Peer Instruction in Computing: the Role of Reading Quizzes

Daniel Zingaro
Department of Computer Science
University of Toronto
Toronto, ON, Canada
daniel.zingaro@utoronto.ca

Cynthia Bailey Lee
Computer Science and Engineering Department
University of California, San Diego
La Jolla, CA, USA
clbailey@cs.ucsd.edu

Leo Porter
Department of Mathematics and Computer Science
Skidmore College
Saratoga Springs, NY, USA
leo.porter@skidmore.edu

ABSTRACT

Peer Instruction has recently gained interest in computing as an effective active learning pedagogy. The general focus of PI research has been on the in-class portion of PI: multiple choice questions and group discussion. Here, our focus is the reading quizzes completed by students for purposes of class preparation. These quizzes contain content questions but also ask for difficulties or confusion with course material. Consistent with expectations, we demonstrate that providing correct responses to quiz questions positively correlates with other course assessments. Somewhat counter-intuitively, we find that identifying confusions, noting problematic sections, or asking questions about the reading are also correlated with lab grades.

Categories and Subject Descriptors

K.3.2 [**Computer Science Education**]: Computer and Information Science Education

Keywords

CS1; peer instruction; clickers; classroom response; active learning

1. INTRODUCTION

Peer Instruction (PI) is a pedagogical technique developed in physics that has since been used with considerable success in computing. Physics educators realized that standard lectures are ineffective for teaching core concepts and addressing misconceptions [4], and PI has been shown to remedy such concerns as measured through course-based assessments and standardized concept inventories. The core of PI is the multiple-choice question (MCQ) posed by instructors and answered by students. Lecture meetings hence become sequences of posing questions, answering questions, and discussing the questions in small groups. As detailed below, such activity driven by MCQs has received significant attention from those CS educators interested in PI. However, PI consists of much more than this "MCQ core". In particular, in order to engage in productive in-class discussions, students are expected to complete reading and a corresponding assessment prior to each class. Typically, such quizzes, like the MCQs they portend, are graded based on participation and not correctness. To what extent do students complete such reading quizzes? Do students tend to provide complete (correct or incorrect) responses or provide only minimal information? Do students acknowledge confusion or difficulty, and how does this relate to in-class or course-based performance? We seek to broaden understanding of PI in CS through such analyses of reading quizzes. In particular, we demonstrate that quiz completion and quiz correctness positively correlate with other areas of course performance. In addition, we demonstrate that the ability or willingness to identify confusions about the reading also positively correlates with other measures of course performance.

SIGCSE'13, March 6–9, 2013, Denver, Colorado, USA.

2. WHAT IS PI?

Each PI meeting consists of the following steps. The instructor begins with a mini-lecture that quickly overviews what the students read for the pre-lecture quiz. Then, the instructor poses a ConcepTest [4]—a multiple-choice question designed both to focus attention on and raise awareness of the key course concept from the mini-lecture [1]. After individually thinking about and voting on the correct answer (the solo vote), students discuss the question in small groups, reach a consensus, and vote again (the group vote). Students are encouraged to discuss each answer, verbalizing why it is correct or incorrect [16].

While student voting patterns can be estimated using flashcards [10], instructors can more easily obtain and communicate clicker data. Clickers also afford immediate, accurate vote counts to the instructor, as well as compelling graphical displays for the students. Ideally, each response option for a ConcepTest will correspond to a common student misconception, so that the instructor can identify the range of understandings present among the students [6]. Seeing the histogram of results, students come to realize that they are not alone in their confusion [9].

Following the group vote and displaying of response graphs, the instructor leads a class-wide discussion. Students are asked to support various response options as the class moves toward consensus. Once the discussion is complete, a new PI cycle begins.

3. WHY PI IN COMPUTING?

Since 2010, many investigations of PI in CS have been published. The first wave of papers focused exclusively on the gains exhibited between the solo and group votes of an MCQ [16, 18]. To measure such improvements, the normalized gain metric (NG) is frequently used. NG captures the proportion of students who answer incorrectly in the solo vote but correctly in the group vote. For example, if 60% of students answer correctly in the solo vote and 80% answer correctly in the group vote, the NG for the question is 50% (i.e. 50% of "potential learners" demonstrated new understanding). Representative findings include NGs of 41% in CS1 [16] and 29% in a CS1 where all students were taking the course for the second time [18]. These NG increases are particularly promising in light of the fact that peer discussions generally last only a couple of minutes, suggesting that learning occurs in a relatively brief timeframe.

Unfortunately, NG does not discriminate active learning from passive copying. Perhaps NG increases are due partly or mostly to students copying from peers whom they perceive as more knowledgeable. To help determine the extent to which the cited NG measures reflect one or the other of these scenarios, more recent papers have measured learning through the use of *isomorphic questions*. The methodology was first used in 2009 in the context of a biology course [17]. A pair of isomorphic questions tests the same concept, but the questions have different cover stories. In CS, this might reflect two questions both testing understanding of variables, but with different values being assigned or variable assignments occurring in different orders. Answered individually, the second question of a pair can give an idea of what students can now do on their own, outside of their PI groups. In both biology and CS [17, 14], isomorphic questions have been used to demonstrate that significant learning takes place during peer discussion.

The present wave of CS-PI research moves beyond per-question gains to an understanding of how PI as a whole confers conceptual understanding to students. One recent argument suggests that PI enables a form of cognitive apprenticeship, where the MCQs set by the instructor can help inculcate students into the professional discipline of CS [5]. When instructors use MCQs to lead students toward the types of thinking engaged in by experts, PI can provide opportunities for the deliberate practice and scaffolding that have been lauded by decades of educational research [7]. Yet, all of this research presupposes, at least implicitly, the requirement for students to be well-prepared to deliberately engage with PI materials. There is little time to "lecture" in a PI class, so preliminary material cannot be taught in the traditional way.

In the physics PI literature, student preparation typically comes in the form of reading quizzes [4], and it is likely that student experiences on reading quizzes are important in order for the core of PI to be effective. Authors of seminal physics PI research [3] note that it can be effective to administer short, three-question web-based quizzes prior to each lecture. The first two questions relate directly to the content covered in the required pre-class reading, and the third question asks, "what did you find difficult or confusing about the reading? If nothing was difficult or confusing, tell us what you found most interesting. Please be as specific as possible." Instructors can use the responses to these quizzes to support a form of just-in-time teaching whose focus is problematic areas reflected in students' quiz responses.

Though the use of reading quizzes is a PI best-practice, CS-PI courses sometimes do not use such quizzes [15]. When CS instructors do administer reading quizzes, research publications tend to jump straight to an analysis of in-class MCQs without reference to the preliminary reading quizzes. For example, [18] note that reading quizzes were used in the style of [3] and offer a sampling of the types of questions included in the quizzes, but say little in regard to correctness, completion, effort, acknowledgment of confusion, and so on. That said, we do know that students find reading quizzes valuable. For example, in work using clickers, occasionally in the PI format, it was found that 76% of Data Structures students agreed that reading quizzes helped them recognize and focus on difficult course concepts [13].

One recent paper suggests that the form of reading quizzes in CS should leverage the activity-based nature of our subject [8]. Rather than decoupling reading from experimentation, these authors support students in experiment-based learning where book and programming environment are used concurrently. Interspersed with reading, students are guided to emulate expert-like behavior: identifying learning goals, exploring small code segments, making predictions, and so on. Then, at the start of each lecture, students are quizzed on the homework material before the instructor launches into the first PI cycle. Student usage and valuation of the *exploratory homeworks* were promising: 50% of students "almost always" or "always" completed the homeworks; 58% of students "always" felt that homework helped them understand lecture; and those who read the text and wrote code before, during or after that reading outperformed the students who did not write the code. At the same time, only 30% of students completed the homeworks as intended: flexibly moving between reading, experimenting, and exploring (rather than, for example, serially reading and then coding).

The work of [8] is similar to the present work, but with two important differences. First, those authors used exploratory homeworks rather than standard PI reading quizzes, and it is the latter that is currently in more common use. In particular, these authors did not seek input from students on what they find difficult or confusing, which is a key role of PI reading quizzes. Second, they did not report on actual student performance on the quizzes, instead relying on student perceptions of the utility of doing the homework. Here, we seek to explore traditional PI reading quizzes by examining completion, correctness, confusion, and the ways in which these variables relate to course-based performance.

4. SETTING

We implemented PI in a small (40 students) remedial CS1 course for Computer Engineers taught by the primary author at a research-intensive Canadian university. All students in the class were unsuccessful in completing a traditional (non-interactive) 12-week version of the same course the previous semester, and require this course to continue in the engineering program. We used C to teach the major topics of the course: selection, iteration, functions, arrays, recursion, sorting, pointers and memory allocation (linked lists were taught in the prior offering but not in this remedial offering). We met for three fifty-minute lectures each

1 Is the expression 'a' < 'C' true (1) or false (0). Please carefully explain your answer.
2 The following program does not run.

```
#include <stdio.h>

int main (void) {
  int q = 4;
  if (q == 4)
    printf ("The if-statement is running.\n");
    printf ("Another printf.\n");
  else
    printf ("This is the else.\n");
  return 0;
}
```

Try to compile the program. Explain each error that your compiler gives, and suggest a solution.
3 What did you find particularly difficult or confusing about the material you read? If nothing was difficult or confusing, tell us what you found most interesting. Please be as specific as possible: this is meant to help you focus on problematic areas, as well as help me make more meaningful lectures.

Figure 1: The fifth reading quiz.

week for 10 weeks, and were supported by weekly practical lab sessions and weekly tutorials.

Our motivation for using PI in this remedial setting was based on a reported interaction between pedagogy (PI or traditional) and experience [11]. This study found that students with low background knowledge gain as much from PI as students with high background knowledge gain from traditional instruction. It is likely that students in our remedial CS course lack significant background knowledge; therefore, we believed that PI would be particularly effective for achieving learning gains commensurate with those of the successful students from the semester prior.

Before each class, students were required to read one or two sections of our textbook [2], and complete an online reading quiz, typically with three questions. We used the format suggested in [3] and described above (i.e. two or three content questions followed by the "confusion question"). 6% of students' grades were allocated to completion of the quizzes: marks were awarded for effort, not correctness. In general, reading quiz questions were designed to skim the surface of what the students had read, while deeper application of that material was designated for in-class ConcepTests. Figure 1 contains one example reading quiz from week 2; the remainder are available on a PI resource page [19]. An additional 6% of students' grades were allocated to participation in ConcepTests and, again, marks were awarded for effort, rather than correct answers. Further details of the class offering can be found in [18].

5. DATA ANALYSIS AND RESULTS

5.1 Completion and Correctness

In total, there were 27 reading quizzes. On average, students submitted 21.2 reading quizzes, for an average submission rate of 78.5 percent. As will be shown, some of these "submissions" contained empty (or essentially empty) responses.

The reading quizzes contained a total of 52 content-based questions (i.e. not including the "confusion question"). To investigate the quality of the responses submitted, the primary author (also the course instructor) graded each response to a content-based question on a four-point scale:

Table 1: Average per-student responses on content-based questions, broken down by response category.

Category	Percentage
G0-Empty	1.8
G1-Incomplete	10.5
G2-Incorrect	37.6
G3-Correct	50

- G0-Empty: Essentially no response (e.g. submitting "no" or "not sure", or leaving the response completely blank).
- G1-Incomplete: something submitted, but very incomplete (e.g. answering half of a question or tracing only partway through a program).
- G2-Incorrect: complete response, but incorrect.
- G3-Correct: complete and correct response.

Table 1 contains the average per-student percentage of responses that fell in each of these four categories.

G2-Incorrect and G3-Correct reflect complete responses to the question. In the case of G2-Incorrect, students worked through a problem with misconceptions that led them toward incorrect conclusions, but the execution was nevertheless complete. For this reason, we argue that 87.6 percent of responses may be educationally useful as they represent students engaging with questions prior to lecture. 12.3 percent of responses, plus the questions not submitted at all (21.78 percent), may suggest inadequate preparation for lecture. Lastly, we see the rate of G3-Correct responses as surprisingly high: students knew that they were being graded only on submission (not correctness), so there was no immediate mark-based benefit for students to answer questions correctly. In addition, these are students that struggled in the past, and it is encouraging that half of the responses submitted by such "weak" students are correct with no further instructor intervention.

5.2 Confusion Question

On all but two of the reading quizzes, we provided the "confusion question" as the final question for students to answer. As indicated in Figure 1, we suggested to students that their responses might help the instructor create more accurately-targeted lectures. There is also recent evidence that students who are able to acknowledge confusion or difficulty can be well-positioned to excel in post-secondary courses [12].

To determine the extent to which students use this question for such purposes, we read through the responses searching for themes. Themes were initially based on the wording of the question itself, and were refined until all data could be captured and no themes failed to represent data. The following themes emerged:

C0-Interesting: identifies something interesting.

C1-Not Interesting: states that nothing in the reading was interesting.

C2-Confusing: identifies something confusing or unclear.

C3-Not Confusing: states that nothing in the reading was confusing.

C4-Difficult: identifies something that is difficult or "hard".

C5-Not Difficult: states that nothing in the reading was difficult.

C6-Forward: identifies something difficult that has not yet been reached in the course. (This is an artifact of the students having taken the course unsuccessfully once before.)

C7-Problematic: identifies something that cannot be understood or that something is proving problematic or troublesome. (This likely overlaps with student conceptions of difficulty, though with different word choice. For example, students often used phrases such as "I did not understand" or "I had trouble with this concept". When they did not explicitly use the word "difficult" or "hard", we coded the response as problematic rather than difficult.)

C8-Not Problematic: explicitly states that something is understood.

C9-Questioning: poses a question without explicitly indicating confusion or difficulty.

C10-Ambiguous: identifies a topic, but does not indicate whether that topic is interesting, confusing, difficult, etc. For example, a student might indicate simply "recursion".

C11-Other: comments relating to midterms or exams, or comments too short to otherwise classify (e.g. "so far so good").

C12-Empty: completely blank submission or a submission containing no information (e.g. "N/A", "nothing", "no").

Figure 2 provides the average per-student percentages of these codes in the 25 reading quizzes where the "confusion question" was asked. A total of 786 student responses were coded. The sum of the averages is more than 100% because responses could be coded by multiple codes (e.g. a response that asks a question but also acknowledges confusion).

Striking in this data is the number of questions posed to the instructor by students, an indication that areas of difficulty were not only uncovered but explicitly framed in the form of questions to which the instructor could respond. As a case in point, the responses to the quiz in Figure 1 contained eight instances of students asking questions. Two of these asked about the C `switch` statement that was topical but not tested in this quiz. A further two questions asked about the code that had been supplied, and why the code works at all without a `scanf` statement. The remainder of the questions asked how to chain `else if` branches. In our experience, students simply do not ask about course material to this degree when reading quizzes are not used; in such cases we find the vast majority of questions specifically focused on the students' current assignment rather

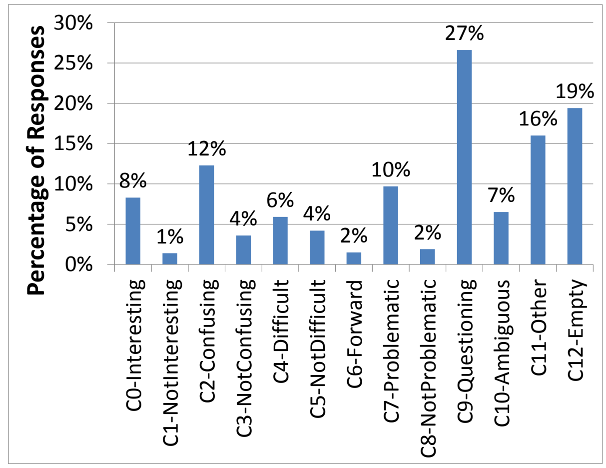

Figure 2: Average per-student response percentages, broken down by codes on the 786 submitted responses to the "confusion question".

than course material per se. We hypothesize that reading quizzes have the desirable effect of encouraging students to ask questions continuously, rather than clustering questions around assignment deadlines.

Unfortunately, the mingling of confusing, difficult, and interesting in the question text also led to 51 ambiguous responses. For example, again in the responses to the quiz in Figure 1, one student provided simply "the flowcharts". This referred to the approach used in our text that depicted control statements in the form of flowcharts, though whether this was problematic or interesting is unclear.

The codes representing students who are not confused or who found nothing difficult mostly reflect a rephrasing of the question text. That is, we received many responses of the form "nothing was confusing or difficult". Given no further elaboration, it is not clear whether this was an attempt to submit content for the question (rather than leaving it blank), or whether they genuinely believed that nothing was confusing or difficult.

The 122 C11-Other responses were made mostly of high-level comments on the course, students' favourite C constructs, reflections on the prior course offering, and suggestions for how lectures could be improved (though of course without providing explicit questions).

5.3 Correlations with Course Performance

Table 2 gives the significant Spearman correlations ($p < .05$) between reading quiz codes and final exam grade. The final exam was worth 45% of students' grade. Perhaps unsurprisingly, students submitting many G1-incomplete responses do poorly on the final exam, and those submitting many G3-Correct responses do well. G2-Incorrect was negatively but insignificantly correlated ($r = -.23$) with final exam grade. None of the codes from the "confusion question" were significantly correlated with exam performance, though C7-Problematic approached significance ($p = .08; r = 0.28$).

Students' unsupervised term work (worth 18% of the final grade) involved the completion of five labs on which students

Table 2: Significant Spearman correlations between reading quiz codes and final exam grade.

Code	Correlation
Content responses submitted	0.32
G1-Incomplete	−0.37
G3-Correct	0.51

Table 3: Significant Spearman correlations between reading quiz codes and unsupervised grade.

Content responses submitted	0.8
G0-Empty	−0.35
G3-Correct	0.72
C2-Confusing	0.4
C7-Problematic	0.31
C9-Questioning	0.57

worked alone. Table 3 provides the significant Spearman correlations ($p < .05$) between codes on the reading quizzes and students' unsupervised term work grade. Number of content responses submitted, a clear effort measure, had the strongest correlation with unsupervised term work. Unlike with final exam score, students willingness to ask questions and acknowledge difficulty and confusion were significantly correlated with unsupervised work. In addition, G2-Incorrect was positively but insignificantly correlated with unsupervised term work ($r = 0.29$).

5.4 Correlations with Clicker Scores

Clicker responses, like reading quizzes, were assessed by contribution rather than correctness. Still, we wondered to what extent scores on reading quizzes correlated with clicker correctness. Our hypothesis was that students who did well on reading quizzes would be well-prepared for lecture and hence correctly answer more clicker questions. We focus here only on solo votes (i.e. those votes occurring prior to group discussion). Our rationale for excluding the group vote is that such scores comingle students' incoming knowledge and knowledge obtained through group discussions. We found significant positive correlations between number of clicker questions answered correctly and both the number of submitted content responses ($r = 0.65$) and number of G3-Correct submissions ($r = 0.48$). There was also a significant correlation between clicker correctness and C9-Questioning ($r = 0.31$). Interestingly, though nonsignificant, there was a moderate negative correlation (-0.16) between clicker correctness and C3-Not Confusing. There were two other insignificant but positive correlations with clicker correctness: C7-Problematic ($r = 0.23$) and C2-Confusing ($r = 0.22$).

6. DISCUSSION

The fact that students submitted almost 80 percent of reading quizzes and that 87.6 percent of question submissions were of acceptable quality suggests that students took reading quizzes seriously. There was no obligation for students to answer questions completely; the instructor explicitly told students that any submission counts. It is heartening that students took it upon themselves to put forth effort rather than simply answering for the sake of answering. This corroborates an end-of-term survey finding reported in [18] where 92 percent of these same students agreed with the claim that reading quizzes helped identify difficult concepts in the reading. The current paper strengthens the belief that students find reading quizzes valuable.

The "confusion question" provided useful data in terms of student questions and confusions, though the wording of the question sometimes led to responses that could not be unambiguously categorized. We argue that receiving this quantity of questions and feedback is further evidence that students cared about the quizzes and used the opportunity to communicate with the instructor. The course instructor was careful to address this feedback personally or at the beginning of the next lecture, which we suspect contributed to the quantity of feedback received. Though feasible in small classes, the workload involved in responding personally to quiz submissions in large classes is likely prohibitive. That said, we believe it is important for students to understand that we take quizzes seriously and treat them as a core component of PI practice.

There were no relationships between codes on the "confusion question" and final exam performance. Yet, codes related to confusion and questioning were highly correlated with unsupervised term work, affirming the relationships discussed by [12]. The reason for this difference is unclear and deserves future consideration. We do note that confusions were largely local ("pointers") rather than global ("what is the relationship between the heap, pointers, and the memory model"), and so perhaps related more strongly to small-scale assignments rather than the integrative exam. Also, assignments provide students some time in which to acknowledge confusions, obtain answers or clarifications, and then apply that knowledge in the context of their deliverable. Such feedback-driven progress is not possible in an exam setting.

Correlations between reading quiz correctness and clicker questions suggest that the quizzes might prepare students for and increase learning in the course. Importantly, we cannot rule out alternate hypotheses: it is equally plausible, for example, that stronger students do well on both quizzes and clicker questions, or that those putting forth more effort on the quizzes also put forth more effort toward answering clicker questions correctly. The relationship does give us confidence that students are finding reading quiz questions and clicker questions to be similar in terms of content and/or difficulty.

7. FUTURE WORK

There are several areas requiring future work. First, we suggest that the "confusion" question be disaggregated to make interpretation of responses more straightforward. For example, the question could be divided into prompts that ask for confusing, difficult, and interesting aspects of the course reading; students could be required to respond to at least one of these prompts. Students could also be provided a space explicitly for asking questions. An analysis of responses to these new questions would be interesting as a comparison to what we have offered here. Second, it will be interesting to compare traditional PI reading quizzes as discussed here and other preparatory assignments such as exploratory homeworks, effectively bridging the span between our work and that presented in [8]. Third, student perceptions of reading quizzes would be useful in the context of our correlation-based causation hypotheses. Do students feel that reading quizzes prepare them for lectures, labs, exams, or the course in general? Do they see relationships between quiz questions and clicker questions? Further student perceptions of reading quizzes are required to determine the accuracy of our correlation-based causation claims.

8. CONCLUSION

Analysis of reading quizzes is underaddressed in the current PI literature in computing. We provide one such analysis here, focusing on completion, correctness, confusion, and relationships to course-based assessments and PI questions. We find that students complete a large portion of reading quiz questions and that their responses suggest a reasonable level of effort and correctness. Students' responses to the "confusion question" demonstrate similar levels of attention: students ask many questions, acknowledge significant confusion, and note particular difficulties. We demonstrate that reading quiz performance correlates with performance in other areas of the course, suggesting that reading quizzes target important course material and may confer advantages to students who expend (or are willing to expend) the effort required to answer the reading quizzes correctly. We also find correlations between identifying confusions and questions on the one hand and performance on unsupervised term work and clicker questions on the other. We encourage further investigations into the utility of reading quizzes, their relationship to learning and the rest of the course, and their applicability to other CS courses in which PI is implemented.

9. REFERENCES

[1] I. D. Beatty, W. J. Gerace, W. J. Leonard, and R. J. Dufresne. Designing effective questions for classroom response system teaching. *American Journal of Physics*, 74:31–39, 2006.

[2] J. Carter. *Introduction to Computer Science Using C.* McGraw-Hill Ryerson, 2008.

[3] C. H. Crouch and E. Mazur. Peer instruction: Ten years of experience and results. *American Journal of Physics*, 69:970–977, 2001.

[4] C. H. Crouch, J. Watkins, A. P. Fagen, and E. Mazur. Peer instruction: Engaging students one-on-one, all at once. In E. F. Redish and P. J. Cooney, editors, *Research-Based Reform of University Physics.* American Association of Physics Teachers, 2007.

[5] Q. Cutts, S. Esper, M. Fecho, S. R. Foster, and B. Simon. The abstraction transition taxonomy: Developing desired learning outcomes through the lens of situated cognition. In *ICER '12: Proceedings of the Eighth international Workshop on Computing Education Research*, New York, NY, USA, 2012. ACM.

[6] Q. Cutts, G. Kennedy, C. Mitchell, and S. Draper. Maximising dialogue in lectures using group response systems. 7th IASTED international conference on computers and advanced technology in education. `www.dcs.gla.ac.uk/~quintin/papers/cate2004.pdf` (accessed August 19, 2011)., 2004.

[7] K. A. Ericsson, R. T. Krampe, and C. Tesch-Roemer. The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, 100(3):363–406, 1993.

[8] S. Esper, B. Simon, and Q. Cutts. Exploratory homeworks: An active learning tool for textbook reading. In *ICER '12: Proceedings of the Eighth international Workshop on Computing Education Research*, New York, NY, USA, 2012. ACM.

[9] J. K. Knight and W. B. Wood. Teaching more by lecturing less. *Cell Biology Education*, 4:298–310, 2005.

[10] N. Lasry. Clickers or flashcards: Is there really a difference? *The Physics Teacher*, 46:242–244, 2008.

[11] N. Lasry, E. Mazur, and J. Watkins. Peer instruction: From Harvard to the two-year college. *American Journal of Physics*, 76:1066–1069, 2008.

[12] E. Mazur. ICER 2011 keynote address, 2011.

[13] R. P. Pargas and D. M. Shah. Things are clicking in computer science courses. In *SIGCSE '06: Proceedings of the 37th SIGCSE technical symposium on Computer science education*, pages 474–478, New York, NY, USA, 2006. ACM.

[14] L. Porter, C. Bailey-Lee, B. Simon, and D. Zingaro. Peer instruction: Do students really learn from peer discussion in computing? In *ICER '11: Proceedings of the Seventh international Workshop on Computing Education Research*, New York, NY, USA, 2011. ACM.

[15] B. Simon and Q. Cutts. CS principles pilot at University of California, San Diego. *ACM Inroads*, 3(2):61–63, June 2012.

[16] B. Simon, M. Kohanfars, J. Lee, K. Tamayo, and Q. Cutts. Experience report: Peer instruction in introductory computing. In *SIGCSE '10: Proceedings of the 41st SIGCSE technical symposium on Computer science education*, pages 341–345, New York, NY, USA, 2010. ACM.

[17] M. K. Smith, W. B. Wood, W. K. Adams, C. Wieman, J. K. Knight, N. Guild, and T. T. Su. Why peer discussion improves student performance on in-class concept questions. *Science*, 323:122–124, 2009.

[18] D. Zingaro. Experience report: Peer instruction in remedial computer science. In *Ed-Media 2010: Proceedings of the 22nd World Conference on Educational Multimedia, Hypermedia & Telecommunications*, pages 5030–5035. AACE, 2010.

[19] D. Zingaro. Pi-cs resource page. `www.danielzingaro.com/pics.php`, 2012.