order thinking processes that occur in the process of solving mathematical problems are characterized by the application of multiple criteria, which may not be known in advance. Problem solving thought patterns involve multiple steps in a complex analysis that requires elaboration and judgment (Resnick, 1987). Thus, there is a need for an assessment tool to document the intricacies of the thinking processes, rather than documenting only the final solution. The SOLO taxonomy was offered by Biggs and Collis (1982) to address these concerns and illustrates a methodology for conceptualizing a model of integrating problem solving processes. The closed format version of this approach broke down the cycle of learning into four stages: unistructural – using one obvious piece of information; multistructural – forming closure with the use of separate pieces of information in the stem; relational – using two or more closures that relate to an integrated understanding of the stem; and extended abstract – deriving a general principle or hypothesis that is derived or suggested by the information. Concepts derived from the taxonomy supported an understanding of the rationale used to think through the problem solving process, and provided a foundation for the present study to devise support clicker questions that were used at various stages in the solution process.

Criticisms for using multiple choice questions to measure higher order thinking have been modified over the past several decades. As the need to teach higher order

thinking skills related to mathematical problem solving took root in the 1980's, severe criticism of objective test questions as a means of assessment was prevalent. Berlak (1985) and Romberg, Zarinnia, and Collis (1990) noted a limitation of the multiple-choice format because of the inherent nature to assess specific items of content that call for one right answer. Further criticism related to the lack of creative demands during the assessment process by virtue of requiring one choice among given alternatives. Most notable to the concerns of this research was the perceived inability of multiple choice questions to evaluate higher order thinking processes, contending that because concepts are tested independently of one another, there was no way to assess a student's cognitive thought approach to a given problem (Romberg, Zarinnia, & Collis, 1990).

Over the past twenty-five years the potential of constructing multiple-choice questions to evaluate higher order thinking has evolved (Bender, 1980; Douglas, Wilson, & Ennis, 2012; Dowd, 1992; Hansen & Dexter, 1997; Lin & Singh, 2012; Lockwood, 2003; Ray, 1978; Stupans, 2006; Torres, Lopes, Babo, & Azevedo, 2011; Wayne, 1982). While Frederiksen (1984) made a strong case for including other question formats when evaluating higher learning levels, both Sternberg (1985) and Ennis (1985) used the format when asked to develop measures to test critical and creative thinking, sensing the potential for designing questions that challenge original thought processes. Current scholarship from the Teaching Effectiveness Program (2014) provided guidance for tailoring questions for the purposes of the current study. Of particular import was the application of creating items that utilized novel language and unique examples, requiring students to apply analysis in a meaningful way (see Appendix A for examples).

The decision to use clickers in this research fulfilled several goals. First and foremost, there is a growing body of research that supports the use of clickers in the mathematics classroom to support tracking formative assessment and the enhanced potential to support learning outcomes (DeBourgh, 2008; Kolikant, Calkins, & Drane, 2010; Liu & Stengel, 2011; Popelka, 2010). Problem solving skill sets in mathematics involve conceptualizing the rationale that organizes given information into a solution, and clicker technology provided a window into students' understanding of the specific steps of that process. Research showed clickers provided instant feedback that allowed for immediate response to student needs (Ribbens, 2007), as well as a medium for motivating whole class participation (Caldwell, 2007; Miller, Ashar, & Getz, 2003). Clickers also provided a format that provided detail record keeping throughout the formative assessment process (Oerman & Gaberson, 2006). Several suggestions from a recent study by Liu and Stengel (2011) guided use of clickers in the present study. (1) Targeted questions were included as part of student's grades in an effort to keep students focused and motivated to participate. (2) Clicker questions were used to break down the steps of complex word problems in an effort to teach the thinking processes that were an integral part of problem solving. (3) The teacher used verbal encouragement to encourage students to use reasoning to answer questions rather than random guessing by reminding them that similar questions would be part of a graded exam. (4) Based on instant feedback from clicker responses, the teacher utilized data to probe for the reasoning behind questions choices. (5) Response time was monitored by the teacher to represent a compromise between the fastest and slowest students, to encourage some students to work faster and not bore students who were already finished.

Research also provided insight into issues that presented problems for practitioners who used clickers. Time spent learning to use the software (Hatch, Murray, & Moore, 2005) and the effort incurred in setting up, using, and taking down the software was mentioned (Stuart, Brown, & Draper, 2004). Prior studies also showed a problem for teachers who tried to incorporate revised pedagogy into

a curriculum that required a defined a body of knowledge, where the outcomes of learning took precedence over the process (Elias & Merriman, 2005; Miller et al., 2003). The type and format of the questions designed for clicker use was found to be a problem due to the demands of class time to move on to subsequent topics (Stuart, Brown, and Draper, 2004), and Uhari, Renko, and Soini (2003) noted the need to include slides in math problems that showed how to arrive at solutions, rather than simply polling correct/incorrect answers.

The current study sought to explore the suggestions from prior research by breaking down the reasoning processes needed to solve word problems through the application of classroom methodology that included use of clickers. Additionally, the present study has interest in documenting student attitudes toward using clickers as well as data that tracks effectiveness in supporting students to improve reasoning skills associated with higher order thinking questions.

RESEARCH DESIGN

A pilot study was conducted during the summer term to allow for: refinement of the survey that was used to document changes in student sentiment toward classroom pedagogy that included using clickers to scrutinize higher order thinking math problems; development of clicker questions that were important for teaching and assessing the use of higher order thinking skills for solving word problems; and training for the teacher to use and implement the technology. Teaching was conducted in the traditional manner (classroom lecture and class participation without using clickers) for the first two weeks of the summer session, and clicker technology was implemented for the last two weeks that focused more intently on breaking down the reasoning processes behind solving mathematical word problems. Students completed a survey at the mid-term and at the end of the course to measure sentiment related to how the class was taught. Survey results from students, quiz and test scores, clicker questions data, and teacher notes were used for analyzing outcomes.

The main study was conducted the following fall semester utilizing a similar format. Students were taught using the traditional method for the first half of the semester and with clickers for the second half of the semester. Pedagogy using clickers was characterized by a more analytical approach to solving word problems. A PowerPoint that included clicker questions was used to break down the reasoning process of individual problems and immediate feedback was available to the teacher concerning students' ability to grasp concepts. Clicker questions were categorized and tracked into two categories, factual and higher order thinking. Fact based questions were those that could be solved using a pre-learned algorithm, while higher order thinking questions involved a reasoning process that was not based on a pre-learned formula (see Appendix A for details). The re-worded survey was issued at the mid-point of the semester and again at the end of the course to determine any changes in student attitudes towards how the course was taught.

Participants

The sample of the study consisted of students in a college algebra class in a small private college in the southeastern United States. The teacher (and co-author of the current study) had 7 years prior teaching experience, a superlative record documenting teaching ability, and no prior experience using clicker technology. The summer course consisted of 6 students and the fall course consisted of 21 students, all who self-enrolled in the courses.

Data collection instruments

The pilot study was used to refine a survey that was administered at the mid-term and at the end of the term. Each question in the survey related to one of 5 dependent variables that were of interest to the present study: Organization of Course Content (questions 1, 10, and 15); Satisfaction of Pedagogy (questions 2, 3, 5, 9 and 18); Motivation to Learn (questions 4, 7, 8, and 19); Support of Higher Order Thinking (questions 6, 11, 12, 13, and 14); and Self-Efficacy (questions 16 and 17). Comprehensive exam scores were recorded at the mid-term and end of the course and student responses to clicker questions were also used to evaluate the research questions.

Procedures

Based on responses from the pilot study, the questions on the survey were refined to better reflect the attitudes of students as they pertained to each of the 5 variables of interest (mentioned above). The pilot study was also used to reevaluate the questions that would be used as examples of higher order thinking word problems in the main study and the details of what reasoning steps would be required to solve them.

The main study took place in the first half of the subsequent fall semester. Classes were taught by giving examples of problems, showing solutions, and having students attempt to solve similar problems on their own, while the teacher circulated around the classroom offering support when needed. At mid-term, students took a comprehensive exam based on the standard content of a college algebra course and a survey designed to measure their sentiment about how the course was being taught (see section on Data Collection Instruments). After the mid-term classes were taught with clickers. Students responded to clicker questions throughout the classes that were categorized as being factual (those that could be solved using a pre-learned algorithm) and higher order thinking (those that needed original reasoning processes to solve). Examples of solving higher order thinking questions were presented and broken down into the specific reasoning processes that were necessary. Subsequent clicker questions presented other examples of higher order thinking questions that tracked whether students could grasp each step of the solution. Clicker questions were presented in a PowerPoint format as short answer questions that required a correct answer. Clicker software tabulated results broken down by factual and higher order thinking types of questions at the end of each class. At the end of the semester a comprehensive exam was given to test overall competency and students re-took the survey that was originally administered at the mid-term.

DATA ANALYSIS

Pilot study

Mid-term and final exams were given during the pilot study. Mid-term exam scores reflected outcomes after teaching that used the traditional methodology as mentioned above; final exam scores reflected outcomes after teaching that used clicker technology. Scores on the final improved on average 7.5 points higher compared with the mid-term scores, indicating a positive trend, although not statistically significant (See Table 1).

Table 1. Mid-term and final exam score comparison_ - paired samples test

	N	corr.	M	df	p
Exam1-Exam2	6	.357	-7.5	5	.057

Figure 1 illustrates the percent of correct scores when using clickers to test students' knowledge in the classroom. The pilot study was the first time any of these students used clicker questions as a classroom learning tool. The positive trend indicates that their ability to focus and participate in solving problems within the context of classroom learning improved with continued exposure to the technology.

Figure 2 gives a breakdown of how students performed answering factual vs. higher order thinking questions. The positive trend that is illustrated for both types of questions indicates that the skill sets for solving both factual and higher order thinking questions improved with ongoing practice.

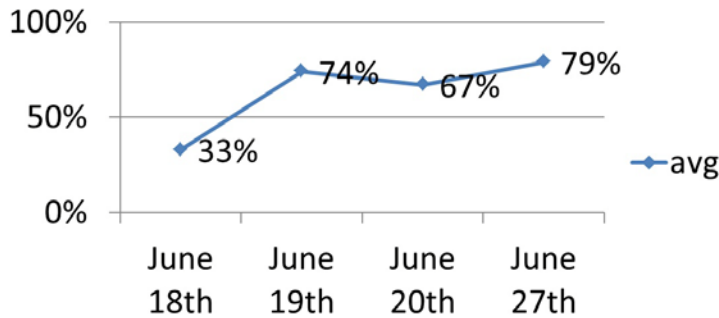


Figure 1. Average class scores using clickers

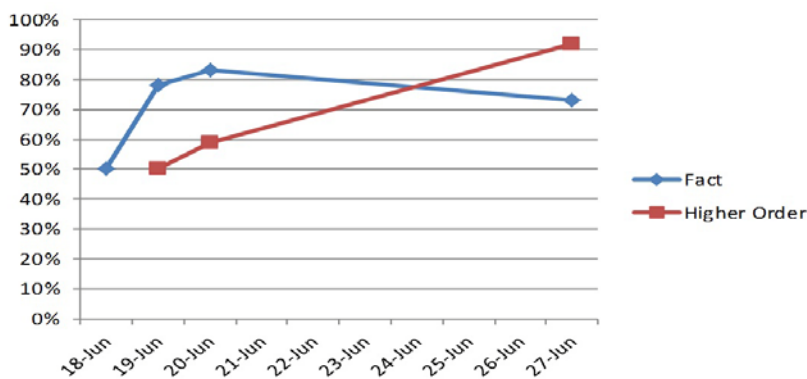


Figure 2. Comparison of clicker use with factual and higher order thinking questions

The following three points that were implemented in the course procedures for the main phase of the study stemmed from student feedback and teacher experience with clickers in the pilot phase of the study. (1) In an attempt to motivate students to participate in solving clicker problems, 1 extra credit point was given for participation and 10% of extra credit points were added to the final grade. (2) To further motivate focus during the use of clicker problems, students were notified that word problems used in the clicker exercises would be similar to those that would be included on quizzes and exams. (3) Due to student enthusiasm for working in small groups during the pilot study, collaboration between groups of 2 or 3 students was encouraged during the main study.

Main Study

A comprehensive exam was given at the end of the first 8 weeks of class, representing the time period when students were taught in the traditional method of teacher lecture. A second comprehensive exam was given at the end of the second 8 week period of the semester, representing the time period when students were using clicker technology as a pedagogical tool for learning the material. Figure 3 details the scores on exam 1 to check for outliers that may affect reliability in

comparing scores. With a standard deviation of 11.85, all scores fall within 3 standard deviations of the mean (81.28).

Figures 3 and 4 details the scores of all students who took the mid-term and final exams. Figure 4 details the scores of exam 2 to check for outliers that may affect reliability in comparing scores. With a standard deviation of 12.92, all scores fall within 3 standard deviations of the mean (87.11).

Table 2 details the comparison of scores from exam 1 (prior to using clickers) and exam 2 (after using clickers). On average, student scores on the second exam were 5.8 points higher than on the first exam, which is a positive trend, although not statistically significant.

Exam 1 and Exam 2 each had higher order thinking questions that students did not attempt to answer. Table 3 indicates that a higher percentage of students attempted to answer these questions during Exam 2, indicating a stronger sense of

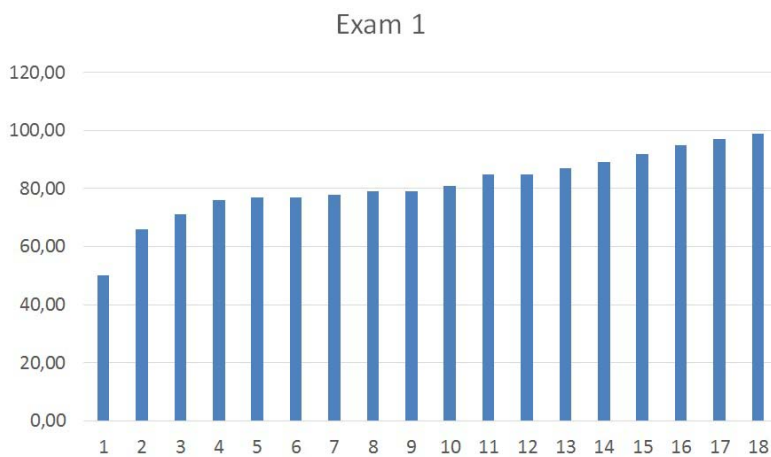


Figure 3. Student scores on Exam 1

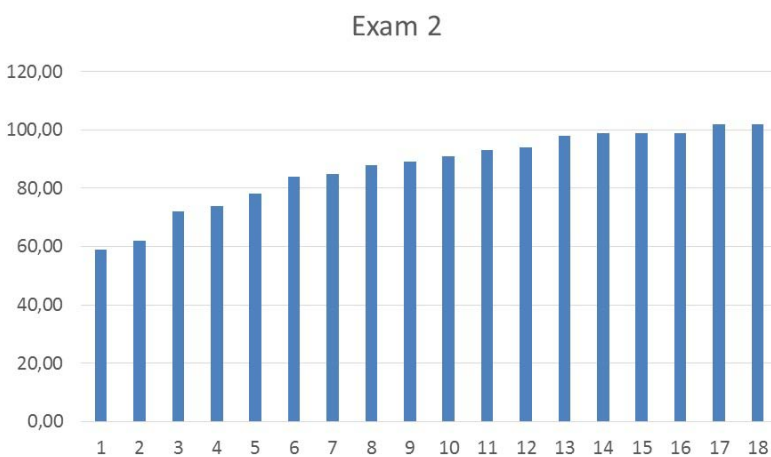


Figure 4. Student scores on Exam 2

Table 2. Paired samples test

	N	T	df	MD	p
Exam1-Exam2	18	-1.861	17	-5.83333	.080

Table 3. Percentage of students who answered word problems

Exam	Word Questions Left Unanswered	Word Questions Attempted
Exam 1 (prior to using clickers)	6 students (28.57%)	15 students (71.43%)
Exam 2 (after using clickers)	3 students (15.79%)	16 students (84.21%)

self-confidence in solving word problems after working with the clicker pedagogy.

Students responded to a survey that was given both at the mid-term and at the end of the course that measured sentiments about how the course was taught. The average correlation between scores on the two surveys was .361.

A chi square test was run comparing the responses in each of the 5 categories analyzed in the study. Table 4 details the results. Of the 19 questions in the survey, there were 3 questions that related to Organization, 5 questions that related to Satisfaction with Pedagogy, 4 questions related to Motivation, 5 questions related to Support for solving Higher Order Thinking Problems, and 2 questions related to Self-Efficacy. Table 4 illustrates the results, which shows there was a positive trend for students favoring the teaching pedagogy that included use of clickers and the process of breaking down the reasoning process for solving higher order thinking problems into logical steps, however, none of the individual categories showed a significant difference.

Figure 6 aggregates the two questions of the Likert survey that relate to negative sentiment, disregards the neutral responses, and aggregates the two questions that relate to positive sentiment. The results visually clarify the increased positive sentiment from the class towards the teaching methodology that included use of clickers.

Table 4. Chi Square Tests for each dependent variable

Category	Chi Square	DF	P
Organization	5.9294	4	0.2045
Pedagogy	2.9541	3	0.3988
Motivation	1.2969	1	0.2548
Higher Order Thinking	5.8180	4	0.2132
Efficacy	1.5284	2	0.4657

Enhancement of Perceived Quality of Teaching after Using Clickers ($\chi^2(df=4, N=628)=10.6653, p=.0306$)

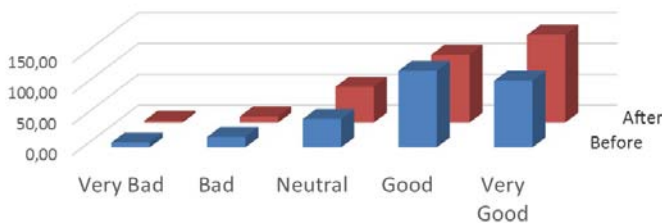


Figure 5. Student survey results after using clickers

Enhancement of Perceived Quality of Teaching after Using Clickers ($\chi^2(4, N=527)=4.49, p=.034$)

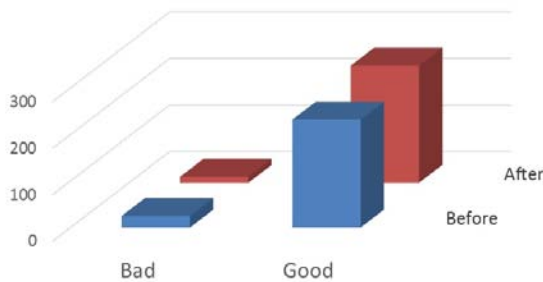


Figure 6. Aggregate of student survey after using clickers

In order to further test the positive trend shown in each individual category, a chi square test was run on the entire survey to compare overall sentiment changes between the two teaching methods (before mid-term and after mid-term). Figure 5 illustrates the results by tracking each question in the Likert survey, showing that there was a significant positive difference in the attitudes of students after being taught with the clicker methodology.

For each class session, results of how well students answered factual and higher order thinking questions were tracked. Results (see Figure 7) showed a slight positive trend for answering both types of questions as the class progressed, with students generally able to answer factual questions better than higher order thinking questions.

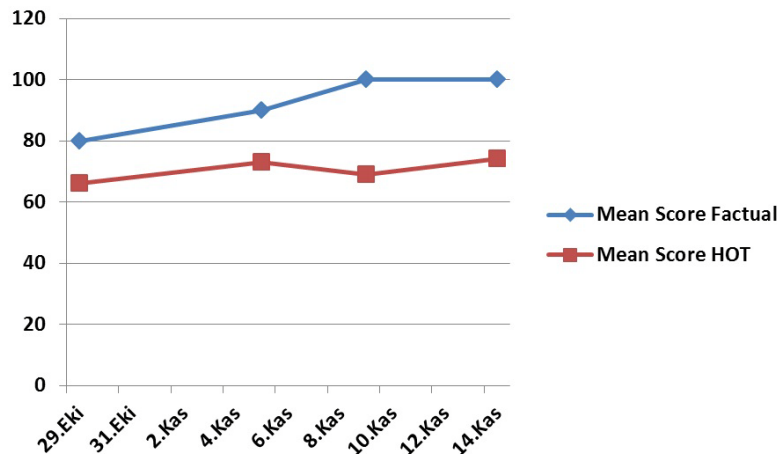


Figure 7. Correct responses to factual vs. HOT questions

DISCUSSION

This study represented a first time experience using clickers for the teacher and the students of both the pilot and main studies; thus, there was a learning curve for the teacher to use clickers as an instructional tool and for the students to use the device as a learning platform. The significant positive sentiment of students after being taught with the innovative pedagogy that broke down reasoning processes and tracked understanding with formative assessment that utilized clicker technology, demonstrates the potential benefit that this teaching methodology has for supporting students to solve higher order thinking problems and encouraging engagement.

Previous research (DeBourgh, 2008) noted the amount of time and effort required to design effective questions, and some of the frustration on the part of the teacher was due to this extra effort. As Stuart, Brown, and Draper (2004), Miller, Ashar, and Getz (2003), and Elias & Merimam (2005) noted, there may be problems due to implementing a change in pedagogy that emphasizes the process of learning into a busy semester with high content demands that require clear performance indicators. The time spent breaking down the reasoning processes for specific higher order thinking questions was perceived as ponderous and time consuming for both the teacher and the students. With a set amount of material that needed to be covered in the curriculum, there was not time to devote extra attention to working on solving higher order thinking problems in such detail. Due to the nature of higher order thinking problems, there wasn't a specific formula to teach that could prepare the students for solving them. Examples of reasoning processes were taught for related multi-stepped problems, and there was hope that by giving the students the experience of breaking down the reasoning process for one problem,

there would be an element of transmission in applying that knowledge for different problems. Indeed, when a subsequent problem was given that required a similar reasoning process, 81% of students answered the problem correctly. However, the time spent in this process was perceived to be out of step with the usual pace of the course material, and there was frustration on the part of the teacher that intense focus on one aspect of the curriculum would compromise coverage of other material. The teacher also expressed frustrations concerning a lack of focus among several students for completing clicker questions in a timely manner, causing some students to wait too long after completing their work.

Absences also may have been a contributing factor that contributed to student frustration during the clicker part of the course. Students that missed previous classes were reticent to participate in classroom activities that involved clickers, claiming that they were unprepared and would rather not expose their lack of knowledge. In most uses of clickers the individual students' response remained anonymous; however, when using the technology for the purpose of making the teacher aware of which students did not understand a specific element of mathematical reasoning, it was important to identify individual students so clarification of reasoning processes could take place for those that required it. The result, however, was exposing those who made errors. Motivation strategies to encourage participation included encouraging students and pointing out that similar questions would be on future exams, but these strategies were less successful for those students who missed prior class time. Upon initiating the research, there was a high degree of optimism on the teacher's part for using clickers, with every intention of making the technology a regular part of instructional repertoire in future classes. However, after completing the study, there was some reluctance to continue their use, as traditional teaching methods, learned through the years, proved to be more comfortable and not fraught with similar issues.

Despite these negative sentiments, efforts to focus on higher order thinking processes with clicker technology may have contributed to improved student outcomes, and there was significant positive changes in student sentiment towards the course after experiencing the refocused methodology that utilized clickers. There was also a nearly 6 point improvement in exam scores after using clickers (table 2). More time and experience for both the teacher and students in using this type of pedagogy on an ongoing basis might abate the sense of frustration, while providing a foundation for approaching solutions to more complex problems.

The increased positive sentiment from students after using clickers resonates with previous research on the topic. Miller, Ashar, and Getz (2003) and Uhari, Renko, and Soini (2003) found an increase in student interaction with course material using this methodology. More engagement may leave students with a sense of ownership in course format, translating into positive sentiment towards organization and pedagogy. The perception of better organization may also be related by the logical sequence of presenting step by step reasoning processes for solving HOT questions. While the process proved to have a degree of frustration for both teachers and students in terms of execution, students may have sensed a higher degree of order to the sequence of course material presentation. With a small sample ($N=21$), results of a chi square did not show significance when testing for each independent variable in the survey, although there was a clear positive trend. However, by aggregating all responses, the chi square test showed a positive change in sentiment that was significant, indicating that with larger sample sizes, each of the dependent variables may also show significant differences. These results imply that implementation of clicker use in this framework may well be worth the effort.

In order for clickers to emerge as a useful tool that will affect how learning takes place in the classroom, it would be important to address the extent of professional development that is provided and the expectations of the curriculum. One of the

advantages in using clickers, noted by Oerman & Gaberson (2006), was the capacity for clickers to provide a platform for formative instruction, enabling the teacher to respond immediately where further learning needs to occur. In theory the software does provide this capability, but results from the current study indicate that professional development and extended practice with the software may be necessary to fully develop this skill set. Uhari, Renko, and Soini (2003) reported needing a half a day of training, along with peer support, for a typical user. The teacher in this study trained for several days in terms of learning how to manage the technical aspects of the software and used clickers for the first time during the pilot phase of the study. The ability to manipulate the software is, however, only one aspect of the learning curve. Applying the skills necessary to react to formative assessment in real time and make immediate decisions that require amendments to previously scripted lessons requires considerably longer training sessions and redefines expectations of classroom management acumen that should be an outcome of professional development training.

CONCLUSIONS AND FUTURE STUDY

For teachers to take the time to delve into the reasoning processes that are required for solving higher order thinking problems, an adjustment to end of course expectations in terms of expected content that must be covered should be made. This suggestion is aligned with the sentiment of the common core standards, exemplified in the emphasis to focus on teaching analysis and problem solving skills over rote memorization. Measuring how well students can answer questions for which no explicit algorithms can be learned is not as straightforward as teaching a set pattern for solving problems and testing recall of that knowledge, and the process for teaching such skill sets may not be able to be compartmentalized into rigid time slots. However, unless the teacher has license from administrators to take the time to explore this aspect of learning, it will be a frustrating digression from the rigors of an already challenging job.

Future studies could benefit by using a larger sample size and organizing a curriculum that has greater built in flexibility to spend more time on the challenges of solving complex problems that require original thought processes. Conducting the research over a longer time frame would also help verify results.

AUTHORS' NOTE

The authors would like to thank Dr. Illie Puiu Vasilescu, psychology professor at Union College, for his contributions to the study.

REFERENCES

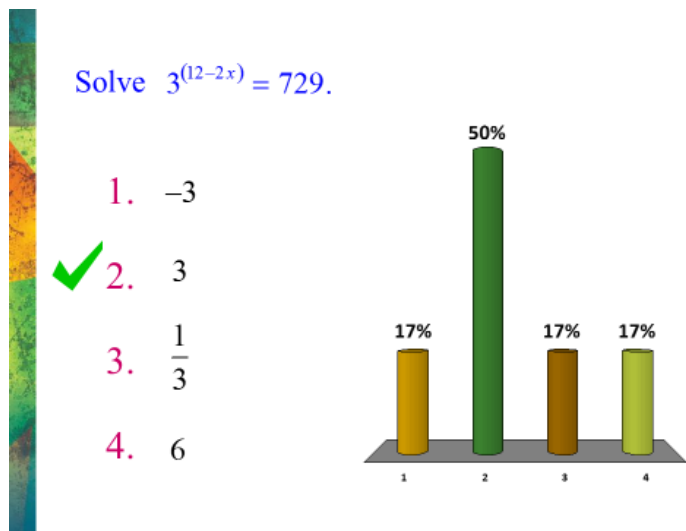
- Bender, T.A. (1980). Processing multiple choice and recall test questions. Paper presented at the *Annual Meeting of the American Educational Research Association*. Boston, MA. Retrieved from <http://www.eric.ed.gov/contentdelivery/servlet/ERICServlet?accno=ED189160>
- Berlak, H. (1985). Testing in a democracy. *Educational Leadership*, 43(2), 16-17.
- Biggs, J.B. & Collis, K. F. (1982). *Evaluating the quality of learning: the solo taxonomy*. New York: Academic Press.
- Caldwell, J. E. (2007). Clickers in the large classroom: Current research and best practice tips. *CBE Life Sciences Education*, 6(1), 9-20.
- Collis, K. F. (1982). The solo taxonomy as a basis of assessing levels of reasoning in mathematical problem solving. Proceedings from the *Sixth International Conference for the Psychology of Mathematical Education*. Antwerp, Belgium: University of Antwerp.

- Collis, K. G., Romberg, T.A., & Jurdak, M. E. (1986). A technique for assessing mathematical problem-solving ability. *Journal for Research in Mathematics Education*, 17(3), 206-221.
- Common Core State Standards Initiative (2015). *About the common core state standards*. Retrieved from <http://www.corestandards.org/about-the-standards/>
- DeBourgh, G. A. (2008). Use of classroom "clickers" to promote acquisition of advanced reasoning skills. *Nurse Education in Practice*, 8, 76-87.
- Douglas, M., Wilson, J., & Ennis, S. (2012). Multiple-choice question tests: A convenient, flexible and effective learning tool? A case study. *Innovations In Education And Teaching International*, 49(2), 111-121.
- Dowd, S. B. (1992). *Multiple-choice and alternate-choice questions: Description and analysis*. Retrieved from <http://files.eric.ed.gov/fulltext/ED351376.pdf>
- Elias, J. L., & Merriam, S. B. (2005). *Philosophical foundations of adult education* (3rd ed.). Malabar, FL: Krieger Publishing Company.
- Ennis, R. (1985). Large scale assessment of critical thinking in the fourth grade. Paper presented at *Issues in the Development of a Large-Scale Assessment of Critical Thinking Skills*. The American Educational Research Association Annual Meeting. Chicago, Illinois.
- Frederiksen, N. (1984). The real test bias, *American Psychologist*, 39(1), 1-10.
- Hansen, J. D., & Dexter, L. (1997). Quality multiple-choice test questions: Item-writing. *Journal of Education for Business*, 73(2), 94.
- Hatch, J., Murray, J., & Moore, R. (2005). Manna from heaven or "clickers" from hell: Experiences with an electronic response system. *Journal of College Science Teaching*, 34(7), 36-39.
- Kolikant, Y.B.D., Calkins, S., & Drane, D. (2010). "Clickers" as catalysts for transformation of teachers. *College Teaching*, 58, 127-135.
- Lin, S., & Singh, C. (2012). Can multiple-choice questions simulate free-response questions? *AIP Conference Proceedings*, 1413(1), 47-50. doi:10.1063/1.3679990
- Lockwood, D.F. (2003). *Higher order thinking in teaching senior science*. Retrieved from <http://members.shaw.ca/donlockwood/mcquestions.htm>
- Liu, W.C. & Stengel, D. (2011). Improving student retention and performance in quantitative courses using clickers. *The International Journal for Technology in Mathematics Education*, 18(1), 51-58.
- Miller, R. G., Ashar, B. H., & Getz, K. J. (2003). Evaluation of an audience response system for the continuing education of health professionals. *The Journal of Continuing Education in the Health Professions*, 23(2), 109-115.
- National Education Association (2015). *An educator's guide to the "four Cs"*. Retrieved from <http://www.nea.org/assets/docs/A-Guide-to-Four-Cs.pdf>
- Oermann, M. H., & Gaberson, K. B. (2006). *Evaluation and testing in nursing education* (2nd ed.). New York: Springer Publishing Company, Inc.
- Popelka, S. R. (2010). Now we're really clicking! *Mathematics Teacher*, 104(4), 290-295.
- Ray, W. (1978). Writing multiple-choice questions: The problem and a proposed solution. *The History Teacher*, 11(2), 211-218.
- Resnick, L.B. (1987). *Education and learning to think*. Washington, DC: National Academy Press.
- Ribbens, E. (2007). Why I like personal response systems. *Journal of College Science Teaching*, 37(2), 60-62.
- Romberg, T.A, Zarinnia, E.A., Collis, K.F. (1990). A new world view of assessment in mathematics. In G. Kulm (Ed.), *Assessing Higher Order Thinking in Mathematics* (pp. 21-38). Washington, DC: American Association for the Advancement of Science.
- Teaching Effectiveness Program. (2014). *Writing multiple choice items that demand critical thinking*. University of Oregon. Retrieved from <http://tep.uoregon.edu/resources/assessment/multiplechoicequestions/sometechniques.html#problemsolution>
- Standards (2012). Retrieved from: www.corestandards.org/ October 30, 2012.
- Sternberg and Baron. (1985). A triarchic approach to measuring critical thinking skills: a psychological view. Paper presented at symposium, *Issues in the development of a Large-Scale Assessment of Critical Thinking Skills*. The American Educational Research Association annual Meeting. Chicago, Illinois.

- Stuart, S. A. J., Brown, M. I., & Draper, S. W. (2004). Using an electronic voting system in logic lectures: One practitioner's application. *Journal of Computer Assisted Learning, 20*, 95-102.
- Stupans, I. (2006). Multiple choice questions: Can they examine application of knowledge? *Pharmacy Education, 6*(1), 59-63. doi:10.1080/15602210600567916
- Torres, C., Lopes, A., Babo, L., & Azevedo, J. (2011). Improving multiple-choice questions. *US-China Education Review B1*, 1-11.
- Uhari, M., Renko, M., & Soini, H. (2003). Experiences of using an interactive audience response system in lectures. *BMC Medical Education, 3*(12). Retrieved from <http://www.biomedcentral.com/content/pdf/1472-6920-3-12.pdf>
- Wayne, W. (1982). Relative effectiveness of single and double multiple-choice questions in educational measurement. *The Journal of Experimental Education, 51*(1), 46-50.



Appendix A: Sample Factual Problem



Sample Higher Order Thinking Questions

Suppose a farmer has 1000 yards of fencing to enclose a rectangular field. What is the largest area that the farmer can enclose?

The first step of solving this problem is:

1. Set up a quadratic equation for the area
2. Set up an equation for the perimeter
3. Divide 1000 by 4
4. Find the area by multiplying 500x500

If l is the length and w is the width of the fenced part of the field and the total length of the fencing is 1000, then the equation for the perimeter is:

1. $4lw = 1000$
2. $2l \times 2w = 1000$
3. $2l + 2w = 1000$
4. $2(lw) = 1000$

The next step of solving this problem is:

1. Write an equation for the area of the fenced part of the field
2. Find the area of the whole field
3. Solve the equation for the perimeter
4. Divide the rectangular area into two triangles

The area (A) of the fenced part of the field is:

1. $A = w+l$
2. $A = wl$
3. $A = w/l$
4. $A = w(l+2)$

Then the next step of solving this problem is:

1. Solve the perimeter equation for either l or w so you can plug it into the area formula
2. Use a quadratic equation to solve for the area
3. You cannot find the area because the values for l and w are not known
4. Solve for either l or w in the area equation

In solving the equation $2w+2l=1000$ for w , the expression is:

1. $w = 1000-2l$
2. $w = 500-l$
3. $w = 1000-l$
4. $w = 500 - 2l$

Now substituting $w = 500 - l$ into the area formula we have:

1. $A(l) = l^2 - 500l$
2. $A(l) = 500 - l^2$
3. $A(l) = 500l - l^2$
4. $A(l) = l - 500 l^2$

When graphing the parabola that represents $A(l)$ the y-coordinate of the vertex will represent the greatest area.

1. True
2. False