MODESTUM
EURASIA Journal of Mathematics, Science and Technology Education, 2024, 20(8), em2480
ISSN:1305-8223 (online)
OPEN ACCESS
Research Paper
https://doi.org/10.29333/ejmste/14778

Peer instruction's Achilles' heel: An analysis of its ineffectiveness in confronting counterintuitive physics questions

Yerbol Ospanbekov [1], Samat Maxutov [2], Yerbol Sandybayev [3], Aknur Boranbekova [1], Nuri Balta [2*]

[1] Abai Kazakh National Pedagogical University, Almaty, KAZAKHSTAN
[2] SDU University, Kaskelen, KAZAKHSTAN
[3] Nazarbayev Intellectual School of Biology and Chemistry, Almaty, KAZAKHSTAN

Received 02 May 2024 ▪ Accepted 28 June 2024

Abstract

This paper examines peer instruction, an educational technique that enhances student-student interaction and student-teacher interaction, influencing students' performance on counterintuitive physics questions. The study employed a quantitative approach with 40 participants, 22.5% of whom were male and 77.5% were female. The counterintuitive dynamics test was used to measure students' counter intuitions in dynamics. The study results on counterintuitive questions showed no noticeable effect from peer instruction. Surprisingly, more students shifted from correct to incorrect responses after group discussions, leading to a decrease in the overall accuracy rate. Although peer instruction did not show a clear impact on counterintuitive physics questions, it is important for educators to recognize the complexity of peer instructions. Exploring different teaching methods that use peer interaction might help improve learning outcomes for counterintuitive physics questions.

Keywords: counterintuitive physics questions, group discussions, peer instruction, student performance

INTRODUCTION

Peer instruction is an educational technique that enhances student-student interaction and student-teacher interaction, influencing students' performance on conceptual questions (Mazur, 1997). This study examines the effectiveness of peer instruction, especially its impact on counterintuitive questions. Counterintuitive questions, which present scenarios that challenge common sense, have become a significant tool in science education to promote deeper understanding and higher-order thinking among students (Campanario, 1998).

The conventional teaching methods often fall short in addressing counterintuitive concepts in physics, which are crucial for developing a robust understanding of physical principles. This research aims to explore whether peer instruction can effectively enhance students' comprehension and performance on such questions.

This study seeks to evaluate the impact of peer instruction on students' performance and confidence in solving counterintuitive physics questions. It specifically investigates the shift in correctness and confidence levels before and after peer discussions.

LITERATURE REVIEW

Counterintuitive Questions

Counterintuitive questions have emerged as a useful tool for challenging students' traditional thinking in science education. By providing scenarios that challenge normal common sense, these questions aim to stimulate interest, drive motivation, and foster higher-order thinking. They inspire students to dive deeper into the underlying scientific ideas, creating a deeper comprehension of scientific concepts and improving their problem-solving skills (Campanario, 1998). Offering a unique way to evaluate students'

✉ Os_Erbol@mail.ru ✉ samat.maxutov@sdu.edu.kz ✉ sandybaev58@gmail.com ✉ aknur_0404@mail.ru
✉ baltanuri@gmail.com **(*Correspondence)**

Contribution to the literature

- This study challenges the belief that peer instruction always improves student outcomes by showing its ineffectiveness with counterintuitive physics questions.
- The research shows a shift from correct to incorrect answers after discussions, critically evaluating peer discussions in counterintuitive contexts.
- The study reveals that increased confidence doesn't always mean correct answers, especially for counterintuitive questions.
- This research highlights how cognitive biases like confirmation bias affect student performance on counterintuitive questions during peer instruction.
- The findings suggest the need for specialized strategies, like pre-discussion scaffolding and structured interactions, to improve peer instruction effectiveness.

understanding and problem-solving abilities, counterintuitive dynamic questions have emerged as a significant area of interest in physics education in recent years. In this review, we draw insights from distinct studies (Alvermann et al., 1995; Balta & Asikainen, 2019; Balta & Eryılmaz, 2017; Balta et al., 2020; Balta & Moğol, 2016; Guzzetti 2000) that collectively shed light on the effectiveness of using counterintuitive dynamic questions in science education.

Balta and Eryılmaz (2017) introduced the counterintuitive dynamics test (CIDT) as an assessment tool to evaluate students' understanding and comprehension of counterintuitive concepts in physics education. The development process of CIDT comprised of expert reviews and pilot studies. The final version of CIDT with 30 questions had a strong reliability coefficient (0.826) reflecting its potential as a powerful tool to measure students' comprehension of counterintuitive dynamic concepts. In their study, Balta and Eryılmaz (2017) compared CIDT with force concept inventory (FCI), claiming that FCI assesses common sense beliefs and misconceptions of the students while CIDT specifically targets counterintuitivity.

An experimental study conducted by Balta and Moğol (2016), compared the effectiveness of teaching with counterintuitive physics questions versus traditional teaching method. They found that using counterintuitive physics questions significantly improved students' achievement in understanding Newton's law of motion. This study found that counterintuitive questions created debates among students which further led to class discussion and exploration of the problems.

Another mixed method study was conducted by Balta and Asikainen (2019), who investigated Olympians' and regular students' successes in solving counterintuitive dynamics problems. The main goal of the study was to identify the success rates among Olympiad competitors and novice high school students in relation to solving counterintuitive dynamic problems in physics to provide adjustments in teaching strategies. According to the quantitative results, Olympiad competitors outperformed regular high school students in solving counterintuitive questions. Experts outscored novices on all counterintuitive questions, and the disparity widened as the degree of difficulty in the problems arose. The qualitative findings highlight distinctions in the problem-solving approaches between the two groups. For instance, high school students often employed superficial problem-solving approaches and provided incomplete answers, while Olympiad students offered more comprehensive solutions to the given problems.

A study carried out by Balta et al. (2020) explored the students' success in solving counterintuitive physics problems. Their findings can be categorized into three patterns. First, deep interpretation and success, students who were able to see the hidden structure and underlying physical principles of the problem tend to come up with scientifically accepted counterintuitive responses. Second, superficial and intuitive interpretation, when students focus on literal meaning of the problem, their responses tend to be more superficial and based on everyday knowledge. This type of problem-solving approach has low rates of successful responses. Third, some students were able to provide quick responses that were accurate even if students couldn't come up with scientific reasoning and correct mathematical calculations. This suggests the potential development of physics intuition (Balta et al., 2020).

Collectively, those studies underscore the value of using counterintuitive dynamic questions as a pedagogical tool in science and physic classes. The above research findings suggest that prospective teachers would benefit significantly from counterintuitive questions, which lead to expanded understanding of physics concepts and improved problem-solving skills.

Peer Instruction

Peer instruction is a student-centered approach designed to support students' engagement in classrooms and improve educational outcomes overall. Mazur (1997) critiqued the traditional way of lecturing, claiming that students passively absorb the provided information without clear understanding of the main idea of the underlying concepts. Mazur (1997) defined peer instruction as a pedagogical approach that promotes students to think and discuss concepts with their peers during class. He outlines several stages in the application of peer instruction. First, the instructor presents multiple-choice questions known as concept tests. Second, students individually respond to these questions. Third, students engage in group discussions to deliberate their responses. Subsequently, students vote again on the question after peer discussions. Finally, the instructor facilitates a class-wide discussion, addressing any misconceptions and providing correct answers.

To begin with, the studies collectively have shown the effectiveness of application of peer instruction as an educational approach to enhance students' learning outcomes, motivation, confidence and develop critical thinking skills in various educational settings (Chien et al., 2015). For instance, Al-Hebaishi (2017) explores the impact of the peer instruction method on the conceptual comprehension of pre-service teachers. The results showed statistically significant differences in treatment and comparison group while examining peer instruction and students' achievement.

In a similar vein, the research conducted by Pilzer (2001), investigated the effectiveness and application of peer instruction in physics and calculus courses. The study underscores the development of peer instruction approach by Mazur (1997) and enumerates its application stages. The results of the study are in agreement with the previous studies (Al-Hebaishi, 2017; Mazur, 1997). The study results indicate significant improvement in students' conceptual understanding and standard problem-solving skills in physics classes where peer instruction was used. Pilzer (2001) claims that using concept tests during lecture classes resulted in a deeper grasp of the material.

The literature review uncovered studies that focus on peer instruction in relation to class size (Porter et al., 2013); the impact of peer instruction on students' fail rates (Porter et al., 2013b); positive students' attitudes towards peer instruction (Porter et al., 2016; Zhang et al., 2017); the effectiveness of using peer instruction (Al-Hebaishi 2017; Cummings & Roberts, 2008; Fagen et al., 2002; Pilzer 2001; Turpen & Finkelstein, 2009); learning outcomes and motivation (Crouch & Mazur, 2001; Gok, 2012a; Zingaro 2014); and a comparison of peer instruction with standard lecturing (Simon et al., 2013). Collectively, the above research findings suggest that the application of peer instruction in science classes successfully affects students by developing their conceptual understanding, motivation to learn, overall learning experience while reducing fail rates.

Although studies conducted by Zhang et al. (2017), and Turpen and Finkelstein (2009) demonstrate positive outcomes of peer instruction implementation on students learning outcomes, they also touched upon a few potential obstacles. For instance, faculty resistance was one of the potential challenges. According to the study findings, some instructors are hesitant to apply peer instruction strategy in their teaching as they are not familiar with the approach, need training to use it effectively, some instructors apply it superficially, challenges in shifting from traditional way of teaching, and lack of awareness of the potential benefit of research-based and student-centered teaching method (Turpen & Finkelstein, 2009). Another common difficulty mentioned in earlier papers relates to resource allocation such as technology and time constraints as implementing peer instruction productively needs time and dedication.

Confidence

Generally, confidence refers to one's self-belief and certainty in one's abilities to successfully master academic problems, achieve desired learning outcomes, and actively participate in the learning process. Numerous studies investigated students' confidence in relation to academic achievement (Brooks & Koretsky, 2011; El-Sayed 2013; Elizabeth et al. 2016), conceptual learning and motivation (Gok 2012a; Vickrey et al., 2015), teaching style (Mallow, 1995) and various factors (Parsons et al., 2009; Pratama, 2018; Tullis & Goldstone, 2020).

Various factors, including instructional methods, peer engagement, school environment and students' internal factors, can significantly influence students' confidence level. An earlier study done by Mallow (1995) demonstrated that instructors' teaching style impacts students' self-belief and confidence in the subject. According to Mallow (1995), German physics instructors applied more interactive, interpersonal, and engaging teaching strategies which encouraged students' active participation thus developing students' confidence. In contrast, American instructors used traditional teaching method with more teacher-centric approach affecting students' low confidence level. A more recent study by Parsons et al. (2009) found a positive correlation between students' confidence and their achievements in mathematics.

In terms of the impact of peer instruction on students' confidence, several studies including Gok (2012a), Tullis and Goldstone (2020), and Vickrey et al. (2015) found correlation between peer instruction and its impact on students' confidence level. Gok (2012a) found that peer instruction increased student active involvement during the learning process which contributed to higher levels of confidence in learning the concept matter. Similar findings were found by Tullis and Goldstone (2020). Moreover, Tullis and Goldstone (2020) state that peer instruction scaffolded students to identify their misconceptions, correct those misconceptions and improve their understanding of the concepts and consequently increase confidence level. In a similar vein, the meta-analysis by Vickrey et al. (2015) found that involving in peer discussion led students to explore and deepen their understanding of the concepts, and by deepening their knowledge students increased their confidence level.

Another bunch of empirical studies that have positive results in relation to the use of peer instruction and its effect on students' confidence level are reviewed below. Studies conducted by Brooks and Koretsky (2011), Elizabeth et al. (2016), and El-Sayed (2013) focused on the impact of peer instruction on students' performance and confidence. These studies aimed to determine whether engaging in peer discussions influenced students' confidence in their answers and understanding of the material. For example, El-Sayed (2013) found that nursing students who utilized peer instruction experienced improved academic performance in their nursing administration course. Additionally, participants in the research reported enhanced communication skills and an overall positive learning experience with the peer instruction approach. Similar positive learning experience was found by Elizabeth et al. (2016) within mathematics discipline. The scholars also claim potential benefits on utilizing peer instruction approach in math courses as it enhances students understanding of math concepts and affects students' academic achievement.

The literature review revealed the positive effects of peer instruction on student confidence, learning outcomes, attitudes, and academic performance in various educational settings. Moreover, previously identified challenges such as instructor resistance, time constraints, and resource allocation (Turpen & Finkelstein, 2009) in implementing peer instruction highlight the importance of addressing these barriers to promote its effective use to improve student achievement and learning experience. However, the effect of peer instruction in counterintuitive environments needs to be explored.

Research Questions

1. Does peer discussion effective in counterintuitive questions?
2. Does peer discussion increase students' confidence in solving counterintuitive questions?
3. Is confidence linked to correctness of counterintuitive questions?
4. Does the initial confidence of students influence the quality of peer discussion?
5. Does high achievers of counterintuitive questions effective in peer discussion?
6. Does increase in confidence predict correctness of counterintuitive questions?

Hypotheses

Based on the literature review and research questions, we formulated the following hypotheses for this study. These hypotheses are grounded in existing research and aim to investigate the effectiveness of peer instruction on counterintuitive physics questions.

H1: Peer discussion increases the correctness of responses to counterintuitive questions.

Evidence supporting this hypothesis comes from multiple studies indicating that peer instruction enhances students' understanding and performance. For instance, Crouch and Mazur (2001) and Fagen et al. (2002) demonstrated that peer instruction significantly improves students' conceptual understanding and problem-solving skills in physics.

H2: Peer discussion increases students' confidence in solving counterintuitive questions.

Several studies, including those by Gok (2012a) and Tullis and Goldstone (2020), have found that peer instruction not only boosts students' engagement but also their confidence in their answers. These studies suggest that discussing concepts with peers helps students solidify their understanding and feel more assured about their responses.

H3: There is a positive correlation between students' confidence and the correctness of their responses.

Research by DeSoto and Roediger III (2014) indicates a positive relationship between confidence and accuracy in students' responses. This hypothesis is based on the premise that students who are more confident in their answers are more likely to provide correct responses.

H4: High achievers are more effective in influencing their peers' responses during discussions.

Studies by Carter et al. (2003) and Cortright et al. (2005) suggest that high-achieving students can positively influence their peers during collaborative learning activities. This hypothesis posits that students with a strong understanding of counterintuitive questions will help their peers improve their responses through discussion.

H5: An increase in confidence predicts the correctness of responses after peer discussion.

Evidence from studies by Gok (2013) and Vickrey et al. (2015) supports this hypothesis, indicating that as students' confidence increases through peer discussions, their likelihood of providing correct answers also increases. This hypothesis aims to explore the predictive relationship between changes in confidence and the accuracy of responses.

METHODS

This research employs a quantitative approach to investigate the impact of peer discussion on students'

A block of mass M is hung on a rope X. An identical rope Y is connected to the bottom of the block as shown in the figure. Rope Y is pulled down suddenly with a big enough force.

Which rope will break first?

a. X

b. Y

c. Both will break at the same time.

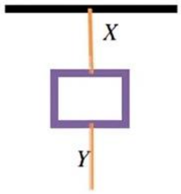

Figure 1. A counterintuitive question I (Balta & Eryilmaz, 2017)

performance and confidence in solving counterintuitive questions in physics.

Participants

The study was carried out at the end of the second semester of academic year 2022/2023. CIDT (Balta & Eryilmaz, 2017) was administered to a cohort of 40 Kazakh prospective teachers (22 and 18 students in each group) enrolled at the faculty of science within the Al-Farabi Kazakh National University. From the participants, nine (22.5%) were male and 31 (77.5%) females. Their average age was about 23. They were from middle income class. Before taking the test, these students had completed two semesters of a calculus-based introductory physics course, which covered mechanics initially and then moved on to electricity and magnetism. This cohort comprised students at the culmination of their second year of study who were aiming to become high school physics teachers.

Participants were taking the course "solving physical problems I" when we collected data. This course is designed for students specializing in Physics with the aim of preparing them as future teachers in modern schools. The course had a duration of 45 hours each week and aims to achieve several key learning outcomes such as equipping students with effective teaching strategies and techniques. The specific topics covered in the course include the fundamentals of dynamics, Newton's laws (I, II, and III), inertial reference systems, weight, and power. Additionally, the course is intended to contribute to the development of educational achievements, particularly in preparing future physics teachers for the national qualification testing. Moreover, it is a comprehensive course that aims to prepare future physics teachers with the knowledge, skills, and abilities required to effectively teach physics, solve problems, and contribute to the educational process in secondary schools and vocational schools.

Instrument

CIDT was previously verified and assessed by Balta and Eryilmaz (2017) to measure students' counter intuitions in dynamics. They validated the CIDT by administering it to 229 students. The final version of CIDT with 30 items was established after analyzing items from the data obtained from a pilot study with 89 students. The KR 20 reliability score was .826 and Ferguson's delta discriminatory power was .952. An example of a counterintuitive question from CIDT is, as follows (**Figure 1**).

The answer to this question appears to students so easy and obvious that rope X will break first.

The dynamics concepts in CIDT are grouped into three main categories and 11 subcategories. Classification of the dynamics concepts analyzed in the CIDT are as follows.

First law

Inertia–No motion: Item 3

Inertia–Motion: Item 1 and item 2

Second law

Only one object–With friction: Items 11, 12, 14, and 24

Only one object–Without friction: Items 6, 8, 20, and 25

System–Without friction: Items 4, 15, and 18

Tension related: Items 5, 9, 26, and 27

Action reaction related: Items 13, 16, and 28

Kinematics related: Items 7, 22, and 23

Equilibrium related: Item 19 and item 29

Third law

Impulsive force: Item 17 and item 21

Continuous force: Item 10 and item 30

Procedure

CIDT was administered to two classes during course hours. CIDT was uploaded to Socrative (an online assessment tool, see Balta & Güvercin, 2016) and questions were posed one by one by the first author in his physics class. Per each question, students made individual selections of their answers, noted their level of certainty, engaged in discussions regarding their answers with their peers, and subsequently indicated any potentially altered answers along with their renewed confidence level. Students made individual selections of their answers and noted their level of confidence on a scale from 1 to 10, where 1 indicated "not confident at all" and 10 indicated "extremely confident." The utilization of peer instruction incorporating confidence levels followed these steps:

1. Pose a counterintuitive question (1-2 min).
2. Collect responses from students.
3. Collect their level of confidence.
4. Facilitate discussions among students (2-3 min).
5. Gather students' responses once again.
6. Collect visit their confidence levels for a second time.

Applying all these steps took approximately 6 minutes for each question. A total of 30 questions adds up to 180 minutes which took us 4 class hours to collect data.

Data Analysis

The data preparation was done in MS Excel and analysis was conducted by using Jamovi version 2.4.8. Descriptive statistics were done to have a rough idea about the effect of peer discussion as well as the shift in accuracy from before discussion to after discussion for individual questions. Comparison of pre-discussion and post-discussion achievement scores and confidence scores were compared through Wilcoxon rank test (first and second RQs). Correlation analyses were done to see the relationship between correctness before the discussion and after the discussion (third RQ. To explore whether confidence is linked to correctness we computed Kendall's tau correlation coefficient (fourth RQ). To see if the correctness in both pre and after discussion is an indicator of achievement, we calculated the Phi correlation coefficient (fifth RQ). Finally, we carried out simple regression analysis to observe if correctness can be predicted from the change in confidence (sixth RQ).

1. Does peer discussion effective in counterintuitive questions?
2. Does peer discussion increase students' confidence in solving counterintuitive questions?
3. Is confidence linked to correctness of counterintuitive questions?
4. Does the initial confidence of students influence the quality of peer discussion?
5. Does high achievers of counterintuitive questions effective in peer discussion?
6. Does increase in confidence predict correctness of counterintuitive questions?

Ethical Considerations

The research subjects were regarded as independent individuals, and their choices regarding participation were honored. Prior to commencing data collection, consent from the participants was obtained. The students were assured that their test scores would remain confidential and would not affect their classroom evaluations. Furthermore, ethical issues were sought from the Ethical Review Committee at the Abai Kazakh National Pedagogical University, and it was granted.

Table 1. Wilcoxon W test for the comparison of pre- and post-discussion scores

			Statistics	p-value
Pre	Post	Wilcoxon W	113	0.286

RESULTS

Does Peer Discussion Effective in Counterintuitive Questions?

Peer instruction with counterintuitive questions was carried out in two classes of each 22 and 18 students. Thirty counterintuitive physics questions were solved during the implementation of peer instruction. Namely, we collected a total of 40*30 = 1,200 responses from students. Of the 1,200 pre discussion responses 291 (%24.3) were correct and 909 (%75.7) were incorrect. Similarly, 297 (%24.7) post discussion responses were correct and 903 (%75.3) were incorrect. This means that only a total of 24.7% of students exhibited either an improvement in the number of correct responses or maintained their level from before the discussion to after the discussion.

Initially, we examined the collective achievement of students in counterintuitive questions. Prior to engaging in discussion, the students exhibited a mean and median score of 7.28 (with a standard deviation of 1.77) and 7.00, respectively. Following the discussion, the mean and median scores were 7.42 (with a standard deviation of 1.62) and 7.00, respectively. Considering a maximum attainable score of 30, it becomes evident that students' success in counterintuitive questions remains notably low both prior to and following peer discussions. The shift from pre-discussion to post-discussion is virtually negligible.

Then, we statistically investigated the alteration in correctness due to peer discussions. An overall comparison of pre-discussion and post-discussion scores are compared through Wilcoxon rank test because of ordinal nature of data. The comparison of these scores through the Wilcoxon rank test is indicated in **Table 1**.

As indicated in **Table 1** there is no statistically significant difference between pre-discussion and post-discussion scores (p = .286). In other words, peer discussion is not effective in increasing students' success in counterintuitive questions.

We also delved into the shift in accuracy from before discussion to after discussion for individual questions, and these outcomes are graphically represented in **Figure 2**. The data illustrates an inconsistent

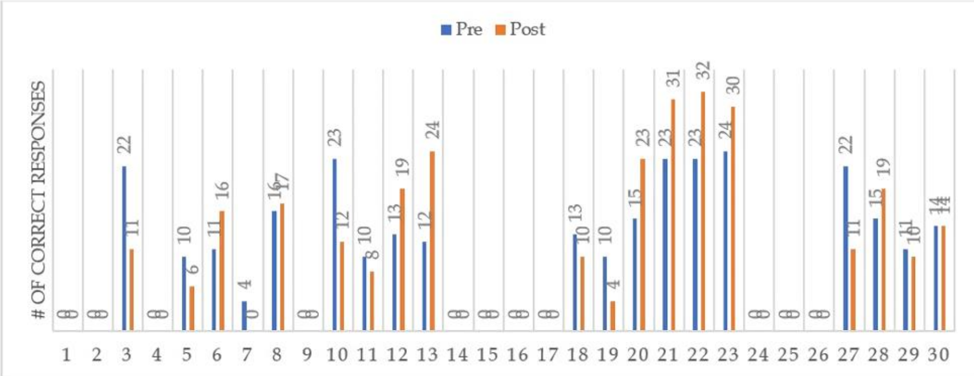

Figure 2. Students' correct responses before and after discussion for each question (Source: Authors' own elaboration)

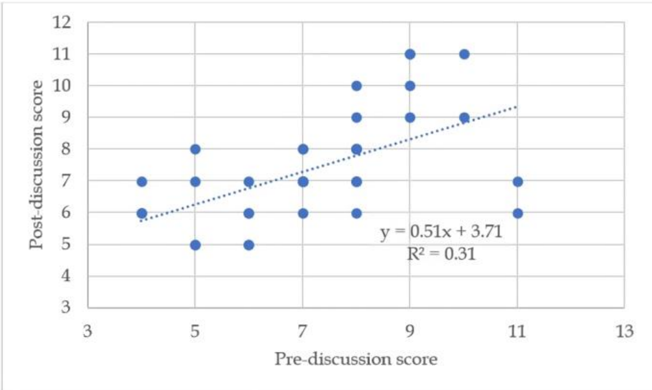

Figure 3. The relationship between pre- and post-discussion correctness (Source: Authors' own elaboration)

improvement in accuracy of responses when transitioning from pre-discussion to post-discussion.

Figure 2 illustrates that for questions 1, 2, 4, 9, 14, 15, 16, 17, 24, 25, and 26, students did not provide correct responses both prior to and following the discussion (36.7%). In contrast, for questions 3, 5, 7, 10, 11, 18, 19, 27, and 29, peer discussions led to a decrease in the number of correct responses (30.0%). On the other hand, questions 6, 8, 12, 13, 20, 21, 22, 23, and 28 saw an increase in the number of correct responses because of peer discussion (30.0%). The 30th question exhibited no change before and after the discussion. Notably, the 7th question initially had four correct responses before the discussion, but none remained after the discussion.

Peer discussion exhibited its most pronounced impact on questions 13th, 21st, and 22nd, with correct response counts rising by 12, 8, and 9, respectively. Conversely, a significant decline in correct responses was observed in questions 3, 10, and 27, where each question experienced a loss of 11 correct answers after the discussion.

We visualized the changes in correctness from before the discussion to after the discussion for each question in **Figure 3**. This is a correlation graph for all 40 students in two classes.

The dashed line in **Figure 3** indicates the line of best fit for the observed data. Each point represents the number of correct responses before and after peer discussion for a student. As indicated by the equation of the best fit line number of correct answers increased from pre-discussion to post discussion by .51 points.

We analyzed the enhancement in performance from before to after discussions by tracking the accuracy of responses during the discussion. **Figure 4** illustrates the progression of the percentage (and count) of accurate responses from before to after the discussion. In the upper row, we display whether students' initial answers were correct or incorrect; the middle row indicates whether students agreed or disagreed with their peers; the final row depicts whether students' responses were correct or incorrect after the discussion. Moreover, **Figure 4** presents the level of confidence linked with each pathway. The bottom part of each entry presents the average confidence of the students.

In a general sense, 26% of accurate responses transitioned to being incorrect, whereas a mere 9% of inaccurate responses shifted to being accurate after the discussion. Among questions where students were initially correct but held differing opinions from their peers, only 3% of answers transformed into incorrect responses post-discussion. Among questions where

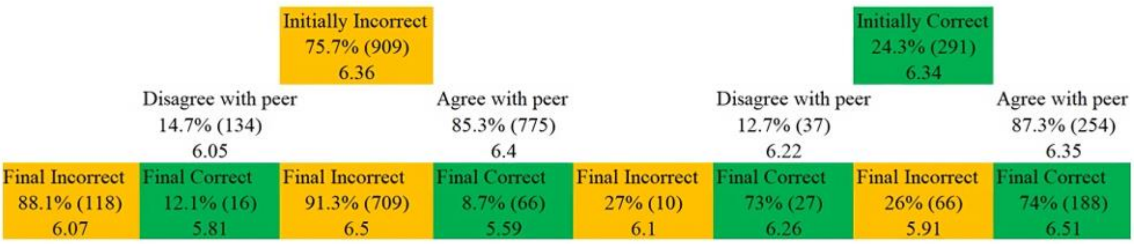

Figure 4. The progression of responses from before-discussion to after-discussion (Source: Authors' own elaboration)

Table 2. Descriptive statistics for pre- and post-confidence of each class

Class	N	Mean	Median	SD	SE
G1-1stCon	660	6.38	6.00	1.40	0.0543
G2-1stCon	540	8.22	8.00	1.09	0.0426
G1-2ndCon	660	6.31	6.00	1.23	0.0529
G2-2ndCon	540	8.19	8.00	1.11	0.0477
All-1st confidence	1,200	6.35	6.00	1.32	0.0382
All-2nd confidence	1,200	8.21	8.00	1.10	0.0318

Table 3. Confidence changes from before the discussion to after the discussion

Comparisons		Statistic	p	Effect size*
G1-1stCon	G2-1stCon	560	< .001	-0.993
G1-2ndCon	G2-2ndCon	800	< .001	-0.986
1st confidence	2nd confidence	2765	< .001	-0.99

Note. *Rank biserial correlation

students were initially wrong and held conflicting opinions with their peers, only 2% were modified to the correct answer.

Does Peer Discussion Increase Students' Confidence in Solving Counterintuitive Questions?

Table 2 displays the descriptive statistics for pre and post confidence of each class. There were 22 (G1) and 18 (G2) students in each class which resulted in 660 (22*30) and 540 (18*30) responses.

Students rated their confidence level between 1 and 10. The average initial confidence was 6.35 while the final confidence was 8.21. The significance of change for all students and for each class was tested with Wilcoxon W because of the ordinal nature of data. **Table 3** indicates the statistically significance confidence changes from before the discussion to after the discussion.

There is a significant increase in the confidence because Wilcoxon W revealed $p < .001$ for the difference between initial and final confidences of each class as well as overall participants (**Table 3**).

Is Confidence Linked to Correctness?

We hypothesis that the variation in the frequency of switching between correct and incorrect answers could potentially be attributed solely to differences in confidence. In other words, it is expected for answers that are more confidently given to have a higher likelihood of being correct. To explore whether confidence is linked to correctness, we computed Kendall's Tau correlation between confidence and correctness of the answers both prior and after the discussion, as demonstrated in **Table 4**. The first column is for pre-discussion and the second is for post-discussion data.

The average correlation (Kendall's tau) between both initial and final confidence, and initial and final correctness (overall -.001, and .036, respectively) was found to be an insignificant correlation ($p > .05$) indicating that greater confidence was not associated with being correct.

Table 4. The Kendall's tau correlation between correctness and confidence before and after discussion

Class	Kendall's tau (sig.)	Kendall's tau (sig.)
G1	-.035 (p = .316)	.022 (p = .533)
G2	-.005 (p = .889)	-.040 (p = .311)
Overall	-.001 (p = .966)	.036 (p = .175)

Table 5. Regression analysis of predicting correctness–All post response

Predictor	Estimate	SE	t	p
Intercept	.182	0.023	7.84	< .001
changeInConf (2nd-1stconf)	.036	0.011	3.37	< .001

Does the Initial Confidence of Students Influence the Quality of Peer Discussion?

We hypothesized that the more initial confidence will lead to quality discussion, and which then affect both own and partners' final correctness. To test this, we calculated the correlation between students' initial confidence and the response after discussion. The Kendall's Tau coefficient was found to be -.047 indicating an insignificant ($p > .05$) negative correlation. In other words, initial confidence did not affect the final correctness.

This was further confirmed by the association between final confidence and final response. Kendall's tau correlation (.036) between final confidence and final response was found to be an insignificant correlation ($p > .05$) indicating that greater confidence after peer discussion was not associated with final correct answer.

Does High Achievers Effective in Peer Discussion?

We further hypothesized that the correctness in both pre and after discussion is an indicator of achievement and those students would change their partners opinion. To test this argument, we looked at the association between students who are correct in their both responses and if their partners have correct answer. Phi correlation coefficient was calculated (as both variables are dichotomous) to be a significant moderate value of .476 ($p < .001$). This finding confirms the hypothesis that high achievers are effective in counterintuitive questions in influencing their partners.

Does Increase in Confidence Predict Correctness?

To respond to this question, we constructed a simple regression (**Table 5**) for the difference in final and initial confidence (independent variable), and the response after peer discussion (dependent variable).

It is found that increase in confidence significantly predicted the final response ($\beta = .036$, $p < .001$).

Moreover, we found a correlation coefficient of r = .097. Which means that 9.4% ($R^2 = 0.94$) of the final correctness can be predicted by the increase in the confidence.

DISCUSSION

To assess the impact of peer instruction on student learning, we conducted an analysis of student responses and their confidence levels before and after group discussions in two separate physics classes. The results indicated that engaging in discussions with a partner did not lead to a significant enhancement in accuracy across both classes. In this research, it was observed that just 24.7% of students either enhanced their performance by providing more correct responses or maintained their pre-discussion level, while a significant 75.3% of the student cohort demonstrated a decline in their responses after the peer instruction discussion. There is evidence in other research that switching in directions is negatively associated with student learning, that is, switching from right to wrong and from wrong to different wrong. For example, Miller et al. (2015) observed 27% of switching to wrong answer, and Tullis and Goldstone (2020) observed 12% of switching to wrong answer. However, the extent of the adverse impact of peer instruction, in our study, outweighed the positive effects observed. This finding clearly contradicts prior research. In many different research areas, the effect of peer discussion has proved to be positive (Balta et al., 2021; Crouch & Mazur, 2001; Fall et al., 2000; Giuliodori et al., 2006; Simon et al., 2013; Tullis & Goldstone, 2020).

Numerous studies in the field of physics education have demonstrated a substantial impact of peer instruction on students' conceptual learning. (Crouch & Mazur, 2001; Cummings, & Roberts, 2008; Dancy et al., 2016; Gok, 2012a, 2012b; Lasry et al. 2008; Miller et al., 2015). Contrary to all these findings we demonstrated the negative effect of peer instruction in counterintuitive questions. Cognitive biases, such as confirmation bias or anchoring, can influence how students interpret information (Royce et al., 2019). These biases can make it challenging for students to accept counterintuitive explanations that go against their intuitions. Furthermore, students may not have prior exposure to counterintuitive physics concepts (Balta et al., 2019). This lack of familiarity can make it challenging to grasp and apply counterintuitive concepts.

Among the 30 CDIT questions only nine (approximately one-third) demonstrated improvements because of peer discussions, while eleven questions did not exhibit any improvement stemming from these discussions. Despite prior research suggesting an increase in the number of correct responses (Giuliodori et al., 2006; Smith et al., 2009), our findings did not reveal the positive effects of peer instruction, particularly in the case of counterintuitive questions. Regrettably, more students transitioned from correct responses to incorrect ones than the reverse, resulting in a decrease in accuracy following the discussions. This finding contradicts the findings of both Bian et al. (2018), and that of Tullis and Goldstone (2020) who indicated that significantly more students changed their answers from incorrect to correct than from correct to incorrect. The disagreement of our results with previous findings can be attributed to the nature of counterintuitive questions. As these questions counter to students' institutions the success rate in these questions is quite low (Balta & Eryilmaz, 2017). Out of the 30 questions, 36.7% remained unanswered correctly both before and after the peer discussion. This also signifies a low rate of success in addressing counterintuitive questions. Moreover, overreliance on intuition may lead to such results (Van Dooren, 2008). Physics often requires students to rely on mathematical models and principles that may contradict their intuitions.

To understand the ineffectiveness of peer instruction in our study, it is important to analyze the characteristics of the items that showed improvement after peer discussion and those that did not. Based on the CIDT questions and the data from **Figure 2**, we can identify specific patterns. Items that showed improvement after peer discussions generally shared these characteristics: Items involved simpler scenarios and required less complex reasoning. For example, items related to basic applications of Newton's laws without additional complicating factors like friction or multiple forces (e.g., items 12, 20, 22, and 23). Items that had a straightforward focus on a single concept, such as tension or basic kinematics, saw better results (e.g., items 6, 13, and 21). Items that aligned more closely with everyday experiences and intuitions also showed improvements, as these were easier for students to discuss and reach a consensus (e.g., item 8 and item 28).

In contrast, items that did not improve or showed a decline in correct responses typically had these characteristics: Items that required integrating multiple concepts or involved multi-step problem-solving, leading to confusion and incorrect peer guidance (e.g., items 4, 15, and 16). Items that strongly contradicted everyday experiences or had counterintuitive elements were particularly problematic, as they were difficult to resolve correctly through peer discussion alone (e.g., items 1, 2, 14, 17, 24, 25, and 26). Items where the problem setup could be interpreted in multiple ways often led to mixed and incorrect responses during discussions (e.g., items 9, 18, and 29).

Moreover, our analysis revealed that peer discussions did boost students' confidence levels. This is what generally existing research confirms (Bian et al., 2018; Gok, 2013; Tullis & Goldstone, 2020; Zingaro, 2014). However, while many studies have yielded positive correlations between confidence and accuracy (DeSoto & Roediger III, 2014), we found no significant correlation between confidence and accuracy. Increased confidence was not necessarily indicative of a correct response, and greater confidence after peer discussions did not reliably lead to a final correct answer. This phenomenon can also be ascribed to the inherent characteristics of counterintuitive questions. Students tend to perceive the answers to such questions as self-evident and straightforward (Campanario, 1998). Furthermore, these answers often come across as highly persuasive to students, leading them to approach these problems with a superficial approach and respond confidently based on their intuition. However, it's worth noting that the correct answer typically contradicts one's intuitive reasoning (Balta et al., 2019).

As expected, high-achieving students were effective in influencing their partners in switching to correct response. Carter et al. (2003) found that high-achieving students collaborating with their lower-achieving peers used a greater number of words, engaged in more conversational exchanges, and demonstrated a higher frequency of supportive actions compared to when they collaborated with another high-achieving partner.

CONCLUSIONS

To examine if peer instruction benefits student learning, we analyzed student answers and confidence before and after discussion across two physics classes. Discussing a question with a partner did not significantly improve accuracy across classes. Of the 30 CDIT questions only nine (about one third) benefited from peer discussion, conversely eleven questions did not benefit from peer discussion. Even do previous research indicates an improvement in the number of correct responses (Smith et al., 2009) our results for counterintuitive questions did not show the effect of peer instruction. Unfortunately, more students switched from correct answers to incorrect answers than vice versa, leading to a disimprove in the number of accuracies following discussion. The summary of other findings are as follows: peer discussion increase students' confidence; no confidence was linked to accuracy; greater confidence was not associated with being correct; greater confidence after peer discussion was not associated with final correct answer; high achievers were effective in counterintuitive questions in influencing their partners; increase in confidence significantly predicted the final response; a partner's greater confidence did not increase the likelihood of changing to their answer.

The relatively small sample size of 40 students limits the generalizability of our findings. It may not represent the broader population of students or other institutions with different demographics and academic backgrounds. Since our study focused on a specific group of Kazakh students at one faculty, the results may not apply to a more diverse student population, including those from different cultural or educational backgrounds. There might be a possibility of response bias in self-reported data or participant behavior during the study, which could affect the accuracy of the results.

Suggestions and Future Implications

Firstly, the observed negative impact of peer instruction on correctness in counterintuitive questions calls for a thorough examination of the underlying reasons. Future research should delve into cognitive biases, such as confirmation bias and anchoring, that may influence students' interpretation of information and hinder their acceptance of counterintuitive explanations.

Secondly, the study highlights the need for broader investigations into the generalizability of these findings across diverse academic disciplines beyond physics. The observed patterns and challenges in peer instruction may vary in different subject areas, and extending this research to other fields can provide valuable comprehensions into the hints of peer learning effectiveness.

Thirdly, the low success rate in addressing counterintuitive questions highlights the importance of developing targeted instructional strategies to enhance students' ability to tackle such questions. Tailoring pedagogical approaches to address the unique challenges posed by counterintuitive concepts can contribute to improved learning outcomes.

Fourthly, to improve the effectiveness of peer instruction for counterintuitive questions, several strategies can be employed. Pre-discussion scaffolding is essential to provide background information and foundational knowledge before engaging in peer discussions. Mini-lectures or guided activities can help clarify complex concepts (Lasry et al., 2008), while visual aids, demonstrations, and simple examples can prime students on the key principles underlying the counterintuitive questions. Structured peer discussions, such as think-pair-share, where students first think individually, then discuss with a partner, and finally share with the larger group, can help ensure focused and productive discussions (Crouch & Mazur, 2001). Instructor facilitation is crucial during peer discussions. Instructors should actively monitor discussions, provide hints, and correct misconceptions in real-time (Turpen & Finkelstein, 2009). After discussions, offering immediate feedback on the correctness of answers and explaining the reasoning behind the correct solutions can reinforce learning (Smith et al., 2009). The use of concept tests designed to target common misconceptions can help students recognize and correct their intuitive errors (Fagen et al., 2002).

Author contributions: All authors have equally contributed to the study and agreed with the results and conclusions.

Funding: This research is funded by the Science Committee of the Ministry of Education and Science of the Republic of Kazakhstan (Grant No. AP14972674).

Ethical statement: The authors stated that the research subjects were regarded as independent individuals, and their choices regarding participation were honored. Prior to commencing data collection, consent from the participants was obtained. The students were assured that their test scores would remain confidential and would not affect their classroom evaluations. Furthermore, ethical issues were sought from the Ethical Review Committee at the Abai Kazakh National Pedagogical University, and it was granted on 26 September 2023 (Approval code: 1).

Declaration of interest: No conflict of interest is declared by the authors.

Data sharing statement: Data supporting the findings and conclusions are available upon request from the corresponding author.

REFERENCES

Al-Hebaishi, S. M. (2017). The effect of peer instruction method on pre-service teachers' conceptual comprehension of methodology course. *Journal of Education and Learning, 6*(3), 70-82. https://doi.org/10.5539/jel.v6n3p70

Alvermann, D. E., Hynd, C. E., & Qian, G. (1995). Effects of interactive discussion and text type on learning counterintuitive science concepts. *The Journal of Educational Research, 88*(3), 146-154. https://doi.org/10.1080/00220671.1995.9941293

Balta, N., & Asikainen, M. A. (2019). A comparison of Olympians' and regular students' approaches and successes in solving counterintuitive dynamics problems. *International Journal of Science Education, 41*(12), 1644-1666. https://doi.org/10.1080/09500693.2019.1624990

Balta, N., & Eryılmaz, A. (2017). Counterintuitive dynamics test. *International Journal of Science and Mathematics Education, 15*, 411-431. https://doi.org/10.1007/s10763-015-9694-6

Balta, N., & Güversin, S. (2016). Increasing undergraduate students' exam performances in statistics course using software Socrative. *The Turkish Online Journal of Educational Technology, Special Issue*, 314-321.

Balta, N., Japashov, N., Abdulbakioglu, M., & Oliveira, A. W. (2019). High-school students' cognitive responses to counterintuitive physics problems. *Physics Education, 55*(1), 015003. https://doi.org/10.1088/1361-6552/ab4df9

Balta, N., Kaymak, S., Almas, A, & Nurbavliyev, O. (2021). The impact of peer instruction on ninth grade students' trigonometry knowledge. *Bolema: Boletim de Educação Matemática, 35*, 206-222. https://doi.org/10.1590/1980-4415v35n69a10

Balta, N., & Moğol, S. (2016). Sezgiye ters fizik soruları ve bu soruların öğrenci başarısına etkisi [Counterintuitive physics questions and their impact on student success]. *Pegem Eğitim ve Öğretim Dergisi, 6*(1), 133-146. https://doi.org/10.14527/pegegog.2016.008

Bian, H., Bian, Y., Li, J., Li, Y., Ma, Y., Shao, X., & Xu, J. (2018). Peer instruction in a physiology laboratory course in China. *Advances in Physiology Education, 42*(3), 449-453. https://doi.org/10.1152/advan.00153.2017

Brooks, B. J., & Koretsky, M. D. (2011). The influence of group discussion on students' responses and confidence during peer instruction. *Journal of Chemical Education, 88*(11), 1477-1484. https://doi.org/10.1021/ed101066x

Campanario, J. M. (1998). Using counterintuitive problems in teaching physics. *The Physics Teacher, 36*(7), 439-441. https://doi.org/10.1119/1.879917

Carter, G., Jones, M. G., & Rua, M. (2003). Effects of partner's ability on the achievement and conceptual organization of high-achieving fifth-grade students. *Science Education, 87*(1), 94-111. https://doi.org/10.1002/sce.10031

Chien, Y. T., Lee, Y. H., Li, T. Y., & Chang, C. Y. (2015). Examining the effects of displaying clicker voting results on high school students' voting behaviors, discussion processes, and learning outcomes. *EURASIA Journal of Mathematics, Science and Technology Education, 11*(5), 1089-1104. https://doi.org/10.12973/eurasia.2015.1414a

Cortright, R. N., Collins, H. L., & DiCarlo, S. E. (2005). Peer instruction enhanced meaningful learning: ability to solve novel problems. *Advances in Physiology Education, 29*(2), 107-111. https://doi.org/10.1152/advan.00060.2004

Crouch, C. H., & Mazur, E. (2001). Peer instruction: Ten years of experience and results. *American Journal of Physics, 69*(9), 970-977. https://doi.org/10.1119/1.1374249

Cummings, K., & Roberts, S. G. (2008). A study of peer instruction methods with high school physics students. *AIP Conference Proceedings, 1064*(1), 103-106. https://doi.org/10.1063/1.3021227

Dancy, M., Henderson, C., & Turpen, C. (2016). How faculty learn about and implement research-based instructional strategies: The case of peer instruction. *Physical Review Physics Education Research, 12*(1), Article 010110. https://doi.org/10.1103/PhysRevPhysEducRes.12.010110

DeSoto, K. A., & Roediger III, H. L. (2014). Positive and negative correlations between confidence and accuracy for the same events in recognition of categorized lists. *Psychological Science, 25*(3), 781-788. https://doi.org/10.1177/0956797613516149

Elizabeth, O., Mutsotso, A. N., & Masibo, E. N. (2016). Effect of peer teaching among students on their performance in mathematics. *International Journal of Scientific Research and Innovative Technology, 3*(12), 10-24.

El-Sayed, S. H. (2013). Effect of peer teaching on the performance of undergraduate nursing students enrolled in nursing administration course. *Journal of Nursing Education and Practice, 3*(9), 156-166

Fagen, A. P., Crouch, C. H., & Mazur, E. (2002). Peer instruction: Results from a range of classrooms. *The Physics Teacher, 40*(4), 206-209. https://doi.org/10.1119/1.1474140

Fall, R., Webb, N. M., & Chudowsky, N. (2000). Group discussion and large-scale language arts assessment: Effects on students' comprehension. *American Educational Research Journal, 37*(4), 911-941. https://doi.org/10.3102/00028312037004911

Giuliodori, M. J., Lujan, H. L., & DiCarlo, S. E. (2006). Peer instruction enhanced student performance on qualitative problem-solving questions. *Advances in Physiology Education, 30*(4), 168-173. https://doi.org/10.1152/advan.00013.2006

Gok, T. (2012a). The effects of peer instruction on students' conceptual learning and motivation. *Asia-Pacific Forum on Science Learning and Teaching, 13*(1), Article 10.

Gok, T. (2012b). 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. https://doi.org/10.1007/s10763-011-9316-x

Gok, T. (2013). A comparison of students' performance, skill and confidence with peer instruction and formal education. *Journal of Baltic Science Education, 12*(6), 747-758. https://doi.org/10.33225/jbse/13.12.747

Guzzetti, B. J. (2000). Learning counter-intuitive science concepts: What have we learned from over a decade of research? *Reading & Writing Quarterly, 16*(2), 89-98. https://doi.org/10.1080/105735600277971

Lasry, N., Mazur, E., & Watkins, J. (2008). Peer instruction: From Harvard to the two-year college. *American journal of Physics, 76*(11), 1066-1069. https://doi.org/10.1119/1.2978182

Mallow, J. V. (1995). Students' confidence and teachers' styles: A binational comparison. *American Journal of Physics, 63*(11), 1007-1011. https://doi.org/10.1119/1.18046

Mazur, E. (1997). Peer instruction: Getting students to think in class. *AIP Conference Proceedings, 399*(1), 981-988. https://doi.org/10.1063/1.53199

Miller, K., Schell, J., Ho, A., Lukoff, B., & Mazur, E. (2015). Response switching and self-efficacy in Peer Instruction classrooms. *Physical Review Special Topics-Physics Education Research, 11*(1), Article 010104. https://doi.org/10.1103/PhysRevSTPER.11.010104

Parsons, S., Croft, T., & Harrison, M. (2009). Does students' confidence in their ability in mathematics matter? *Teaching Mathematics and Its Applications: An International Journal of the IMA, 28*(2), 53-68. https://doi.org/10.1093/teamat/hrp010

Pilzer, S. (2001). Peer instruction in physics and mathematics. *PRIMUS, 11*(2), 185-192. https://doi.org/10.1080/10511970108965987

Porter, L., Bailey Lee, C., Simon, B., & Zingaro, D. (2011). Peer instruction: Do students really learn from peer discussion in computing? In *Proceedings of the 7th International Workshop on Computing Education Research* (pp. 45-52). https://doi.org/10.1145/2016911.2016923

Porter, L., Bouvier, D., Cutts, Q., Grissom, S., Lee, C., McCartney, R., Zingaro, D., & Simon, B. (2016). A multi-institutional study of peer instruction in introductory computing. In *Proceedings of the 47th ACM Technical Symposium on Computing Science Education* (pp. 358-363). ACM. https://doi.org/10.1145/2839509.2844642

Porter, L., Garcia, S., Glick, J., Matusiewicz, A., & Taylor, C. (2013). Peer instruction in computer science at small liberal arts colleges. In *Proceedings of the 18th ACM Conference on Innovation and Technology in Computer Science Education* (pp. 129-134). ACM. https://doi.org/10.1145/2462476.2465587

Pratama, M. A. (2018). Factors affecting students' confidence in public speaking. *Journal of Languages and Language Teaching, 5*(2), 67-70. https://doi.org/10.33394/jollt.v5i2.357

Royce, C. S., Hayes, M. M., & Schwartzstein, R. M. (2019). Teaching critical thinking: a case for instruction in cognitive biases to reduce diagnostic errors and improve patient safety. *Academic Medicine, 94*(2), 187-194. https://doi.org/10.1097/ACM.0000000000002518

Simon, B., Parris, J., & Spacco, J. (2013). How we teach impacts student learning: Peer instruction vs. lecture in CS0. In *Proceeding of the 44th ACM Technical Symposium on Computer Science Education* (pp. 41-46). ACM. https://doi.org/10.1145/2445196.2445215

Smith, M. K., Wood, W. B., Adams, W. K., Wieman, C., Knight, J. K., Guild, N., & Su, T. T. (2009). Why peer discussion improves student performance on in-class concept questions. *Science, 323*, 122-124. https://doi.org/10.1126/science.1165919

Tullis, J. G., & Goldstone, R. L. (2020). Why does peer instruction benefit student learning? *Cognitive Research: Principles and Implications, 5*, 1-12. https://doi.org/10.1186/s41235-020-00218-5

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), Article 020101. https://doi.org/10.1103/PhysRevSTPER.5.020101

Van Dooren, W., De Bock, D., Janssens, D., & Verschaffel, L. (2008). The linear imperative: An inventory and conceptual analysis of students' overuse of linearity. *Journal for Research in Mathematics Education, 39*(3), 311-342.

Vickrey, T., Rosploch, K., Rahmanian, R., Pilarz, M., & Stains, M. (2015). based implementation of peer instruction: A literature review. *CBE–Life Sciences Education, 14*(1), Article es3. https://doi.org/10.1187/cbe.14-11-0198

Zhang, P., Ding, L., & Mazur, E. (2017). Peer instruction in introductory physics: A method to bring about positive changes in students' attitudes and beliefs. *Physical Review Physics Education Research, 13*(1), Article 010104. https://doi.org/10.1103/PhysRevPhysEducRes.13.010104

Zingaro, D. (2014). Peer instruction contributes to self-efficacy in CS1. In *Proceedings of the 45th ACM Technical Symposium on Computer Science Education* (pp. 373-378). ACM. https://doi.org/10.1145/2538862.2538878