Using WeBWorK, a Web-Based Homework Delivery and Grading System, to Help Prepare Students for Peer Instruction

Adam R. Lucas
Department of Mathematics and Computer Science
Saint Mary's College of California
Moraga, CA. 94575, USA
arl3@stmarys-ca.edu

Using WeBWorK, a Web-Based Homework Delivery and Grading System, to Help Prepare Students for Peer Instruction

Abstract: As an instructor who commonly devotes as much as a third of class time to clicker use and peer instruction, it is essential for me to cut down on time spent transferring information. To cover topics in sufficient depth I require students to carefully read the textbook before coming to class. To help them focus on key points in the reading and track their progress, I use WeBWorK, a system of online exercises. In this study I investigate the benefits and challenges of using WeBWorK in a lower division Introduction to Programming class and an upper division Probability and Statistics class. In the lower division class, WeBWorK significantly improved peer instruction performance and resulted in students reading on average 45% longer. WeBWorK and clicker scores correlate well with other measures of learning and provides early warning signs of students having difficulty.

Keywords:

WeBWorK, peer instruction, clicker, reading comprehension

1 INTRODUCTION

In the learning process, it is important to distinguish between the transfer of information and the assimilation or synthesis of that information [14]. Traditionally the transferring process happens during lecture and the assimilation process is left for the students to do after class in problem sets. One drawback in this approach is that students learn

more if they actively participate in their learning [10, 9, 13]. Furthermore, studies support that peer collaboration can help students achieve problem solving skills that they cannot acquire alone [19]. Using peer instruction (PI), teachers around the country are beginning to turn the traditional information transfer model upside down. The responsibility of gathering information now rests on the shoulders of the students. They must read material before coming to class, so that class time can be devoted to discussion, peer interaction and time to assimilate and think [15].

Figure 1: During the pair/share phase of Peer Instruction students are encouraged to discuss with pencil and paper with the peer directly across from them before sharing out to the larger group.

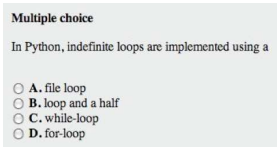
PI was developed in the 1990's at Harvard University by Eric Mazur. Although originally conceived as an interactive teaching method for introductory physics classes, it has proved successful in Mathematics and other disciplines [17, 16, 1, 11]. PI is a multi-step process. First, multiple-choice questions that are known as ConcepTests are presented on a classroom screen. Students possess their own pre-assigned clicker, a hand held radio frequency response device, with which they can anonymously input their selection. After a few minutes, a histogram reflecting how students voted is displayed on the screen. This process is referred to as the "think" phase. Students then form small groups where they actively discuss the question and answers for several minutes. Next, they re-vote and a new histogram, which now includes the correct an-

swer, is displayed. Students can be productively engaged during this "pair/share" phase as shown in Figure 1. The instructor then reviews the question and answer with the class as a whole.

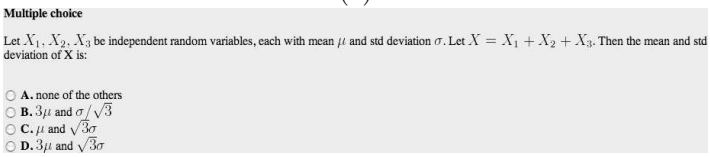
Since the clicker is pre-assigned to each student, the instructor is able to review the details of how each student voted. In my class students received one point for voting and an additional point for getting the right answer. The clicker software adds up points from the "think" and "pair/share" phases for each student and presents them as a cumulative average for the instructor to review. This is known as the "clicker score".

To help students with reading, before coming to class I have students answer multiple choice reading comprehension questions using WeBWorK. WeBWorK is one of a growing number of web-based homework systems available today. Arnold Pizer and Michael Gage in the Mathematics Department at the University of Rochester designed it specifically for mathematics in the mid 1990s. Today it is the largest open source homework system, free for instructors and students to use at over 150 colleges and high schools. In addition to multiple-choice responses, this system can grade free response numerical answers, free response answers involving mathematical expressions, and, in fact, any type of answer for which it is possible to write programmed instructions to determine correctness. Each student sees individualized problems, gets immediate feedback on whether or not the problem is correct, and is encouraged to continue reworking the problem until he or she gets the correct answer. Each assignment receives a "WeBWorK score" which reflects the number of questions answered correctly as well as the number of incorrect attempts. Using WeBWorK as a reading comprehension tool can be compared with the GoodQuestions project's method of running computer-grade pre-class reading quizzes before their calculus classes [16].

Figure 1a,b are WeBWorK reading comprehension questions I used in an Introduction to Computer Programming class and an upper level Probability and Statistics class at Saint Mary's College of California. The student is expected to answer these questions by referring to assigned reading in the textbook. An important area of current investiga-



(a)



(b)

Figure 2: A sample WeBWorK reading comprehension question used in my (a) Introductory Programming class, and (b) Probability and Statistics class.

tion is what kinds of pre-class reading quiz questions are likely to help students learn more from reading their textbook. Studies have shown that pre-class computational questions do a good job preparing students to answer in-class conceptual questions [2].

In this paper I will discuss the benefits and challenges of using WeBWorK for pre-class reading quizzes. Currently, there have been a number of studies done about the effectiveness of WeBWorK, used as a homework system, on student learning [3, 4, 5, 6, 7, 8, 12]. What is clear from these studies is that WeBWorK doesnt have a negative effect on student learning and in some cases has a positive one. The greatest gains come when WeBWorK is supplemented with a lab where students can get help entering their answers into the computer [8]. Because answering multiple choice questions with WeBWorK is straightforward, no lab in my classes was necessary. I will show that in a lower division Programming class there was a significant improvement in PI performance. I will also show that WeBWorK scores together with clicker scores provide the instructor with insight into which students are likely to continue having difficulty.

2 METHODS

Data sources included WeBWorK scores, clicker scores, video taped exit interviews, anonymous student surveys, end of semester student evaluations, and homework and test scores. Equipment used included a standard video camera, tripod, clicker hardware and software, and WeBWorK software. Information from student surveys, student evaluations, and exit interviews was used to determine the advantages and disadvantages of using WeBWorK. clicker scores were used to assess the effect of WeBWorK on PI performance. WeBWorK scores, clicker scores, homework, and test scores were correlated using Excel. I developed reading comprehension questions for my Probability and Statistics class and adapted them from the textbook's instructor supplements for my Introduction to Programming class.

My Introduction to Programming class consisted of 12 students with varied backgrounds and majors. Although the course has no prerequisites, students were expected to enter the class with a strong interest in learning how to program. After every class there was a (WeBWorK) reading assignment and a programming assignment. My Probability and Statistics class had 5 junior and senior Math majors. There were weekly homework assignments in addition to nightly WeBWorK reading assignments.

Classes were 60 minutes long and met three times a week. My classes didn't have a separate problem-solving workshop or lab. With rare exception a portion of each class was devoted to PI. I would devote the first 10 minutes of class to reviewing problems student had with WeBWorK or the homework. The next 30 minutes were devoted to explaining and illustrating the concepts from the reading, and the last 20 minutes were dedicated to PI. Usually there was only time for three or four ConcepTests. The use of PI in the last third of the class was meant both to test student's understanding of the reading and deepen that understanding.

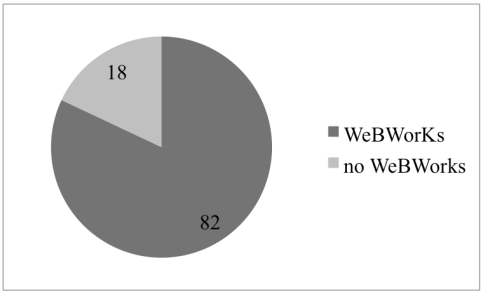


Figure 3: 82% of my students prefer having WeBWorK because it gives them an idea what the instructor wishes them to focus on in the reading and it gives them feedback on the correctness of their answers.

3 RESULTS

WeBWorK was popular with students once they were convinced of its benefits. Periodic anonymous self-surveys and an exit interview revealed that students' perception of WeBWorK changed during the course of the semester. At first many students saw WeBWorK on top of their regular nightly assignments as nothing more than busy work. For example, one student writes: *"WeBWorK just adds more work and doesn't help explain the material."* Halfway through the semester I experimented with requiring students to read the textbook without WeBWorK for several weeks. At the end of this trial period student began to appreciate having WeBWorK. One student for example writes: *"Before I really didn't like the WeBWorK questions and didn't feel they helped me at all. But I guess having that review at the end is nice."* Students wrote that they missed using WeBWorK questions as a guide as to what is important in the reading. They also realized that without something holding them accountable for reading, that in their busy schedules, they became less diligent about keeping up with the nightly reading. Students were glad when we began using WeBWorK again. Figure 3 shows that by the end of the semester, 82% of the students in my classes preferred having WeBWorK to no WeBWorK. To collect this data I asked them in an anonymous self-survey: *"Do you prefer having WeBWorK questions*

or no WeBWorK questions?".

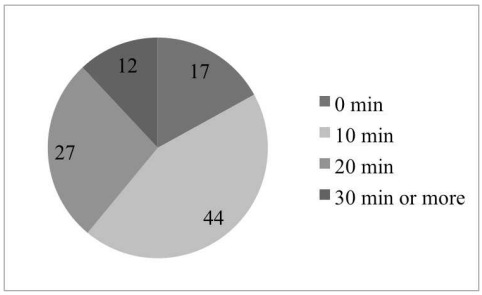
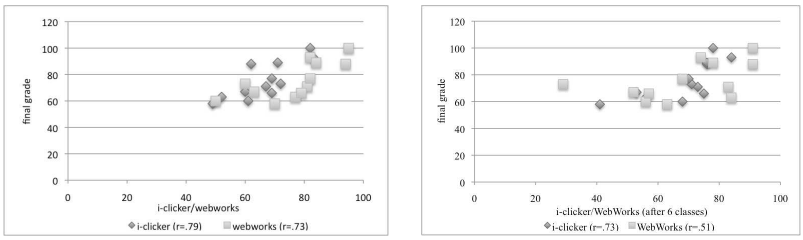


Figure 4: In my lower level class the majority of my students (83%) spent more time reading their textbook when they are required to answer WeBWorK questions before coming to class. 44% of my students read 10 minutes longer for each class using WeBWorK.

In each of my classes I had students keep track of how long they spent reading with and without WeBWorK. In my upper level class, there was no difference in time spent reading. These students claimed that they preferred having WeBWorK because it helped them focus on their reading but didn't feel it changed their study habits. Students in my lower level class, on the other hand, reported in an anonymous self-survey that they were reading, on average, 45% more using WeBWorK. Without WeBWorK students read on average 30 minutes per night. Figure 4 shows that 44% of my students increased their reading time by 10 minutes per night and 83% of students increased their reading time by 10 minutes or more per night.

3.1 WeBWorK And Clicker Scores Correlate Well With Overall Performance

I found that WeBWorK and clicker scores served as a meaningful measure of student performance in both classes. In my Introductory Programming class, WeBWorK and clicker scores had a correlation coefficient of $r = .73$ and $r = .79$ with student's final grade (see figure 5a). Here the clicker and WeBWorK component of the final grade was tem-



(a) (b)

Figure 5: (a) WeBWorK and clicker scores provide the instructor with meaningful data on student performance. (b) As soon as after 6 classes WeBWorK and clickers provide insight into which students are likely to have difficulty in the course.

porarily removed from the calculation. This compares with a correlation of $r = .77$ from daily homework assignments. The WeBWorK score, reflects a student's diligence outside of the classroom. Students who regularly read the textbook and are diligent about doing WeBWorK tend to have solid study skills and are motivated to succeed. WeBWorK scores however do not necessarily reflect student understanding. One student writes in an anonymous self-survey: *"When I don't understand what I read I usually don't understand WeBWorK, so I am basically just copying from the book."* WeBWorK's high correlation with overall performance was due, in part, to my reinforcing in class what students were reading. clicker scores, by contrast, give an indication of how much students are understanding from their reading and class discussion. Students who have natural ability to pick up concepts quickly, and a strong academic background, but have poor study skills, may have a high clicker and test scores but poor homework and WeBWorK scores.

Figure 5b shows the correlation of WeBWorK and clicker scores after the first 6 classes with overall performance. WeBWorK had a correlation of $r = .51$ and clickers had a correlation of $r = .73$. This compares to a homework score correlation of $r = .55$ after 6 classes. The lower correlation in WeBWorK is due in part to the small sample size (there was an outlier significantly affecting the statistics). Factors such as students not having the textbook for the first week, or being unaware of how WeBWorK was graded, probably also contributed. clicker scores on the

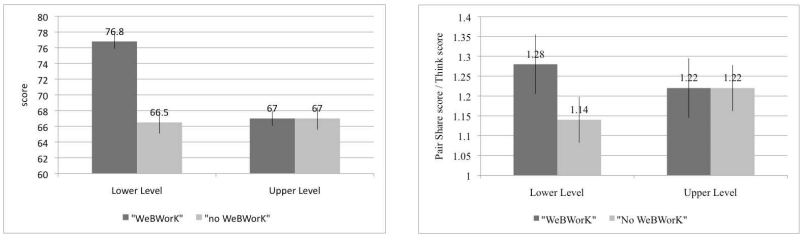
other hand remained a very good early indicator of future performance.

In contrast to homework, the WeBWorK and clicker data is available for the instructor to look at almost instantaneously. To have such meaningful data so quickly may help instructors identify weaker students early in the semester so that interventions can be made in a timely fashion.

3.2 Learning Outcomes

Does WeBWorK have a noticeable effect on PI performance? PI performance was measured in two ways. First, I looked at student's average clicker score since high clicker scores are indicative of students having a command of the course material. Second, I examined the average ratio of the "pair share" phase score to the "think" phase score. This ratio reflects student engagement and student learning. When students haven't done sufficient preparation for class, they don't have the confidence and background to contribute much in the "pair share" phase. As a result, students "pair share" score tends not to be substantially higher than their "think" score. I wrote questions with the aim of making them sufficiently difficult that only 40-50% of the class got the question correct during the "think" phase. This helped avoid situations where everyone got perfect scores on both phases. As well, I structured PI to discourage students from being led astray by higher status or dominant group members. Students need to be given explicit instructions as to how they should discuss the question and answers [11]. I instructed students to first discuss the question with one other student in detail, with the use of pencil and paper, before engaging in discussion with the group.

During a three week break from WeBWorK during the middle of the semester, PI performance was measured and compared to PI performance in the presence of WeBWorK. Figures 6a and 6b show that WeBWorK had a significant effect on PI performance in my lower level class. A t-test was used to determine a 95% confidence interval for all means. The increase in PI performance, I believe, resulted from students reading more and coming to class better prepared to engage. Students come to class knowing what topics in the reading are important and what



(a) (b)

Figure 6: (a) Average clicker score in my lower and upper level classes with and without WeBWorK. (b) Average ratio of "pair share" to "think" score in my two classes with and without WeBWorK. All error bars indicate a 95% confidence interval.

topics they need to clear up during class discussion. WeBWorK gives students automatic feedback so that they can test their understanding. This is especially important for students in lower level classes who lack experience reading mathematics.

WeBWorK didn't affect PI performance in my upper division class. I believe this is in part because these students were motivated to read independent of WeBWorK and had the experience to tell what they needed to understand better during class discussion. Another factor is that the students had difficulty reading the textbook and so even when using WeBWorK we would spend considerable time reviewing concepts in class before PI. As a result, students didn't need to rely on reading as much as in the lower division class where I didn't thoroughly cover all of the topics in the reading.

4 DISCUSSION

WeBWorK has both benefits and challenges which we discuss below.

4.1 Benefits Of WeBWorK

Focuses reading. Reading even one page of a dense mathematics text can be intimidated for many students. It is helpful having key words or phrases for students to recognize in the text. When WeBWorK is assigned regularly, students schedule time for it as part of their

homework. One student, for example writes: *" I like WeBWorK because it makes sure I understand the crucial parts of the reading, and that I actually devote time to doing it."*

Provides immediate feedback. By providing students with immediate feedback as to the correctness of their answers, students are encouraged to make multiple attempts until they succeed. One student says in an exit interview: *"If I think I understand something but I get it wrong I can go back and make sure I understand why the answer is the way it is."*

Provides accountability and motivation. WeBWorK individualizes problems to discourage cheating. Hence the instructor knows before entering class how prepared students are going to be for class. Students are motivated to earn "points" by doing WeBWorK.

Free to use. Unlike other web based homework systems, WeBWorK is free for students and instructors to use. There is extensive documentation on how to install WeBWorK at your home institution[1].

Allows coverage of topics in greater depth. By supporting students in their information gathering, more students benefit from the reading. Thus less class time needs to be spent on information transfer and more time can be spent on the synthesis process. I don't feel responsible for introducing every concept in class for which they are responsible and so I have time to pick and choose topics to discuss in greater depth.

Serves as an early warning sign. Together with clicker data, WeBWorK gives an early indication of which students most need help.

Promotes independent learning. Students using WeBWorK develop critical reading skills and get to know their textbook very well.

A launching pad for class discussion. I often begin class by projecting the previous nights WeBWorK questions on a screen. Students direct me as to the questions they want to discuss.

[1] see http://webwork.maa.org

4.2 Disadvantages Of WeBWorK

In spite of its many benefits WeBWorK isn't appropriate in every situation.

Requires a readable textbook. If students have great difficulty reading the textbook, WeBWorK becomes less effective. Students lose confidence when they repeatedly can't answer the WeBWorK questions and may stop doing it. Students have difficulty pinpointing their areas of confusion and as a result the instructor may find him or herself lecturing.

Extra work. Having pre and post class assignments can seem like a lot of work for students. A student writes in an anonymous self-survey: *"I like WeBWorK as long as it doesn't add a bunch of time to the homework."* For this reason, I wouldn't suggest assigning more than 8 questions per class. Students who automatically read the text book independent of WeBWorK might see it as busy work. One student for example writes: *"I feel like it is just more work. I already read ahead before WeBWorK and I feel now with it I am trying to use that time to do more problems on WeBWorK."* Students have given me grief over assigning so much work. However, at the same time they prefer having it, and understand that they cannot have PI, which they enjoy, without having pre-class reading assignments.

WeBWorK can also add to the workload of the instructor. It is possible to spend a lot of time creating WeBWorK questions for a single class. For this reason, I tend to choose a textbook with reading comprehension questions as an instructor supplement or find some other source of ready made WeBWorK questions.

4.3 Gathering WeBWorK Questions

I tend to choose textbooks that come with multiple choice reading comprehension questions. Converting these questions to WeBWorK's Perl based PG language amounts to cutting and pasting into a template (provided by WeBWorK) for multiple choice type questions. As well, there

is a large national database of WeBWorK questions tied to specific textbooks. Although these questions are primarily designed as homework questions, there likely are appropriate reading comprehension problems. It is also possible to adapt questions from ConcepTests. There is a growing archive of good ConcepTest questions. Readers are invited to explore the supplemental ConcepTest book accompanying the Hughes-Hallett's Calculus book [18], the article on Good Questions [16], the Project Math QUEST website[2], and my PI website[3].

4.4 Optimizing The Use Of WeBWorK

At the beginning of the semester the instructor should explain to students the benefits that WeBWorK will provide them. I have found that students really enjoy peer instruction and using clickers. I tell them that we cannot use clickers without doing WeBWorK because there won't be enough time to cover the necessary material. WeBWorK will help them know what to focus on in the reading. As well, I tell them that they can earn participation "points" by doing WeBWorK regularly. Even if the points are minimal (in my case at most 10% of their final grade), it gives students incentive to do it.

One of the biggest complaints I get from students about WeBWorK has to do with its grading. There is a fine line between making WeBWorK extra credit and making it required. On the one hand students wanted to feel free to learn from their mistakes and not treat WeBWorK as a pre-class reading quiz. On the other hand, students are less likely to do WeBWorK if it isn't required. One student for example writes: *"The pressure to pass without mistakes for a good grade overtakes the learning aspect. Students look up the answers, but don't learn them. The care shifts from learning to passing."* Some students suggest that it be worth participation points but not be graded based on failed answers. So long as it is completed, students should get credit. One student writes in his/her course evaluation *"WeBWorK should be used only to help learning not as a quiz."* My compromise was to give students full credit if

[2]http://mathquest.carroll.edu
[3]http://www.cfkeep.org/users/alucas/iclicker

they completed WeBWorK and worth extra credit if they do it well.

It is advantageous to make WeBWorK due the night before class. It gets students in the habit of organizing their time so that they aren't doing homework in the middle of the night. As well it gives the instructor a chance to design the next day's lectures in response to their performance. I start class by projecting the WeBWorK questions on a screen and letting students direct me as to what they want to discuss. This gives students a sense of control and makes the class more student centered.

It is important to view WeBWorK as a supplement not a substitute to lecturing. Students made it clear to me that they didn't feel prepared to engage with clickers after using WeBWorK alone. One student wrote in an anonymous self-survey: *"WeBWorK doesn't eliminate the need for lecturing at all, because there still will be things that aren't understood, but it helps me to pick out the most important concepts from the text and to make sure they are understood at least."* Many students indicated they wished that at the end of class that I would spend 5-10 minutes introducing the topics that they will encounter next in the reading. Although this may be a good idea, in practice I rarely have time to do this.

4.5 Limitations Of The Study

The biggest limitation of this study is the small size of the classes involved (5 and 12 students). Indeed an outlier can potentially throw off the statistics, as is what happened in Figure 5b. The study should be particularly useful however to instructors with small math classes. Note that to make Figure 6 clickers were used nearly every day during the semester resulting in large data sets.

5 CONCLUSIONS

WeBWorK is a useful tool for instructors using clickers and PI. For one, it gives the instructor control over how students prepare for class. The instructor can choose what key points he/she wants the students to take

away from the reading. Second, it gives the student support in terms of feedback and automatic grading. This makes WeBWorK popular among students. Third, it provides an early warning sign for students having difficulty. Most importantly, WeBWorK allows the instructor to concentrate class time on information assimilation and synthesis rather than information transfer.

WeBWorK has a positive impact in terms of PI performance in lower division classes. In these classes students need the motivation and support that regular WeBWorK reading assignments provide them. However, WeBWorK benefits are diminished if the students have difficulty reading the textbook.

From this self study I have a better appreciation of the challenges of making students independent learners. As a next step I wish to better understand how to design questions that give me the most information about student's understanding of the reading. I would also like to study what kinds of pre-class reading questions are likely to help students learn more from reading their textbook.

ACKNOWLEDGEMENTS

Supported by an Alumni Faculty Grant and Computing and Technology Services at Saint Mary's College of California. The author would like to thank Derek Bruff and Kelly Cline for helpful comments.

REFERENCES

[1] Bruff, D. 2009. *Teaching with classroom response systems: creating active learning environments*. Vanderbilt University, The Jossey-Bass higher and adult education series.

[2] Bruff, D. *Pre-Class Reading Assignments in Statistics*. `http://www.cfkeep.org/html/snapshot.php?id=81181051005778`. Accessed August 2009.

[3] Cassady, J. C., J. Budenz-Anders, G Pavlechko, and W. Mock. 2001. *The Effects of internet-based formative and summative as-*

sessment on test anxiety, perceptions of threat, and achievement. Annual Meeting of the American Educational Research Association. Seattle, WA. 1-12.

[4] Denny, J., and C. Yackel. *Implementing and teaching with WeBWorK at Mercer University.* Proceedings of the 2005 ASCUE Conference. 2005. 85-93.

[5] Dufresne, R., J Mestre, D. M. Hart, and K. A. Rath. 2002. The effect of web-based homework on test performance in large enrollment introductory physics courses. *Journal of Computers in Mathematics and Science Teaching.* 21(3): 229-251.

[6] Gage, M., A. K. Pizer, and V. Roth. 2003. *WeBWorK: generating, delivering, and checking math homework via the internet.* In Proceedings of the Second International Conference on the Teaching of Mathematics. New York: Wiley.

[7] Hauk, S., and A. Segalla. 2005. Student perceptions of the web-based homework program WeBWorK in moderate enrollment college algebra classes. *Journal of Computers in Mathematics and Science Teaching.* 24(3): 229-253.

[8] Hirsch, L., and C. Weibel. Statistical evidence that web-based homework helps. *MAA Focus.* 23 (2003): 14.

[9] Johnson, D.W., R.T. Johnson and K.A. Smith. 1998. Maximizing Instruction Through Cooperative Learning. *ASEE Prism* 7(6): 24-29.

[10] Johnson, D.W., R.T. Johnson, and K.A. Smith. 1991. *Active Learning: Cooperation in the College Classroom.* Interaction Book Company.

[11] Lucas, A. 2009. Using peer instruction and i-clickers to enhance student participation in Calculus. *PRIMUS.* 19(3): 219-231.

[12] LaRose, P. G., and R. Megginson. 2003. Implementation and assessment of on-line gateway testing at the University of Michigan. *PRIMUS*. 13(4): 289-307.

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

[14] Mazur, E. 1997. *Peer Instruction: A User's Manual*. Englewood Cliffs NJ: Prentice Hall.

[15] Mazur, E. 2009. Farewell, Lecture? *Science*. 323(2): 50-51.

[16] Miller, R.L., E. Santana-Vega, and M. Terrell. 2006. Can Good Questions and Peer Discussion Improve Calculus Instruction? *PRIMUS*, 16(3): 193-203.

[17] Pilzer, S. 2001. Peer Instruction in Physics and Mathematics. *PRIMUS*. 11(2): 185-192.

[18] Pilzer, S., M. Robinson, D. Lomen, D. Flath, D. Hughes Hallet, B. Lahme, J. Morris, W. McCallum, J. Thrash. 2003. *ConcepTests to Accompany Calculus*, Third Edition. Hoboken NJ: John Wiley & Son.

[19] Vygotsky, L.S. 1978. *Mind in Society: Development of Higher Psychological Processes*. Harvard University Press. 14th ed. March 7: p. 86.

BIOGRAPHICAL SKETCH

Adam Lucas is currently an Assistant Professor of Mathematics and Computer Science at Saint Mary's College of California. He is also a Research Associate in the Department of Pharmaceutical Chemistry at UCSF. After a degree in Biochemistry from McGill University he did his PhD in the area of Representation Theory at MIT. In 1999, he received an NIH postdoctoral fellowship to study protein folding at UCSF. Adam's current interests include polymer statistical mechanics,

complex networks, and the use of peer instruction and WeBWorK in mathematical education.