The effects of peer instruction on students' conceptual learning and motivation

Tolga GOK

University of Dokuz Eylul, Torbali Technical Vocational School of Higher Education Izmir, TURKEY

E-mail: tolga.gok@deu.edu.tr; gok.tolga@gmail.com

Received 20 Oct., 2011
Revised 19 Mar., 2012

Contents

Abstract

This aim of this study was investigate the effects of peer instruction on college students' conceptual learning, motivation, and self-efficacy in an algebra-based introductory physics course for nonmajors. Variables were studied via a quasi-experiment, Solomon four-group design on 123 students. Treatment groups were taught by peer instruction. Control groups were taught by traditional didactic lecture method. To assess the effects of peer instruction, students were administered Force Concept Inventory and Motivated Strategies for Learning Questionnaire. Factorial analyses indicated that the treatment groups acquired significantly more conceptual learning, and were significantly more self-efficacious than students in the control groups. It was found that there were no significant differences in motivation between groups.

Keywords: Conceptual Learning; Motivation; Peer Instruction; Physics Education; Self-Efficacy

Introduction

Physics educators have realized that many students learn very little physics from traditional lectures. Several researchers have carefully documented college physics students' understanding of a variety of topics, and have concluded that traditionally taught courses do little to improve students' understanding of the central concepts of physics, even if the students successfully learn problem-solving algorithms (Crouch & Mazur, 2001; Crouch, Watkins, Fagen, & Mazur, 2007; Mazur, 1997). Simultaneously, researchers studying learning in higher education have established that students develop complex reasoning skills most effectively when actively engaged with the material they are studying, and have found that cooperative activities are an excellent way to engage students effectively. In response to these findings, many methods have been devised to improve student understanding of physics, ranging from modifications of traditionally taught courses to complete redesign of courses. One of the methods is peer instruction "PI" (Mazur, 1997).

PI is a student-centered approach to teaching. PI modifies the traditional lecture format to include questions designed to engage students and uncover difficulties with the material (Crouch & Mazur, 2001; Mazur, 1997; Porter, Bailey, Simon, Cutts, & Zingaro, 2011). PI provides a structured environment for students to voice their ideas and resolve misunderstandings by talking with their peers. By working together to learn new concepts and skills in a discipline, students create a more cooperative learning environment that emphasizes learning as a community in the classroom (Hoekstra, 2008; Kalman, Bolotin, & Antimirova, 2010; Turpen & Finkelstein, 2009). Research studies suggested that this type of cooperative learning environment could help promote deeper learning, as well as greater interest and motivation (Cross, 1998; Keiner & Burns, 2010; Simon & Cutts, 2012). Other research studies also showed that experts are able to monitor and regulate their own understanding (Bransford, Brown, & Cocking, 2000; Schoenfeld, 1992). When students were taught to apply peer instruction on how to learn effectively,

they engaged in a process called metacognition which was an ability to evaluate and monitor one's own cognitive process, so that a reasonable assessment could be done about future performance (Desoete, 2009; Shimamura, 2000). These metacognitive abilities enabled expert to employ different strategies to improve their learning. PI could also help students develop better metacognitive skills, as they checked their own understanding during pre-class reading and in-class questions. The method helped students when they did not understand a concept, when they were unable to answer a question on the reading, or when they could not give complete explanations to their peers during in-class discussion (Turpen & Finkelstein, 2010). With this formative, internal feedback, students could learn how to better assess their own understanding during the learning process.
PI encourages students to take responsibility for their own learning and emphasize understanding over simple task completion. Crouch & Mazur (2001) investigated the attitudes of students in the class performed of PI by the help of Maryland Physics Expectation Survey. The results revealed insignificant change in class attitudes over the semester. The convince-your-neighbor discussions systematically increased both the percentage of correct answers and the confidence of the students. Also the survey showed that student' satisfaction -an important indicator of student success- increased as well.

A great deal of research studies on cognition and learning indicated that students learned by using their existing knowledge, beliefs, and skills to create new knowledge (Bransford et al., 2000). Therefore, pedagogies in which instructors are made aware of students' incoming knowledge can enhance learning. PI provides opportunities for instructors and students to recognize background knowledge relevant to the pre-class reading, their initial vote, and discussion.

PI structures time during class around short, conceptual multiple-choice questions, known as ConcepTests. The best in-class ConcepTests often take advantage of the experiences and thinking students bring to the classroom about the material so that students can recognize their ideas and build on them. These questions are targeted to address student difficulties and promote student thinking about challenging concepts (Mazur, 1997).

The ConcepTests procedure is depicted in Fig. 1. After a brief presentation by the instructor, the focus shifts from instructor to student, as the instructor encourages students to think about the material by posing a ConcepTest. After 1-2 minutes of thinking, students commit to an individual answer. If too few students respond with the correct answer, the instructor may revisit the concept using lecture or try a different ConcepTest. If a large majority of students responded correctly, the instructor typically gives a brief explanation and moves on the next topic or ConcepTest. If an appropriate percentage of students answer ConcepTest correctly, the instructor asks students to turn to their neighbors and discuss their answers.

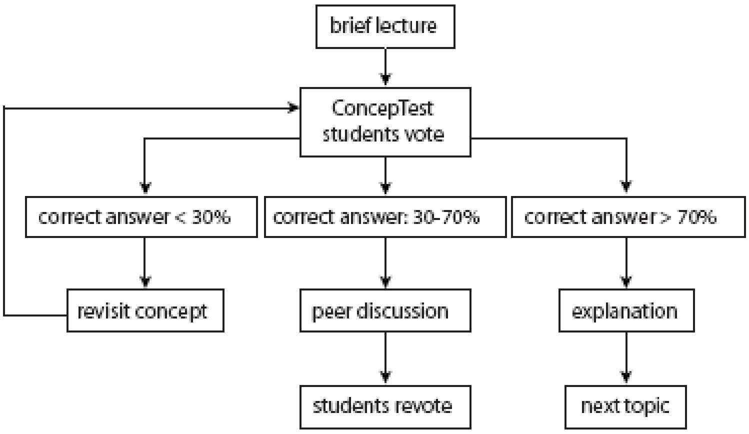

Figure 1. A peer instruction implementation procedure (Lasry et al., 2008)

Students talk in pairs or small groups and are encouraged to find someone with a different answer. The instructor circulates throughout the room to encourage productive discussions and guide student thinking. After several minutes, students answer the same ConcepTest again. The instructor then explains the correct answer and, depending on the student answers, may pose another related ConcepTest or move on to a different topic.

With the constant feedback from the reading assignments (the reading of a passage assigned by the instructor) and ConcepTests, the instructor can monitor student progress and help guide students to use their previously held ideas to understand new concepts and theories. Additionally, the flexibility of a PI lecture makes it easy for instructors to spend more time on concepts that are difficult for students by giving more focused, short presentations or asking more ConcepTests. In an interactive classroom, instructors are paying attention to student thinking throughout the learning process.

Rationale for the Present Study

Generally, research studies done concerning peer instruction have been focused on conceptual learning of the students so far (Crouch & Mazur, 2001; Crouch et al., 2007; Fagen, Crouch, & Mazur, 2002; Lasry, Mazur, & Watkins, 2008; Nitta, 2010). These studies performed have shown that peer instruction is effect on conceptual learning. But, any research studies on students' motivation and self-efficacy related to peer instruction have not met in the open literature as of 2011.

The main purpose of this study was to examine the effects of peer instruction on students' conceptual learning, motivation, and self-efficacy. The research questions investigated in this study were as follows:

1. Are there any effects of using peer instruction on students' conceptual learning?

2. Are there any effects of using peer instruction on students' motivational orientation?

Method

Experimental Design and Procedure

Students self-select into courses on the basis of personal choice, subjects could not be randomly assigned to treatment and control groups, nor could equal numbers of students be enrolled in each section. This limitation was addressed by using a "quasi-experimental design", as outlined by Campbell & Stanley (1963), Cook & Campbell (1979). Quasi-experimental designs assume that subjects cannot be randomly assigned to treatment or control groups, and thus, groups may be unequal as for as students' gender, majors, ability, background, etc. (the classes themselves were, however, randomly selected as treatment or control groups). The experimental design and analysis chosen, the Solomon four-group design (Campbell & Stanley, 1963), attempts to account statistically for any dissimilarities between treatment and control groups, but this is indeed a limitation of quasi-experimental designs.

The Solomon four-group design involves assignment of subjects to four groups. Two of the groups are pretested, and two groups are not. One of the pretested groups and one of the unpretested groups are subjected to the experimental treatment. The other two groups serve as controls. All four groups are then posttested (Campbell & Stanley, 1963) (Table 1).

Table 1 Solomon Four-Group Design

Pretest	Treatment	Posttest
O	X	O
O		O
	X	O
		O

X, treatment; O, dependent variables.

The Solomon four-group design offers rigorous control of most sources of internal and external validity and allows for increased generalizability vs. other experimental designs, because the four design elements are paralleled (Campbell & Stanley, 1963; Cook & Campbell, 1979). The paralleled elements control for the possible effects of a pretest on students' subsequent performance and determine both the main effects and interactions of testing. If the pretest cues the students, both pretest groups will have higher posttest scores than the groups that do not receive the pretest. If there is an interaction between the pretest and the experimental treatment, so that the pretest provides an advantage to those students who receive only the treatment, the pretest-treatment-posttest group will have higher posttest scores than the treatment-posttest group (Abraham & Cracolice, 1994; Campbell & Stanley, 1963; Cook & Campbell, 1979). So, this design allowed for the investigation of variables as well as interaction effects. All students were required to buy a student textbook, which serve as a template for basic course content, lecture notes, and review problems. The study was performed by the same instructor. All sections followed the same course outline and were taught similar paces.

The study was conducted in an algebra-based physics course (concerning Newtonian Concepts). The primary objective of the course was to have students be able to describe and explain the kinematics, first, second, and third law, superposition principle, and kinds of force.

This study was performed in the two-year college classroom in Turkey. A total of 132 students in algebra-based introductory physics were initially included when this study was initiated. However, due to an attrition rate that was not unexpected in this subject area, only 123 completed all requirements of the study. Because instruments were administered on different days at the beginning and end of the study due to scheduling and because some students were absent on the days the instruments were given, the numbers of students who took the pretest and posttests varied between sections. These numbers were statistically accounted for by the analytical procedures used.

The number of students who completed the study in the treatment groups included 62 students. In this group, 53.89% were male, 46.11% were female. A total of 61 students participated in the control groups. Like the treatment group, most of the students in this group were male (52.62%), with fewer female (47.38%). The course content was the same for all classes participating in the study. Other course structure variables, such as the syllabus, text, content, grading procedures, and exam structure/formats were held constant and did not deviate from previous course structure. The control groups were taught using the traditional, didactic lecture method of instruction (i.e., students listened as the instructor lectured on the content).

The treatment groups were taught using peer instruction mentioned in the "Introduction" section. The lecture emphasizing the concepts and the ideas behind the proof were implemented for 5–10 min while avoiding equations and derivations. After the short lecture period, the ConcepTest question was presented. The question was read to the students, making sure there was no misunderstanding about it. Then, they had 1 min to select an answer (more time allows them to fall back onto equations rather than think). Since it was important each student to answer individually, it was not allowed them to talk to one another. After about a minute, the students asked to record their answer and the voting was started. In this study, students used flashcards to answer the ConcepTest in the lectures instead of showing of hands. Flashcard was that each student had a set of six or more cards labeled A–F to signal the answer to a question. According to the number of the correct answer, the instructor decided whether start a discussion between students or not. If it was needed, students formed small groups by picking their friends where they actively discussed the answers for several minutes. Students tried to convince a neighbor of the rightness of that answer. The convince-your-neighbor discussions were always conducted with a few groups of students. Doing so allowed us to assess mistakes being made, to hear how students who have the right answer explain their reasoning, and also to minimize the effect of lack of instructional ability of students. Also the small group of students could join to the larger group if they were finished discussing or if they got stuck. After the discussion period, they were asked to record their revised answer. The instructor then reviewed the answer with the class as a whole. The entire process from beginning to end took approximately 10–15 min.

PI sections used three to four ConcepTests per 75-min class while the traditional section relied solely on lecturing. For a ConcepTest to be most effective, the question requiring higher-level thinking about a concept was selected so that students were not simply recalling something they read or using "trial and error" with equations. ConcepTests were also at an appropriate difficulty level so students were challenged but could reason to the answer with their existing knowledge.

Instruments

The data used in this study were collected with two ways. The first of them was Force Concept Inventory "FCI" (Hestenes, Wells, & Swackhammer, 1992) and the last was Motivated Strategies for Learning Questionnaire "MSLQ" (Pintrich, Smith, Garcia, & McKeachie, 1993). It was expected to find answers to all two research questions with the outcomes obtained from those statistical tools, respectively. The details of them were given as follows:

1. FCI (Hestenes et al., 1992) is a widely used tool to assess student's knowledge about topics in Newtonian Mechanics. FCI test consisting of 29 multiple-choice questions related to force and motion concepts, was used as a pre-and posttest. Internal reliabilities (Kuder-Richardson 21) for these tests were calculated as 0.69 and 0.71, respectively. FCI questions were divided to the 6 categories which were originally used by Hestenes et al. (1992) (Table 2; note that questions can appear in more than one category).
2. Motivated Strategies for Learning Questionnaire "MSLQ" (Pintrich et al., 1993) is a self-report instrument designed to assess college students' motivational beliefs and use of learning strategies. The instrument consists of motivation and learning strategies scales. The motivation scale proposes three general motivation constructs: value, expectancy, and affect. The motivation scale consists of thirty-one items. The reliability and validity of

the scale was performed in Turkish (Karadeniz, Buyukozturk, Akgun, Cakmak, & Demirel, 2008).

To determine the effectiveness of peer instruction on students' conceptual learning, motivation, and self-efficacy, the posttest scores (means) of the conceptual test and the means of MSLQ's categories were statistically analyzed for differences between the groups using 2x2 factorial design (Albanese & Mitchell, 1993).

Table 2 Categorization of the Question on FCI (Hestenes et al., 1992)

	Category	Questions
1. Kinematics	Velocity discriminated from position	20
	Acceleration discriminated from velocity	21
	Constant acceleration entails	
	parabolic orbit	23, 24
	changing speed	25
	Vector addition of velocities	7
2. First Law	with no force	4, 6, 10
	velocity direction constant	26
	speed constant	8, 27
	with cancelling forces	18, 28
3. Second Law	Impulsive force	6, 7
	Constant force implies constant acceleration	24, 25
4. Third Law	for impulsive forces	2, 11
	for continuous forces	13, 14
5. Superposition Principle	Vector sum	19
	Cancelling forces	9, 18, 28
6. Kinds of Force	Solid contact	
	passive	9, 12
	Impulsive	15
	Friction opposes motion	29
	Fluid contact	
	Air resistance	22
	buoyant (air pressure)	12
	Gravitation	5, 9, 12, 17, 18, 22
	acceleration independent of weight	1, 3
	parabolic trajectory	16, 23

Results and Discussion

The Results of FCI

Students' conceptual understanding of Newtonian Mechanics using FCI (Hestenes et al., 1992) in both treatment and control groups was measured. To measure how well students performed after instruction relative to their performance before instruction, the Hake normalized gains (Hake, 1998) for each student was calculated. These sets of individual normalized gains for treatment and control groups were then compared for statistical significance.

Table 3 shows FCI scores before instruction (pretest) and after instruction (posttest), as well as the normalized gains (g) for students in traditional and PI course. Differences between the traditional and PI courses were considered significant for p values less than 0.05. No significant difference existed between PI and control sections before instruction; however, PI section achieved significantly greater normalized gains after instruction ($p<0.01$). Also it was found that the student normalized gains this group was higher found as $g=0.59$.

Table 3 Precourse (Spre) and Postcourse (Spost) FCI Results for PI and Traditional Courses

	n	S_{pre}	S_{post}	g
Treatment				
Pretested	32	41.7	76.2	0.59
Not Pretested	30	-	75.3	-
Control				
Pretested	31	42.1	60.8	0.32
Not Pretested	30	-	59.7	-
Difference		-0.4	15.4	0.27[a]

[a]p<0.01

Also, this test was analyzed statistically to determine whether the experimental intervention had affected students' academic achievement as measured by the test. The 2x2 factorial analysis determined that the treatment groups performed significantly better on FCI than the control groups ($F=5.03$, $p=0.012$). No difference was found between those students who took FCI as a pretest and those who did not. This indicated that students who had taken FCI as a pretest had no significant advantage over those who had not (F=0.02; p=0.921). There was also no significant group by pretest interaction (F=0.34, p=0.543).

The Results of MSLQ

Factorial analyses were applied to determine whether the treatment had affected students' motivation and self-efficacy. Analyses were therefore directed at the motivation (value and affective components) and self-efficacy scales within the motivation section of MSLQ as a measure of students' motivation and self-efficacy. The means ± Standard Deviation "SD" for MSLQ scores are included in Table 4. Also the results of MSLQ categories were analyzed as follows:

Table 4 MSLQ Components

Source of Variation	*n*	Value Component	Affective Component	Self-Efficacy Component
Treatment				
Pretested	32	4.88±0.97	3.62±1.43	5.13±1.12
Not pretested	30	4.79±1.03	4.13±1.31	5.02±1.08
Control				
Pretested	31	4.73±0.79	3.81±1.31	4.61±1.09
Not pretested	30	4.61±1.01	3.95±1.38	4.45±1.15

Value Component Results

The value component of MSLQ motivation scale measures students' interest and goal orientation and the value of the course. Higher means indicate more interest, value, and positive goal orientation in the course and serve as a measure of students' motivation. The factorial analysis revealed no significant differences between the treatment and control groups with regard to students' scores on the value component of the motivation scale of MSLQ (F=1.26, p=0.382), even though the means were higher in the treatment groups (Table 5). No differences were noted between the pretested groups (F=0.00, p=0.847). Likewise, there was no significant pretest by treatment interaction (F=1.07, p=0.311).

Table 5 Value Component of MSLQ (n=123)

Source of Variation	*SS*	*df*	*MS*	*F*	*p*
Main effects Within (error)	124.56	119	1.04		
Treatment vs. control group	1.32	1	1.32	1.26	0.382
Pretest	0.00	1	0.00	0.00	0.847
Group by pretest	1.12	1	1.12	1.07	0.311
Between (model)	2.44	3	0.81	0.77	0.591
Total	127	122	1.04		

SS: sum of squares; df: degrees of freedom; MS: mean squares.

Affective Component Results

The affective component of MSLQ motivation scale measures how much students worry about tests and how often they have distracting thoughts when they take an exam. Higher means indicate more anxiety in testing situations, thus measuring the affective component of motivation. The factorial analysis revealed no significant differences between the treatment and control groups with regard to students' scores on the affective component of the motivation scale of MSLQ (F=0.18, p=0.664; Table 6). No differences were noted between the pretested groups (F=1.97, p=0.152), and there was no significant pretest by treatment interaction (F=0.70, p=0.433).

Table 6 Affective Component of MSLQ (n=123)

Source of Variation	*SS*	*df*	*MS*	*F*	*p*
Main effects Within (error)	235.81	119	1.98		
Treatment vs. control group	0.36	1	0.36	0.18	0.664
Pretest	3.92	1	3.92	1.97	0.152
Group by pretest	1.39	1	1.39	0.70	0.433
Between (model)	5.67	3	1.89	0.95	0.445
Total	241.48	122	1.97		

Self-efficacy Component Results

The self-efficacy component of MSLQ measures students' expectancy of success, their perceptions of self-confidence in understanding the course content, and their control over those beliefs. Higher means indicate the better students believe they will do in the course and be able to master the course material. The factorial analysis revealed a significant difference between the treatment groups and the control groups with regard to their self-efficacy on the self-efficacy section of MSLQ (F=4.37, p=0.041; Table 7). No significant differences were found between those who were pretested and those who were not (F=0.37, p=0.563), and there was no significant pretest by treatment interaction (F=0.02, p=0.879).

Table 7 Self-Efficacy Component of MSLQ (n=123)

Source of Variation	*SS*	*df*	*MS*	*F*	*p*
Main effects Within (error)	152.18	119	1.27		
Treatment vs. control group	5.55	1	5.55	4.37	0.041
Pretest	0.47	1	0.47	0.37	0.563
Group by pretest	0.03	1	0.03	0.02	0.879
Between (model)	6.05	3	2.01	1.58	0.214
Total	158.23	122	1.29		

Conclusion

When the results were evaluated in terms of the students' conceptual learning, peer instruction in the two-year college classroom was found to be more effective at developing students' conceptual understanding than traditional didactic lecture method. The use of PI in an algebra-based introductory physics course did not have an impact on the students' motivation (value and affective component of MSLQ). Students may have failed to see the connection or relevance of the course to their own particular academic major. FCI designed to connect the content to students' own majors were not introduced until the end of the study by simple inquiry, their duration may not have been enough to make a change in the value of students' perception toward the course and thus their motivation. The interest, meaning, and relevance are measures of intrinsic motivation (Pintrich & Garcia, 1995), PI used in this manner may thus have failed to generate the necessary interest and value of perception required to promote motivation change.

For the affective component of the motivation scale of MSLQ, no significant change was detected between the treatment and control groups. To explain these results, it may be relevant to consider the behavior of the students enrolled in the course and the links between motivation and achievement. PI, although resulting in greater achievement gains in students, may not have been able to address students' worry and concern about overtaking exam. Paulsen and Feldman (1999) determined that students who have naive beliefs about learning and knowledge were more likely to be less motivated and have higher levels of test anxiety (Allen, Duch, & Groh, 1996) than were students with more sophisticated beliefs. Non-major students often have little science background and come from various majors. They may have viewed the content from a less sophisticated view because of their lack of science background; however, more work needs to be done in this area.

Besides, even though PI did not significantly benefit students' motivation, its use did improve their achievement and self-efficacy. These findings imply that PI should be directed more at making the content more relevant and meaningful in the future, possibly with more duration and frequency. PI resulted in an improvement of students' self-efficacy, but what exactly is self-efficacy and how does it related to PI? Self-efficacy is characterized by one's beliefs about behavioral outcomes, coupled with expectations about one's ability to engage in, perform, persist in, and be successful at a particular behavior (Allen et al., 1996; Bandura, 1977; Hemenway, Straits, Wilke, & Hufnagel, 2001), in this case science. Because many non-majors students come to the courses with negative attitudes and low-efficacy, it is imperative to nurture feelings of confidence from the beginning.

The use of PI in present study was reported to significantly increase or change students' science self-efficacy by promoting a belief in their own to do science and be successful in learning about it. PI used were indeed designed to help students gain confidence in and had control over their abilities to learn physics and be successful in doing so, but PI was also introduced gradually over the course of the study period based on a ConcepTest continuum. This was to enable students to see the results of their efforts and receive feedback in a relatively risk-free environment and thus help them develop their self-efficacy over the course of the entire study period. Because students were active participants in the learning process, self-efficacy was improved compared with those students who experienced traditional didactic lectures. What these findings suggest is that self-efficacy and classroom success are linked and that an individual's level of engagement in a task and willingness to persist at the task are indicators of success (Paulsen & Feldman, 1999; Pintrich, Marx, & Boyle, 1993a).

References

Abraham, M. R., & Cracolice, M. S. (1994). Doing research on college science instruction. *Journal of College Science Teaching*, *23*, 150-153.

Albanese, M. A., & Mitchell, S. (1993). Problem-based learning: A review of literature on its outcomes and implementation issues. *Journal of the Association of Medical Colleges, 68*, 52-68.

Allen, D. E., Duch, B. J., & Groh, S. E. (1996). The power of problem based learning in teaching introductory science courses. *New Directions Teaching Learning*, *68*, 43-51.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, *84*, 191-215.

Bransford, J. D., Brown, A. L., & Cocking, R. R. (2000). *How people learn: Brain, mind, experience, and schooling*. Washington, DC: National Academies Press.

Campbell, D. T., & Stanley, J. C. (1963). Experimental and quasi experimental designs for research: *A handbook for research on interactions*. Boston, MA: Houghton Mifflin.

Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Chicago, IL: Rand- McNally College Publishing.

Cross, K. P. (1998). Why learning communities? Why now? *About Campus*, *3*(3), 4-11.

Crouch, C. H., & Mazur, E. (2001). Peer instruction: Ten years of experience and results. *American Journal of Physics*, *69*, 970-977.

Crouch, C. H., Watkins, J., Fagen, A. P., & Mazur, E. (2007). Peer instruction: Engaging students one-on-one, all at once. (In E. F. Redish, & P. Cooney (Eds.), American Association of Physics Teachers, College Park: MD).

Desoete, A. (2009). Mathematics and metacognition in adolescents and adults with learning disabilities. *International Electronic Journal of Elementary Education*, *2*(1), 82-100.

Fagen, A. P., Crouch, C. H., & Mazur, E. (2002). Peer instruction: Results from a range of classrooms. *The Physics Teacher*, *40*, 206-209.

Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics*, *66*, 64-74.

Hestenes, D., Wells, M., & Swackhammer, G. (1992). Force concept inventory. *Physics Teacher*, *30*, 141-151.

Hemenway, M. K., Straits, W. J., Wilke, R. R., & Hufnagel, B. (2001). Educational research in an introductory astronomy course.*Innovative Higher Education*, *26*, 178-269.

Hoekstra, A. (2008). Vibrant student voices: Exploring effects of the use of clickers in large college courses. *Learning, Media, & Technology*, *33*(4), 329-341.
Kalman, C. S., Bolotin, M., & Antimirova, T. (2010). Comparison of the effectiveness

of collaborative groups and peer instruction in a large introductory physics course for science majors. *Canadian Journal of Physics*, *88*(5), 325-332.

Karadeniz, S., Buyukozturk, S., Akgun, O. E., Cakmak, E, K., & Demirel, F., (2008). The Turkish adaptation study of motivated strategies for learning questionnaire (MSLQ) for 12-18 year old children: Results of confirmatory factor analysis. *The Turkish Online Journal of Educational Technology*, *7*(4), 108-117.

Keiner, L. E., & Burns, T. E. (2010). Interactive engagement: How much is enough? *The Physics Teacher*, *48*(2), 108-111.

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

Mazur, E. (1997). *Peer instruction: A user's manual*. Upper Saddle River. NJ: Prentice Hall.

Nitta, H. (2010). Mathematical theory of peer instruction dynamics. *Physical Review Special Topics- Physics Education Research*, *6*(2), 020105, 1-4.

Paulsen, M. B., & Feldman, K. A. (1999). Student motivation and epistemological beliefs. *New Directions Teaching Learning*, *78*, 17-25.

Pintrich, P. R., Smith, D. A., Garcia, T., & McKeachie, W. J. (1993). Reliability and predictive validity of the motivated strategies for learning questionnaire (MSLQ). *Educational and Psychological Measurement*, *53*, 801-813.

Pintrich, P. R., Marx, R. W., & Boyle, R. A. (1993a). Beyond cold conceptual change: The role of motivational beliefs and classroom contextual factors in the process of conceptual change. *Review of Educational Research*, *63*, 167-199.

Pintrich, P. R., & Garcia, T. (1995). *Assessing students' motivation and learning strategies*. (Meeting American Educational Research Association, San Francisco, CA).

Porter, L., Bailey, L. C., Simon, B., Cutts, Q., & Zingaro, D. (2011). A multi-classroom report on the value of peer instruction. *In proceedings of the 16th Annual Joint Conference on Innovation and Technology in Computer Science Education*, June 27-29, Darmstadt, Germany.

Schoenfeld, A. H. (1992). *Learning to think mathematically: Problem solving, metacognition, and sense making in mathematics Handbook of research on mathematics teaching and learning: A project of the National Council of Teachers of Mathematics (pp 334-370).* New York: Macmillan.

Shimamura, A. P. (2000). What is metacognition? The brain knows. *The American Journal of Psychology*, *113*, 142-147.

Simon, B., & Cutts, Q. (2012). Peer instruction: A teaching method to foster deep understanding. *Communications of the ACM*, *55*(2), 27-29.

Turpen, C., & Finkelstein, N. D. (2009). Not all interactive engagement is the same: Variations in physics professors' implementation of peer instruction. *Physical Review Special Topics- Physics Education Research*, *5*(2), 020101, 1-18.

Turpen, C., & Finkelstein, N. D. (2010). The construction of different classroom norms during peer instruction: Students perceive differences. *Physical Review Special Topics - Physics Education Research*, *6*(2), 020123, 1-22.