JOURNAL OF BALTIC SCIENCE EDUCATION

ISSN 1648-3898

PEER INSTRUCTION IN THE PHYSICS CLASSROOM: EFFECTS ON GENDER DIFFERENCE PERFORMANCE, CONCEPTUAL LEARNING, AND PROBLEM SOLVING

Tolga Gok

Abstract. *The purpose of this study was to examine the effects of peer instruction (PI) on performance, gender gap, conceptual learning and problem solving in physics. The students enrolled in two sections of a physics course were selected; one section was treatment group (TG) and the other section was control group (CG). 42 students in TG were instructed with PI whereas 56 students in CG were instructed with conventional instruction. Data were collected using Force Concept Inventory, final examination problems, and students' opinions about the instruction. The results indicated that PI had more positive effect on students' conceptual learning and problem solving than conventional instruction. The conceptual learning with PI did not show any statistically significant difference between female and male students in TG. However, the quantitative problem solving performance after PI revealed that male students performed better than female students. The female and male students' conceptual learning and problem solving performance did not change noticeably in CG. The students in TG also changed their perspective on solving a problem and understanding a concept. They found the method helpful to connect the quantitative solution with related concepts.*

Key words: *conceptual learning, gender gap, peer instruction, physics education, problem solving.*

Tolga Gok
Dokuz Eylul University, Izmir, Turkey

Introduction

Conceptual learning and problem solving are critical in physics education. Students' ability to solve problems and to learn fundamental concepts are an important dimension of learning physics (Fraser et al., 2014). Several studies demonstrated that many students complete physics courses retaining the same misconceptions and preconceptions as when they entered the course (McDermott, 2001; Van Heuvelen, 1991). Also, they have great difficulty in solving physics problems and defining concepts after conventional instruction (Mazur, 1997; Meltzer & Manivannan, 2002; Reif, 1995). Many physics education researchers (PER) have found that conventional instruction is not appropriate in facilitating learners in physics courses to comprehend concepts, to solve problems, and to analyze and evaluate their solutions of problems (Crouch & Mazur, 2001; Madsen, McKagan & Sayre, 2013; Puente & Swagten, 2012).

Therefore, several alternative instructional approaches to enhance students' understanding of physics concepts have been implemented to overcome the deficiency of conventional instruction and to generate a better foundation for further physics understanding (Hake, 1998; McDermott, Shaffer, & Rosenquist, 1995; Redish, 2003). McDermott (2001) has reported several inadequacies of conventional instruction. These inadequacies include enabling students to make connections among fundamental concepts, use of formal representations (graphical, diagrammatic, algebraic, etc.), inadequate reference to events in the real world, inability to overcome certain conceptual difficulties (e.g., the diffraction and the interference of light, wave or particle model, etc.), insufficient comprehension of the procedure of learning how to solve qualitative and quantitative physics problems, and improving the scientific reasoning skills necessary to implement the fundamental principles or concepts in basic situations. He suggested some alternative instructional approaches (e.g., *tutorials in introductory physics, physics by inquiry)* in place of conventional instruction in his landmark study.

"Peer Instruction" as an Alternative Approach to Postsecondary Physics Instruction

One alternative educational approach is peer instruction (PI). Mazur and Watkins (2010) have reported PI as "an interactive teaching technique that promotes classroom interaction to engage students and address difficult aspects of the material" (p. 39).

PI uses multiple-choice test items for difficult-to-learn physics concepts to enhance the teaching environment in large physics class and to engage students of non-physics majors. Students represent their responses to test by holding up different coloured indicators (Mazur, 1997). The collected responses are summarized and presented quickly to the whole class. Peer student groups (2-3 student/group) are required to discuss their responses, opine on the differences, and further a shared response. They were asked one more time to express their responses. The rest of the lecture is structured based on this set of responses and any changes from the first expression (Gok, 2014). The peer-to-peer interaction deepens the students` comprehension of fundamental concepts or principles, enables discussing the problem with their peers in class discussions (Crouch & Mazur, 2001).

The studies performed on PI have reported that peer instruction was effective in facilitating conceptual learning (Crouch & Mazur, 2001; Crouch, Watkins, Fagen, & Mazur, 2007; Lasry, Mazur, & Watkins, 2008; Turpen & Finkelstein, 2009; Mazur &Watkins, 2010) and problem solving (Gok, 2012; Gok, 2014; Puente & Swagten, 2012).

A limited number of studies on gender differences related to peer instruction have been found in the extant literature as of 2014. Lorenzo, Crouch and Mazur (2006) only investigated a physics course by using interactive engagement (IE) methods. They showed that teaching with certain interactive strategies not only provided significant increase in comprehension of both genders, but also reduced the gender difference. Therefore, the effects of the IE methods (cooperative group problem solving, interactive lecture demonstrations, just in time teaching, peer instruction, real-time physics, think-pair-share, tutorials in introductory physics, workshop physics, etc.) on gender differences need to be examined in physics education by PER.

Fraser et al. (2014) modified Hake's (1998) definition of the IE methods. They stated that "interactive engagement methods promote conceptual understanding through interactive engagement of students in heads-on (always) and hands-on (usually) activities which yield immediate individual feedback to all students through discussion with peers and/or instructors" (p. 2). There are many studies (Deslauriers, Schelew, & Wieman, 2011; Hoellwarth, Moelter, & Knight, 2005; Meltzer & Manivannan, 2002; Sayre et al., 2012; Watkins & Mazur, 2013) that discuss IE methods. These studies demonstrate that the IE methods have resulted in higher students' average scores than the average scores of conventional lecture strategies in physics courses. Also female students were reported to have benefited more from the IE methods than male students (Baker & Leary, 2003; Heller & Hollabaugh, 1992; Laws, Rosborough, & Poodry, 1999). The studies performed on the IE methods were generally focused on conceptual understanding gains of the students instead of gender differences in both the short and long term. Therefore, these methods should be examined and discussed in terms of gender gap.

Gender Differences in Physics Achievement Related to Newtonian Mechanics

Many studies performed on gender differences in physics education have reported that male students generally outperformed female students on understanding the motion and force concept tests (Coletta, Phillips & Steinert, 2011; Donnelly, MacPhee & Bates, 2012; Kost-Smith, Pollock & Finkelstein, 2010; Madsen et al., 2013; Richardson & O'Shea, 2013). Also, the performance of male students was indicated to be higher than that of female students on quantitative problems and examinations (Kost-Smith et al., 2010; Kost-Smith, Pollock & Finkelstein, 2009). A few studies have shown that female students performed better than male students on homework and the other course components (Donnelly et al., 2012; Kost-Smith et al., 2009; Tai & Sadler, 2001). The extant literature indicates that the use of multiple representations (e.g., graphs, diagrams, and vectors) and conceptual understanding related to force and motion concept tests were not affected by gender differences (Bates et al., 2013; Nieminen, Savineinen & Viiri, 2013).

It has also been reported that most female students had less confidence in their own abilities to do science courses than most male students. Many female students were found to be highly anxious about succeeding in physics courses (Ding & Harskamp, 2006; Hazari, Sadler, & Tai, 2008; Udo et al., 2001).

The gender gap between female and male students may be decreased by the IE methods in physics courses (Docktor & Heller, 2008). The effects of the IE methods on gender gap were examined at both Harvard and the University of Colorado (Lorenzo et al., 2006; Pollock, Finkelstein, & Kost-Smith, 2007). These methods eliminated the gender gap in performance scores on the diagnostic tests (Force Concept Inventory "FCI", Force and Motion

ISSN 1648–3898

Conceptual Evaluation "FCME"). They claimed and argued that the elimination of the gender difference was not dependent on the instructor, but it was dependent on the pedagogical strategies that were used in the classroom, and other contextual factors that were covered such as, student demographics, epistemological beliefs, physics self efficacy, etc.

The purpose of the present research was to examine the gender differences in conceptual learning and problem solving by using different teaching methods (peer instruction in the treatment group and conventional instruction in the control group). The performance of the students was evaluated using Force Concept Inventory and final examination problems. The students' opinions were also collected with written feedback about the teaching method. The research questions investigated were:

1. Are there any differences between female and male students' conceptual learning in the treatment group and the control group?
2. Are there any differences between female and male students' quantitative problem solving skills in the treatment group and the control group?
3. Does the teaching of peer instruction change female and male students' opinions about conceptual learning and problem solving?

Methodology of Research

A 2 (method: peer instruction and conventional instruction) x 2 (gender: female and male) factorial design with a control group was used in this research. Intact sections of a college physics course were assigned to treatment and non-treatment conditions. In this section, participants, instructional approach, data collection, and data analysis were given as follows.

Participants

The research was performed in Torbali Technical Vocational School of Higher Education at Dokuz Eylul University, Turkey. There are four departments (Industrial Glass and Ceramics, Geotechnic, Drilling Technology, Natural Building Stone Technology) in the college offering two-year programs. The study sample consisted of 98 first year college students from two different sections of a physics course where the sections were randomly assigned to two groups. The students were between 18 and 20 years of age. About 50% of the students from this college continue their education at a four-year university. The treatment group (TG) consisted of 42 students composed of 48% female and 52% male. The treatment group was instructed using peer instruction. The other section of 56 students with similar distribution of females (39%) and males (61%) served as the control group (CG). The control group was instructed by means of conventional instruction. The interaction between the students in treatment group and the control group was not allowed. Another task relating with the course except FCI and examinations was not assigned to the students.

Instructional Approach

This study was performed with two groups. The TG students were instructed using peer instruction, and the CG students were instructed by means of conventional instruction. Both groups were taught by the same instructor for five weeks. The primary objective of the course was to accustom students to describing and explaining the principles of kinematics, Newton's Laws, and applications of Newton's Laws by in-class discussions. There was no laboratory associated with this course. Concept tests were used in the teaching of these topics.

All of the concept tests were designed by the instructor to require students a) engagement in known and unknown concepts, (b) exploration of known and unknown concepts, and (c) explanation of known and unknown concepts. The concept tests were formatted as multiple-choice questions with five possible responses. Six or eight concept tests were administered in a 75-min class. The procedures of instruction of the treatment and control groups were explained in detail.

The procedure of peer instruction in the TG was as follows:

- The instructor gave several short presentations on key points in each course instead of presenting the details covered in the textbook.

- The instructor presented one or two concept test(s) - short conceptual question on the subject being discussed - after each short presentation.
- The students were given time to formulate individual answers. The instructor did not allow the students to talk to one another.
- They reported their individual answers.
- During the voting process, coloured flashcards were used to report the students' answers. The students used various colours to indicate their responses: red for A, yellow for B, green for C, blue for D, and white for E.
- They were asked to discuss their answers with peers.
- The peer discussion process is shown in Figure 1. When the number of correct answers reached 30-70% of the responses, the instructor started the whole-class discussion. The concept test was re-examined if the correct answers were below the threshold value of 30%.

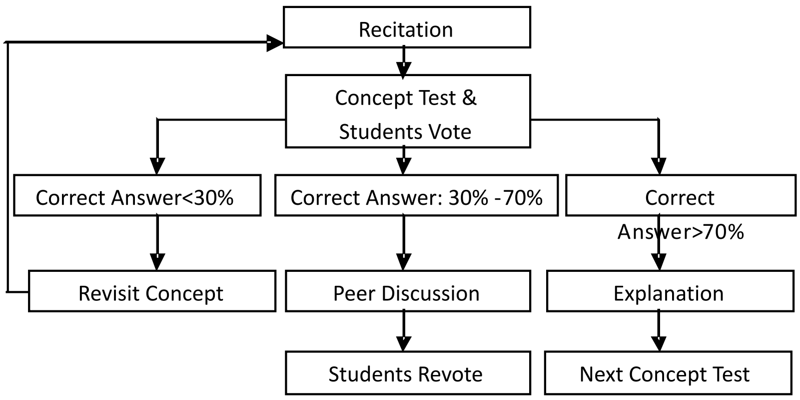


Figure 1: The procedure of peer instruction (Lasry et al., 2008).

- The students revised their answers.
- The instructor provided general feedback concerning their correct answers.

Identical concept tests were administered to the CG with the help of conventional instruction. The instructor gave the presentation related to the concepts and principles and then the instructor asked the concept tests to the students. The students also answered the concept tests by using the flashcards. Finally, the instructor evaluated and discussed their answers in the classroom environment. Also identical quantitative problems were solved both by the TG and the CG according to the performed teaching method. The instructor followed stepwise problem solving strategies (physical representation, concept representation, solution representation, and controlling representation) by problem solving in the courses (See Appendix for a sample problem).

Data Collection

The data of this study were quantitatively collected using Force Conceptual Inventory (Hestenes, Wells, & Swackhamer, 1992) and final examination problems. Also, the opinions of the students in the TG were qualitatively collected from their writing of an anonymous essay on the teaching method.

Measure of Conceptual Understanding of the Students using Force Concept Inventory (FCI)

The FCI was developed in 1992 (Hestenes et al., 1992). The FCI has been commonly used worldwide as a diagnostic tool to measure students' understanding of Newtonian concepts of force and highlights common misconceptions (Donnelly et al., 2012). The FCI, consisting of 30 items, was revised in 1995 by Halloun, Hake, Mosca, and Hestenes (Mazur, 1997). The English version of the FCI was translated into Turkish in this research. The researcher

ISSN 1648–3898

examined the validity and reliability of the FCI using classical test theory. The face and construct validity were established by eight physics professors in Turkey. The clarity of the items was pilot-tested on 95 randomly selected second year students not involved in the present study. These students' responses were submitted to statistical analyses to establish validity and reliability.

The pilot test data were analyzed. The Bartlett's test of sphericity ($p<0.01$) and the Kaiser-Meyer-Olkin value (KMO>0.60) were examined from the principal component analysis. Five factors were extracted by Exploratory Factor Analysis (EFA) with eigenvalues greater than 1.00. The items with factor loading below 0.30 were deleted, which led to the exclusion of six items from the five dimensions. Five sub-factors were retained and accounted for 76% of the total variance. Item loading for the 24 selected items ranged from 0.70 to 0.32. These dimensions and items of the FCI showed similarity to the research of the Scott, Schumayer, and Gray (2012). The first factor (5 items) was "Identification of Forces (IF)", the second factor (5 items) was "Newton's First Law with Zero Force (NFLZF)", the third factor (6 items) was "Newton's Second Law and Kinematics (NSLK)", the fourth factor (3 items) was "Newton's First Law with Cancelling Force (NFLCF)", and the last factor (3 items) was "Newton's Third Law (NTL)".

Measure of Problem Solving Skills of the Students with Final Examination Problems

The problem solving skills of the female and male students in the TG and the CG were evaluated with five problems on the final examination (see Appendix for sample of a final examination problem). The five final examination problems were distributed among each of the sub-factors of the FCI as follows: Identification of Forces, Newton's First Law with Zero Force, Newton's Second Law and Kinematics, Newton's First Law with Cancelling Force, and Newton's Third Law. Final examination problems were selected from the textbook (Cummings, Laws, Redish, & Cooney, 2004) which was used in the course.

Determination of Students' Opinions with Written Essay

Female and male students' opinions in the TG were documented by writing an anonymous essay on peer instruction. Volunteer students (n=28) were asked to determine positive and negative aspects of the applied teaching method at the end of the instruction with the following question: "What do you think about the teaching of peer instruction? Please state positive and negative aspects of the instruction method." No interactions amongst the students were allowed during the writing session, which took about 15 minutes to complete.

Data Analysis

The TG and CG students' responses to the FCI pre-test and post-test were analyzed using IBM SPSS Statistics version 22. Descriptive statistics, fractional gain, analysis of variance (ANOVA), multivariate analysis of variance (MANOVA), and effects sizes were performed. Descriptive statistics included means (M) and standard deviations (SD) for the pretests and posttests. The fractional gains (*g*) were calculated to avoid the limitations in normal gain scores (Hake, 1998, see formula below). Hake suggested specific ranges for various fractional gains: "*high-g*" ($g \geq 0.7$), "*medium-g*" ($0.7 > g \geq 0.3$) and "*low-g*" ($g < 0.3$).

$$\text{Fractional Gain (FG)} < g >= \frac{(\text{posttest\%} - \text{pretest\%})}{(100\% - \text{pretest\%})}$$

An ANOVA test was conducted to test the treatment main effect on the posttest means of the TG and the CG. MANOVA was used for comparing multivariate means of sub-factors. Also, effect sizes were calculated via Partial Eta Squared-PES-($\eta_p^2(\eta_p^2$). PES effects sizes were considered to be "*small*" for $0.01 \leq \eta_p^2 < 0.06$, "*medium*" for $0.06 \leq \eta_p^2 < 0.14$, and "*large*" for $\eta_p^2 \geq 0.14$ as suggested by Kinnear and Gray (2008).

The problem solving performances of the TG and the CG students were evaluated according to the answer key by the instructor and compared with the scores obtained from the final examination that covered five quantitative problems. The students' final examination problems were evaluated on a four-point scale: physical representation of the problem received 1 point, concept representation of the problem received 1 point, solution representation qualitatively and quantitatively received 1 point and controlling representation received 1 point.

The opinions of the voluntary students (28 of 42 students) in the TG were read, coded and categorized by the researcher as being positive and negative. Most of the students (90%) had positive opinions, while the other 10% of the students had negative opinions. The positive opinions of the students focused generally on the verbs "*comprehend, learn, improve, solve, focus, examine, and apply*". The negative opinions of the students were combined with the adjectives "*time-consuming, boring, and nervous*".

Results of Research

The descriptive statistics (Table 1) illustrated that female and male students in the groups obtained similar performance scores at the beginning of the study. Appropriate parametric tests (independent samples *t*-test and 2x2 ANOVA) were used to detect any significant differences between the TG and the CG on the pretest scores. Analyze of the *t*-test revealed no statistically significant differences in the FCI $[t_{df=96} = 0.07, p > 0.05]$. The fractional gains (*g*) of the TG (0.52) and the CG (0.13) were found as "*medium*" and "*low*" respectively. The factorial gains showed that the TG had a larger gain than the CG.

Table 1. Descriptive statistics of the two groups' results on the FCI.

Gender	Group	Pretest			Posttest		Fractional Gain
		N	M	SD	M	SD	*g*
Female	TG	20	4.35	2.51	14.50	2.37	0.52
	CG	22	4.04	1.93	6.77	1.50	0.14
Male	TG	22	3.63	2.08	14.45	3.24	0.53
	CG	34	3.88	1.87	6.32	2.08	0.12
Total	TG	42	3.97	2.30	14.47	2.83	0.52
	CG	56	3.94	1.88	6.50	1.87	0.13

The repeated measures ANOVA was used to determine any significant differences between the TG and the CG mean scores on the posttest for the FCI. The results from 2x2 (method x gender) repeated measures ANOVA confirmed performance differences the TG and the CG. The main effects for gender [$F_{(1,94)}$=0.26; *p*>0.611; $\eta_p{}^2 = 0.003$] and the interaction (gender x method) [$F_{(1,94)}$=0.17; *p*>0.678; $\eta_p{}^2 = 0.002$] between the TG and the CG were calculated. ANOVAs indicated that the results were not statistically significant. The main treatment (instruction method) effect between the TG and the CG [$F_{(1,94)}$=267.03; *p*<0.001; $\eta_p{}^2 = 0.740$] was found to be statistically significantly different in favor of the TG. The PES values for gender and the interaction (gender x method) were found to be "*small*", whereas the PES value for the treatment was calculated to be "*large*".

Figure 2 represents the interaction between the instruction method and gender. The difference between female and male students' mean scores before the instruction was found as 0.72 for the TG. The result of the *t*-test on female and male students' mean scores obtained from the FCI revealed no statistically significant difference $[t_{df=40} = 1.00, p > 0.05]$. After the instruction, the difference between female and male students' mean scores was calculated as 0.05 and the difference in the *t*-test result of FCI was not statistically significant $[t_{df=40} = 0.51, p > 0.05]$.

For the control group, the difference in the FCI between female and male students' mean scores before the instruction (0.16) and after the instruction (0.45) was calculated as 0.29. The differences of the *t*-test results before the instruction $[t_{df=54} = 0.31, p > 0.05]$ and after the instruction $[t_{df=54} = 0.87, p > 0.05]$ were not statistically significant.

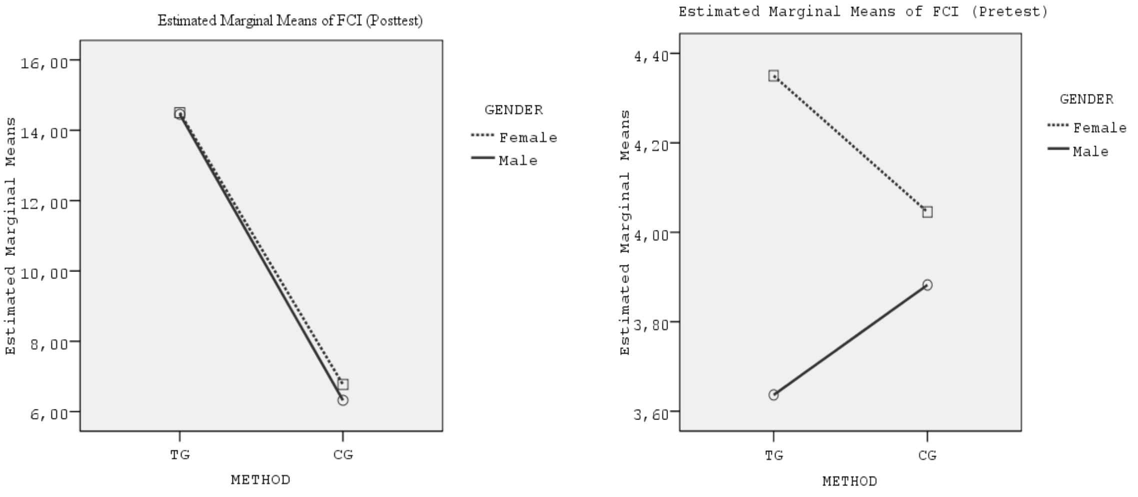


Figure 2: The difference between pre-test and post-test mean scores of female and male students in the TG and the CG.

Table 2 shows the descriptive statistics of the FCI sub-factor scores for the pre-test and post-test of the TG and the CG. When the results of the groups' pre-test and post-test scores were compared, the pretest scores of both groups for the sub-factors appeared to be similar while the posttest scores of the TG for the sub-factors (IF, NFLZF, NSLK, NFLCF, and NTL) were found higher than the posttest scores of the CG for the sub-factors.

Table 2. Descriptive statistics of the two groups' and female and male students' results on the FCI' sub-factor scores.

		IF				NFLZF			
		Pretest		Posttest		Pretest		Posttest	
Gender	Group	M	SD	M	SD	M	SD	M	SD
Female	TG	0.55	0.51	2.80	1.19	1.45	0.51	5.05	1.19
	CG	0.31	0.47	1.00	0.69	1.09	1.10	2.45	1.18
Male	TG	0.36	0.49	2.86	1.24	1.27	0.70	4.40	2.30
	CG	0.52	0.74	0.94	0.73	1.50	0.96	2.00	1.25
		NFLCF				**NTL**			
		Pretest		**Posttest**		**Pretest**		**Posttest**	
Gender	**Group**	**M**	**SD**	**M**	**SD**	**M**	**SD**	**M**	**SD**
Female	TG	0.40	0.82	1.60	0.75	0.55	0.51	1.75	0.78
	CG	0.81	0.66	0.90	0.52	0.59	0.50	0.63	0.78
Male	TG	0.54	0.73	1.68	0.71	0.40	0.73	1.81	0.58
	CG	0.50	0.78	0.70	0.52	0.58	0.55	0.73	0.44

		NSLK			
		Pretest		Posttest	
Gender	Group	M	SD	M	SD
Female	TG	1.40	1.27	3.30	1.08
	CG	1.22	0.97	1.77	0.68
Male	TG	1.04	1.09	3.68	0.64
	CG	0.76	0.95	1.94	0.88

Multivariate analysis of variance (MANOVA) was performed to test gender, method, and interaction effects on the scores of the sub-factors (IF, NFLZF, NSLK, NFLCF, and NTL) between the TG and the CG (Table 3). The pretest results revealed no statistically significant difference for gender, method, and interaction (p>0.001).

Table 3. The results of repeated measures MANOVA for differences between the TG and the CG for the FCI' sub-factors.

		Before Instruction			After Instruction		
Source	Sub-Factors	*F*	*p*	$\eta_p{}^2$	*F*	*p*	$\eta_p{}^2$
Gender	IF	0.01	0.920	0.000	0.00	0.990	0.000
	NFLZF	0.41	0.521	0.004	3.01	0.086	0.031
	NSLK	3.48	0.065	0.036	2.50	0.117	0.026
	NFLCF	0.31	0.575	0.003	0.22	0.638	0.002
	NTL	0.35	0.551	0.004	0.39	0.530	0.004
Method	IF	0.07	0.789	0.001	87.2	0.000	0.481
	NFLZF	0.13	0.715	0.001	62.9	0.000	0.401
	NSLK	1.07	0.303	0.011	88.3	0.000	0.484
	NFLCF	1.47	0.228	0.015	42.0	0.000	0.309
	NTL	0.84	0.362	0.009	68.7	0.000	0.422
Gender x Method	IF	2.62	0.109	0.027	0.09	0.759	0.001
	NFLZF	2.64	0.107	0.027	0.08	0.768	0.001
	NSLK	0.06	0.806	0.001	0.37	0.541	0.004
	NFLCF	2.28	0.134	0.024	1.23	0.270	0.013
	NTL	0.33	0.566	0.004	0.01	0.908	0.000

Quantitative Problem Solving

Problem solving skills of female and male students were measured with final examination (five) problems in the TG and the CG. The achievement percentages of female and male students are summarized in Table 4.

Table 4. The achievement percentage of female and male students according to groups and problems.

Group		Problem I	Problem II	Problem III	Problem IV	Problem V
TG	F	11/20 (55.00%)	12/20 (60.00%)	13/20 (65.00%)	13/20 (65.00%)	14/20 (70.00%)
	M	15/22 (68.18%)	16/22 (72.73%)	17/22 (77.27%)	18/22 (81.82%)	18/22 (81.82%)
CG	F	6/22 (27.27%)	7/22 (31.82%)	6/22 (27.27%)	7/22 (31.82%)	6/22 (27.27%)
	M	9/34 (26.47%)	11/34 (32.35%)	10/34 (29.41%)	10/34 (29.41%)	9/34 (26.47%)

Male students were more successful in quantitative problem solving than female students in the TG. The male students' mean score was about 76%, while the female students' mean score was about 63%. Male and female students' mean scores in the CG were found to be approximately 29%. The students in the TG (70%) were more successful than the students in the CG (29%) when the results of the final examination were evaluated. The female and male students in the TG were positively affected by the teaching method. Especially, the students in the TG initially focused on understanding the fundamental concepts of the problems and they understood the importance of physics concepts, solution, and controlling representations by problem solving with the help of peer instruction. These findings indicated that the teaching method appeared to help the students to determine the concepts/principles, to analyze the solution, and to evaluate the procedure. On the other hand, the female and male students in the CG showed no noticeable change.

Students' Opinions about the Instruction

The general opinions of the students were collected using the student essays. The students' opinions were separated into two groups (cognitive and affective ideas), positively and negatively. Female students (85%) and male students (93%) indicated their positive ideas about the instruction as follows:

Positive Cognitive Ideas

I think that this instruction method is a really *useful* tool for the students to *comprehend* the concepts and fundamental principles.

I *learned* to connect the problems related with topics and concepts.

This instruction method helped me to *remember* the equations and formulas.

I think that conceptual learning is *important* for problem solving.

This instruction method positively *changed* my way of thinking.

I *focused* on the concepts in the problems.

I began to *examine* the difference between relevant and irrelevant concepts in the problems.

I *solved* the difficult problems by simplifying as instructed in the teaching method.

I *got* the habit of thinking about concepts and problems and I believed that I *improved* myself with the help of this method.

Positive Affective Ideas

I would like to *apply* this method to other science courses.

By means of the teaching method, I started to have an *interest* in physics courses.

I *enjoyed* identifying a concept and solving a problem.

Learning is *easy* with the help of this teaching method.

Negative Cognitive Ideas

This method was *time-consuming*.

I sometimes *forget* to bring my flashcards.

Negative Affective Ideas

I would consider physics quite *boring*.

I was *nervous* if I showed the wrong flashcard and my friends would make fun of me.

Discussion

The differences between female and male students' conceptual learning and problem solving performance were investigated with two instruction methods; peer instruction and conventional instruction. The students' opinions were collected with written feedback about peer instruction in the treatment group. This research was mainly focused on three research questions, the findings of which are explained as follows:

With reference to the first research question (Are there any differences between female and male students' conceptual learning in the treatment group and the control group?), the effect of peer instruction on conceptual learning of female and male students in the treatment group was not found to be statistically significantly different. Many findings reported in the literature (Coletta et al., 2011; Donnelly et al., 2012; Kost-Smith et al., 2010; Madsen et al., 2013; Richardson & O'Shea, 2013) indicated that male students' conceptual learning performance generally outperformed female students' performance on Force Concept Inventory (FCI). However, in the present study peer instruction reduced the gender gap in the conceptual learning because of the increased classroom interactivity and collaboration.

The conventional instruction did not cause any noticeable change in female and male students' both conceptual learning and problem solving performance. The conventional instruction could not facilitate the students' comprehension of the physics concepts or fundamental principles, solving problems qualitatively and quantitatively, analyzing their solutions of problems, and improvement an alternative solution way (Crouch & Mazur, 2001; Gok, 2014; Madsen, McKagan & Sayre, 2013). Recently, alternative interactive engagement methods have been investigated by physics education researchers to overcome the deficiency of conventional instruction.

In contrary, the gender difference was observed in the students' problem solving performance. The performance of male students on quantitative problem solving was higher than the female students' performance in the treatment group. The results in the literature (Kost-Smith et al., 2010; Kost-Smith et al., 2009) supported the findings of the study. Male students were more likely to focus on systematically analyzing problems rather than the literal meaning of the problems according to female students in the final examination answers. Analyses of male students' problem solving processes indicated that they applied the task of representing, determining, planning, calculating, and evaluating the problems. Besides, the affecting factors on the results of the research might be listed as follows; a) male students' problem solving performance might be impressed by affective factors (the level of family interest in physics, parental affective support, having parents with science jobs, family' socioeconomic status, etc.) b) male students might be more motivated in a new approach, c) female students might be more anxious about problem solving, d) female students could have low confidence in problem solving, e) metacognitive knowledge, skills and beliefs of female and male students might affect the performance scores of the students. Limited numbers of studies on the gender gap related to affecting factors were conducted in the literature (Ding & Harskamp, 2006; Hazari, Sadler, & Tai, 2008; Redish, 2004; Udo et al., 2001). These studies showed that affecting factors played an important role in female and male students' problem solving performance. The most common factor was reported to be metacognitive knowledge (declarative, procedural and conditional) in problem solving. These findings are in response to the second research question (Are there any differences between female and male students' quantitative problem solving skills in the treatment group and the control group?).

With reference to the third research question (Does the teaching of peer instruction change female and male students' opinions about conceptual learning and problem solving?), the written essays indicated that most of the students were pleased with peer instruction which is very useful to comprehend the concepts, to gain the reasoning ability, and to solve problems. Even so, some of the students were not satisfied with peer instruction and found it as a time-consuming method. This drawback could be eliminated by exchanging the low-technology, coloured flashcard approach with the more technologically advanced tools such as classroom response system "clicker", which will speed up the feedback process and maximize the interaction time for the students in peer discussion.

Conclusions and Implications

In the light of the findings of the research, it could be concluded that peer instruction, one of the interactive engagement methods, had more positive effect on students' conceptual learning and problem solving performance than conventional teacher-centred instruction. Although male students' problem solving performance was higher than female students' problem solving performance with the help of peer instruction, the effect of peer instruction on female and male students' conceptual learning were not observed in this research. The majority of the female and male students declared that peer instruction is a useful method and more effective than conventional instruction.

Finally, the findings of the research surrounding the teaching of peer instruction have established evidence-based practices for interactive engagement in Turkey. This teaching method will require professional development for college physics teachers, but workshops can be organized country-wide or by web-based activities to support the implementation of peer instruction. Besides, the findings of the study should be confirmed through more researches with the same experimental design in different countries.

References

Baker, D., & Leary, R. (2003). Letting girls speak out about science. Journal of Research in Science Teaching, 40, 176-200.

Bates, S., Donnelly, R., MacPhee, C., Sands, D., Birch, M., & Walet, N. R. (2013). Gender differences in conceptual understanding of Newtonian mechanics: A UK cross-institution comparison. European Journal of Physics, 34, 421-434.

Coletta, V. P., Phillips, J. A., & Steinert, J. J. (2011). FCI normalized gain, scientific reasoning ability, thinking in physics, and gender effects. *Proceedings of the 2011 Physics Education Research Conference*, Omaha, Nebraska, US, 1413, 23-26.

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.), *Reviews in physics education research* (Vol. 1, 11 p.). College Park, MD: American Association of Physics Teachers. Retrieved September 3, 2014, from http://www.per-central.org/document/ServeFile.cfm?ID=4990

Cummings, K., Laws, P. W., Redish, E. F., & Cooney, P. J. (2004). *Understanding Physics*. New York: John Wiley & Sons.

Deslauriers, L., Schelew, E., & Wieman, C. (2011). Improved learning in a large-enrolment physics class. *Science, 332*, 862-864.

Ding, N., & Harskamp, E. (2006). How partner gender influences female students' problem solving in physics education. *Journal of Science Education and Technology, 15* (5), 331-343.

Docktor, J., & Heller, K. (2008). Gender differences in both force concept inventory and introductory physics performance. *Proceedings of the 2008 Physics Education Research Conference,* Edmonton, Alberta, Canada*, 1064*, 15-18.

Donnelly, R., MacPhee, C., & Bates, S. (2012). The performance gender gap in undergraduate physics. *Proceedings of the HEA STEM Learning and Teaching Conference.* London, England.

Fraser, J. M., Liman, A. L., Miller, K., Dowd, J. E., Tucker, L., & Mazur, E. (2014). Teaching and physics education research: Bridging the gap. *Report on Progress in Physics, 77* (032401), 1-17.

Gok, T. (2012). The impact of peer instruction on college students' beliefs about physics and conceptual understanding of electricity and magnetism. *International Journal of Science and Mathematics Education, 10*, 417-436.

Gok, T. (2014). An investigation of students' performance after peer instruction with stepwise problem-solving strategies. *International Journal of Science and Mathematics Education*. Advance online publication. doi:10.1007/s10763-014-9546-9.

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.

Hazari, Z., Sadler, P. M., & Tai, R. H. (2008). Gender differences in the high school and affective experiences of introductory college physics students. *The Physics Teacher, 46*, 423-427.

Heller, P., & Hollabaugh, M. (1992). Teaching problem solving through cooperative grouping. Part 2: Designing problems and structuring groups. *American Journal of Physics, 60*, 637-644.

Hestenes, D., Wells, M., & Swackhamer, G. (1992). Force Concept Inventory. *The Physics Teacher*, 30, 141-151.

Hoellwarth, C., Moelter, M. J., & Knight, R. D. (2005). A direct comparison of conceptual learning and problem solving ability in traditional and studio style classrooms. *American Journal of Physics, 73* (5), 459-462.

Kinnear, P. R., & Gray, C. D. (2008). *SPSS 16 made simple*. Hove: Psychology Press.

Kost-Smith, L. E., Pollock, S. J., & Finkelstein, N. D. (2009). Characterizing the gender gap in introductory physics. *Physical Review Special Topics-Physics Education Research, 5* (010101), 1-14.

Kost-Smith, L. E., Pollock, S. J., & Finkelstein, N. D. (2010). Gender disparities in second-semester college physics: The incremental effects of a "smog of bias". *Physical Review Special Topics-Physics Education Research, 6* (020112), 1-17.

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

Laws, P., Rosborough, P., & Poodry, F. (1999). Women's responses to an activity-based introductory physics program. *American Journal of Physics, 67* (7), 32-37.

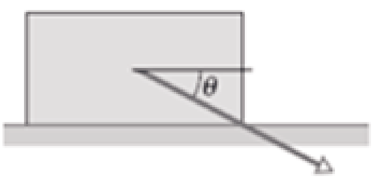
Lorenzo, M., Crouch, C. H., & Mazur, E. (2006). Reducing the gender gap in the physics classroom. *American Journal of Physics, 74* (2), 118-122.

Madsen, A., McKagan, S. B., & Sayre, E. C. (2013). Gender gap on concept inventories in physics: What is consistent, what is inconsistent, and what factors influence the gap? *Physical Review Special Topics-Physics Education Research, 9* (020121), 1-15.
Mazur, E. (1997). *Peer instruction: A user's manual*. Upper Saddle River, NJ: Prentice Hall.
Mazur, E., & Watkins, J. (2010). Just in time teaching and peer instruction. In S. Scott & M. Mark (Eds.), *Just in time teaching: Across the disciplines, and across the academy* (pp. 39-62). Sterling, VA: Stylus.
McDermott, L., Shaffer, P., & Rosenquist, M. (1995). *Physics by inquiry*. John Wiley & Sons, New York.
McDermott, L. C. (2001). Oersted medal lecture 2001: Physics education research – The key to student learning. *American Journal of Physics, 69* (11), 1127-1137.
Meltzer, D. E., & Manivannan, K. (2002). Transforming the lecture-hall environment: The fully interactive physics lecture. *American Journal of Physics, 70* (6), 639-654.
Nieminen, P., Savinainen, A., & Viiri, J. (2013). Gender differences in learning of the concept of force, representational consistency, and scientific reasoning. International Journal of Science and Mathematics Education, 11, 1137-1156.
Pollock, S., Finkelstein, N. D., & Kost-Smith, L. (2007). Reducing the gender gap in the physics classroom: How sufficient is interactive engagement? *Physical Review Special Topics-Physics Education Research, 3* (010107), 1-4.
Puente, S. M. G., & Swagten, H. J. M. (2012). Designing learning environment to teach interactive quantum physics. *European Journal of Engineering Education, 37* (5), 448-457.
Redish, E. F. (2003). Teaching physics with physics suite. Wiley, New York.
Redish, E. F. (2004). A theoretical framework for physics education research: Modeling student thinking. *Proceedings of the International School of Physics "Enrico Fermi" Course CLVI, Research on Physics Education, Italy, 156*, 1-64.
Reif, F. (1995). Millikan Lecture 1994: Understanding and teaching important scientific thought processes. *American Journal of Physics*, *63* (1), 17-32.
Richardson, C. T., & O'Shea, B. W. (2013). Assessing gender differences in response system questions for an introductory physics course. American Journal of Physics, 81 (3), 231-236.
Sayre, E. C., Franklin, S. V., Dymek, S., Clark, J., & Sun, Y. (2012). Learning, retention, and forgetting of Newton's third law throughout university physics. Physical Review Special Topics-Physics Education Research, 8 (010116), 1-10.
Scott, T. F., Schumayer, D., & Gray, A. R. (2012). Exploratory factor analysis of a force concept inventory data set. Physical Review Special Topics-Physics Education Research, 8 (020105), 1-10.
Tai, R. H., & Sadler, P. M. (2001). Gender differences in introductory undergraduate physics performance: University physics versus college physics in the USA. International Journal of Science Education, 23 (10), 1017-1037.
Turpen, C., & Finkelstein, N. H. (2009). Not all interactive engagement is the same: Variations in physics professors' implementation of peer instruction. *Physical Review Special Topics-Physics Education Research, 5* (020101), 1-18.
Udo, M. K., Ramsey, G. P., Reynolds-Alpert, S., & Mallow, J. V. (2001). Does physics teaching affect gender-based science anxiety? Journal of Science Education and Technology, 10 (3), 237-247.
Van Heuvelen, A. (1991). Learning to think like a physicist: A review of research-based instructional strategies. *American Journal of Physics, 59* (10), 891-897.
Watkins, J., & Mazur, E. (2013). Retaining students in science, technology, engineering, and mathematics (STEM) majors. Journal of College Science Teaching, 42, 36-41.

Appendix

(Sample Final Examination Problem)

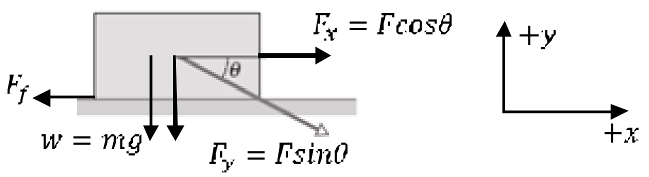
A 3.5 kg block is pushed along a horizontal floor by a force of magnitude 15 N at an angle θ= 40° the horizontal (Fig.). The coefficient of kinetic friction between the block and the floor is 0.25. Calculate the magnitudes of (a) the frictional force on the block from the floor and (b) the acceleration of the block.



Solution Steps:

I. Step: Physical Representation

The problem was visualized with the help of a coordinate system and a free-body diagram as represented below.



ISSN 1648–3898

II. Step: Concept(s) Representation
After the physical representations were illustrated, the fundamental concepts, known and unknown variables in the problem were determined as shown below.

Fundamental Concepts	Known Variables	Unknown Variables
Acceleration	$m = 3{,}5kg$	$f_k = ?$
Net force	$F = 15N$	$a = ?$
Normal force	$\theta = 40^0$	
Force of gravity	$\mu_k = 0.25$	
Force of kinetic friction	$g = 9.81m/s^2$	
Coefficient of kinetic friction		

III. Step: Solution Representation
Qualitative solutions were conducted with the help of needed formulas and equations. A mathematical model was established for findings desired unknown variables. After the desired unknown variables by using the given variables by means of qualitative solution were calculated quantitatively.

Qualitative Solution		Quantitative Solution	
y-direction	x-direction	Calculation for "F_f"	Calculation for "a"
$F_y = Fsin\theta$ $F_y = Fsin\theta$	$F_x = Fcos\theta$	$F_f = \mu_k F_N = \mu_k(w + F_y)$	$F_{net} = ma = F_x - F_f$
$F_N = w + F_y$	$F_f = \mu_k F_N$ $F_f = \mu_k(w + F_y)$ $F_{net} = ma$ $F_{net} = F_x - F_f$	$F_y = Fsin\theta = 15sin40^0$ $F_y = 9.64N$ $F_f = 0.25(34.33N + 9.64N)$	$ma = Fcos\theta - \mu_k(w + F_y)$ $a = \frac{Fcos\theta - \mu_k(w + F_y)}{m}$
		$F_f = 10.99N$	$a = \frac{11.49N - 0.25(34.33N + 9.64N)}{3.5kg}$ $a = 0.14m/s^2$

IV. Step: Controlling Representation
The units, signs, and magnitudes of the unknown variables were checked in this part.

Unit: N for force of kinetic friction; $m/s^2 m/s^2$ for acceleration
Sign: -x for force of kinetic friction; +x direction for acceleration
Magnitude: In this problem, $F_f F_f$ for force of kinetic friction and $a a$ for acceleration were calculated $10.99N$ and $0.14m/s^2$ respectively. This result ($F_f < F_x$ and $F_{net} \neq 0$; $a \neq 0$) was reasonable.

Received: *June 25, 2014* Accepted: *September 10, 2014*

Tolga Gok PhD., Associate Professor, Dokuz Eylul University, Izmir, Turkey.
E-mail: gok.tolga@gmail.com