A Multi-Institutional Study on Peer Instruction: Evaluating Text-Chat with Assigned Group Members vs Verbal Discussion

Xingjian (Lance) Gu
xjgu@umich.edu
University of Michigan
Ann Arbor, MI, USA

Barbara J. Ericson
barbarer@umich.edu
University of Michigan
Ann Arbor, MI, United States

Zihan Wu
zihan.wu@maine.edu
University of Maine
Orono, ME, USA

Margaret Ellis
maellis1@vt.edu
Virginia Tech
Blacksburg, VA, USA

Janice L. Pearce
jan_pearce@berea.edu
Berea College
Berea, KY, USA

Susan Rodger
rodger@cs.duke.edu
Duke University
Durham, NC, USA

Yesenia Velasco
yvelasco@cs.duke.edu
Duke University
Durham, NC, USA

Abstract

In Peer Instruction (PI) an instructor displays a challenging multiple-choice question during lecture that students answer individually, discuss verbally with nearby peers, answer individually again, and finally, the instructor leads a discussion of the question. Peer Instruction typically increases student learning and motivation over traditional lecture. We added a text-chat mode to improve PI for remote synchronous learning. This feature assigns students to discussion groups to maximize the number of groups that have members with different answers. The tool was pilot tested in Winter 2022 and revised. In Fall 2022 and Winter 2023, it was tested at one institution. In Fall 2024, it was tested at four institutions. We conducted a log file analysis of student data from 1394 students and analyzed surveys with 848 student responses. We found that questions answered using the text-chat had a significantly higher improvement than those using traditional verbal discussion, although the two modes were tested with different questions. Interestingly, most of the students preferred to discuss the question verbally, although some preferred the text-chat discussion. These results inform efforts to improve the effectiveness of Peer Instruction and increase its adoption.

CCS Concepts

• **Applied computing** → *Interactive learning environments.*

Keywords

Peer Instruction, interactive ebooks, multi-institutional

ACM Reference Format:
Xingjian (Lance) Gu, Barbara J. Ericson, Zihan Wu, Margaret Ellis, Janice L. Pearce, Susan Rodger, and Yesenia Velasco. 2026. A Multi-Institutional Study on Peer Instruction: Evaluating Text-Chat with Assigned Group Members vs Verbal Discussion. In *Proceedings of the 57th ACM Technical Symposium on Computer Science Education V.1 (SIGCSE TS 2026), February 18–21, 2026, St. Louis, MO, USA.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3770762.3772563

SIGCSE TS 2026, St. Louis, MO, USA

ACM ISBN 979-8-4007-2256-1/2026/02
https://doi.org/10.1145/3770762.3772563

1 Introduction

Peer Instruction (PI), with capitalized initial letters, refers to a specific instructional technique originally developed by Eric Mazur to improve student engagement and learning in physics [9]. In Peer Instruction, during lecture, the instructor displays a challenging multiple-choice question. The students answer the question individually (the first vote), then verbally discuss their answers with nearby students, and then answer again (the second vote). The instructor then leads a discussion of the question and may show the results of the two votes [9].

Prior studies have shown Peer Instruction to be highly effective at increasing student engagement and learning in several fields, including physics [9, 10], biology [17], and computer science [20, 27]. It leads to twice the learning gains in physics over traditional passive lecture as measured by the Force Concept Inventory [9, 16]. In computer science, it has substantially reduced failure rates [19], improved retention [24], and increased final exam performance [27]. This is true for both introductory computing courses [3] as well as advanced computing courses [18].

One limitation of traditional Peer Instruction is that it offers little control over group assignment, potentially leading to less effective student discussions. The most effective discussions during Peer Instruction happen between peers with different answers. Students with different answers can explain to each other their reasoning, help correct misconceptions, and reinforce new concepts, leading to improved understanding even before the instructor's explanation [21, 23, 28]. However, when students discuss with nearby peers, they may have the same answer, which leads to less productive discussions. While there has been extensive research on Peer Instruction, it has mostly been conducted during in-person lectures with students verbally discussing the question with nearby peers

[9]. We also do not know of any empirical evidence on how learning changes when students are assigned to discussion groups that maximize the number of groups where members hold different answers.

To address these issues, we designed and implemented Peer+: a learning tool that facilitates Peer Instruction activities and supports both verbal discussion with nearby peers and text-chat discussion with assigned group members. Students are assigned to groups to maximize the number of groups with students who voted for different answers. We conducted a multi-institutional study to investigate the following research questions:

RQ1: What is the effect of text-chat peer discussion with assigned group members on learning compared to verbal peer discussion with nearby peers?

RQ2: Which discussion mode of Peer Instruction (text-chat with assigned groups or verbal with nearby peers) do students find more engaging and effective, and why?

To answer these questions, we tested the tool and collected both log file data of student usage and qualitative student feedback over several years, and with several institutions. In Winter 2022, we pilot tested Peer+ in a large undergraduate intermediate Python programming course, then fully deployed the system and iteratively modified the tool during Fall 2022 and Winter 2023 semesters. In Fall 2024, the tool was implemented in programming courses at four higher education institutions. We evaluated RQ1 with a quantitative log file analysis of Fall 2022 ($n = 147$), Winter 2023 ($n = 191$), and Fall 2024 ($n = 225$ in institution 1, $n = 199$ in institution 2, $n = 597$ in institution 3, $n = 35$ in institution 4). We addressed RQ2 through student surveys conducted in each of the four institutions in Fall 2024 ($n = 157$ in institution 1, $n = 140$ in institution 2, $n = 521$ in institution 3, $n = 30$ in institution 4).

This research contributes evidence that assigning students to text-chat discussion groups to maximize the number of groups where members hold different answers significantly improves learning compared to in-lecture verbal discussion with nearby peers. This work also sheds light on students' perceptions of the effectiveness of different modes of Peer Instruction. In addition, it provides evidence across multiple institutions that include various class sizes and courses. The findings can help us improve the implementation and effectiveness of Peer Instruction.

2 Literature Review

2.1 Theoretical Foundation of Peer Instruction

The success of Peer Instruction is consistent with the theory of social constructivism by Vygotsky [31]. He agreed with Piaget [32] that learners have to actively construct their own knowledge, but specifically that knowledge is co-constructed in the learner's social context [31]. He proposed that learning can be maximized when a learner is in the *Zone of Proximal Development* (ZPD): when the learner is challenged to do more than they currently can alone, but can achieve by interacting with more knowledgeable others [31]. Peer Instruction activities are designed to have students attempt challenging questions, then discuss their answers with peers, thus situating students in the ZPD and supporting their learning. This study explores whether assigning students to discussion groups to maximize the number of groups where members hold different answers can achieve better learning outcomes than verbal discussion with nearby peers.

Peer Instruction's effectiveness is also consistent with Chi's ICAP theory, which describes four modes of learner engagement: Interactive, Constructive, Active, and Passive [6, 7]. The theory posits that interactive engagement results in the greatest learning, then constructive, then active, and finally passive [7]. An example of passive mode is a typical lecture where the learner is not taking any notes. An example of active mode is solving a problem. An example of constructive mode is a learner writing notes in their own words during lecture. Peer Instruction, by facilitating discussion between learners, is an example of the interactive mode.

2.2 Effectiveness of Peer Instruction

Decades of empirical research on Peer Instruction provide evidence that it typically improves student learning [9]. The percentage of students who get the question correct usually improves between the first vote and the second vote [15, 23, 26, 34]. There is evidence that more difficult questions lead to more learning gains than easier questions [35]. Crouch and Mazur recommend that Peer Instruction questions be challenging, with ideally between 30% to 70% of the students choosing correctly on the first vote [9, 10]. In addition, 74% to over 90% of students typically agree that Peer Instruction is valuable for learning and other instructors should use it [18, 20, 22]. Peer instruction works well, even in large lectures [20, 22]. Even though PI is often implemented in lower-division courses, prior work has shown that it is also effective in upper-division courses, including computer architecture and cybersecurity [11, 19].

While there is strong evidence for the benefits of Peer Instruction, it is mostly from studies with in-person verbal discussion without any group assignment. One study reported an increased student participation rate when students are randomly assigned seating in lecture each week, compared to no group assignment [8]. However, this study did not assign students to groups with peers with different answers, nor compare student learning in assigned groups to unassigned groups. A pair of studies affirmed that remote synchronous PI without group assignment, where students discussed in Zoom breakout rooms, has a comparable effect on learning to in-person discussions [13, 14]. There are also studies that developed tools that support PI, but without testing their effects on student learning in synchronous classroom contexts [2, 30].

3 Tool Design and Implementation

We used a design-based research approach [1] to design and implement a free tool to support Peer Instruction and integrated it into a free e-book platform Runestone [12]. Design-based research creates interventions based on theory, tests them in real educational environments, revises the interventions based on the tests, and contributes to theory [33]. For this study, we pilot-tested the tool in Winter 2022 initially and fully deployed it in institution 1 in Fall 2022. The tool was subsequently updated each semester, with changes to the user interface and fixes for technical problems. The core designs relevant to this study, including text-chat group assignment and Peer Instruction procedures, remained consistent over the study period.

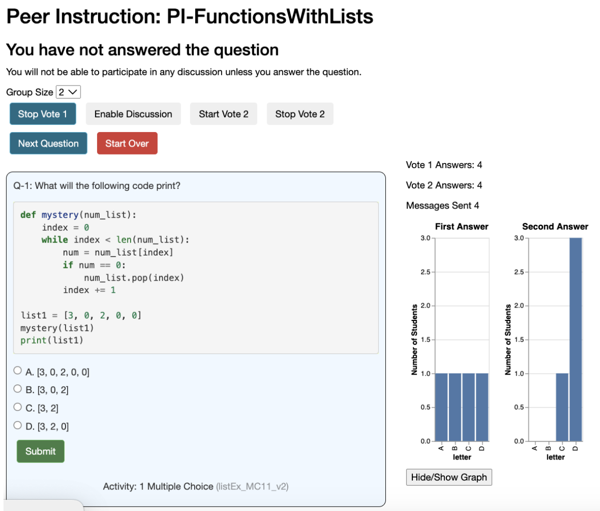

Figure 1: The Instructor PI Session Interface

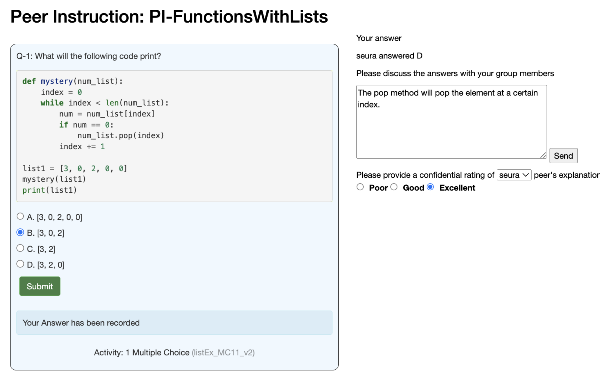

Figure 2: The In-Lecture Student Text-Chat Mode

3.1 Creating a Peer Instruction Assignment

Most of the Peer Instruction questions were created by the course instructors. We also added over a thousand questions to free e-books from a public question bank of Peer Instruction questions for computer science (https://peerinstruction4cs.com/). Instructors can search for questions by topic and/or author and add questions to an assignment. The assignment interface has a check mark for making an assignment a PI assignment. The assignment interface also allows an instructor to make an assignment visible to students.

3.2 In-Lecture Peer Instruction Procedure

The design of the instructor session interface is intended to facilitate the proper PI procedure [9]. The recommended procedure is that students first vote individually, then discuss with peers, and then vote again. Finally, the instructor may show the histograms from the two votes and lead a discussion of the question.

As soon as the instructor starts the PI session, the first question is displayed, and the first vote begins. The number of votes is constantly updated on the instructor interface. When the number of students who voted is close to what is expected, the instructor announces that they will end voting and click the 'Stop Vote 1' button, as shown in Figure 1. Students receive a five-second countdown before voting ends.

The instructor can then either instruct the students to discuss the question verbally with nearby peers or click the 'Enable Discussion' button to enable the text-chat interface, as shown in Figure 2. In the text-chat mode, the default discussion group size was two during both the Fall 2022 and Winter 2023 semesters. In Fall 2024, the default group size was increased to three, since we had received student feedback that sometimes the other person in the two-person text-chat group did not respond or did not engage much in the discussion. The system assigned students to groups in order to maximize the number of groups that had members with different answers.

The instructor interface displays the number of messages sent during text-chat discussion mode. When the number of messages indicates most students have discussed, or when verbal discussions quiet down, the instructor can start the second vote by clicking the 'Start Vote 2' button. The instructor concludes the voting by clicking the 'Stop Vote 2' button, which also displays a five-second countdown. The instructor can then display the histograms of the first and second vote results and lead a discussion about the question.

3.3 Course Contexts and Materials

To address the research questions, academics from four US higher education institutions collaborated to each implement Peer Instruction activities and Peer+ in their courses and collect data. Each course did PI with both discussion modes during the semesters.

We first tested the system in a programming course in institution 1: a large public institution. The course is required for Information majors, but students from other majors take the course as well. The course uses Python and covers programming fundamentals, data structures (lists, dictionaries, and tuples), object-oriented programming, unit tests, debugging, regular expressions, and working with data from files, HTML pages, web servers through APIs, and databases. There were 147 students enrolled in Fall 2022, 191 students in Winter 2023, and 225 students in Fall 2024. Before the creation of Peer+, PI was conducted with iClicker devices and instructor-created PI questions for eight semesters. Those PI questions were added to a free and interactive e-book, which is used in the course for readings, practice, and midterms.

The system was tested at three more institutions starting in Fall 2024. Institution 2 is a medium-sized private research university. PI and Peer+ were tested in an introductory CS1 course that serves both prospective majors and non-majors with an enrollment of 199 students. The course uses Python and emphasizes computational thinking, lists, sets, dictionaries, tuples, and debugging. The students participated in both verbal and text-chat discussions using the tool. Before adopting the tool, the instructional team had used Google Forms to facilitate interactive in-class questions for over

Institution	Semester	Verbal			Text-Chat			t	p
		n	$\overline{x}$	σ	n	$\overline{x}$	σ		
Institution 1	Fall 2022	19	0.32	0.32	31	0.54	0.37	-2.14	0.038*
Institution 1	Spring 2023	31	0.40	0.03	19	0.59	0.06	-2.95	0.006**
Institution 1	Fall 2024	62	0.35	0.26	15	0.52	0.28	-2.16	0.042*
Institution 2	Fall 2024	19	0.39	0.17	12	0.56	0.17	-2.63	0.014*
Institution 3	Fall 2024	32	0.36	0.19	8	0.51	0.14	-2.66	0.018*

*p < .05, **p < .01., ***p < .001

Table 1: Welch's *t*-Test for NC of Verbal Discussion versus Text-Chat. Institution 4 had nine PI questions in total. It is not included in the statistical analysis due to the small sample size.

ten years. In some semesters, the team used duplicate forms to capture both individual (first vote) and post-discussion (second vote) responses. While effective for basic data collection, this approach lacked automated group assignment. The team added the questions from the Google Forms to an ebook for use with Peer+.

Institution 3 is a large public institution. PI and Peer+ were used in a CS2 course taught in Java. Peer+ was deployed in Fall 2024 across 4 sections with 3 different instructors teaching a total of 597 students. This CS2 course is an introduction to software design and data structures, covering topics such as polymorphism, UML diagrams, recursion, and various implementations of queues, stacks, lists, and binary trees. The course includes CS majors and CS minors, but also many incoming first-year students and others in computing-related majors such as Data Analytics, Mechanical Engineering, and FinTech. This instructional team uses an e-book with pre-made lecture videos. Before implementing Peer Instruction, students had weekly multiple-choice questions to complete independently. Variations of these multiple-choice questions were added to Peer+ and made publicly available and searchable.

Institution 4 is a small private liberal arts college. PI and Peer+ were used by a single instructor in two sections of a CS2 Data Structures course taught in C++. There were 17 students in one section and 18 in the other. The course is taken by CS majors, CS minors, other STEM majors, and Business Administration majors. The course covers algorithm analysis early and places a heavy emphasis on computational complexity throughout the standard data structure coverage of stacks, queues, linked lists, recursion, searching, hashing, trees, and graphs. In addition to mini-lectures, the course pedagogy leverages various types of other team-based learning techniques in which students are assigned to small randomized groups of 3-4 students that are rotated every couple of weeks. These same groups were used for the in-person PI groups.

Institutional Review Board (IRB) approval was sought and granted for each of the participating institutions. Students were asked to provide consent to be surveyed. We had a separate IRB, which allowed us to analyze the anonymous log file data.

4 Student Log File Analysis

To address RQ1, we compared the effects of text-chat discussions versus verbal discussions on student learning. Peer+ records students' answers for both vote 1 and vote 2. We analyzed the log file data from institution 1 in Fall 2022 (n=147) and Winter 2023 (n=191), and across all four institutions in Fall 2024 (n=225 in institution 1, n=199 in institution 2, n=597 in institution 3, n=35 in institution 4). We analyzed the data from each semester and each institution independently since the learning contexts are different. Some of the PI questions were changed between semesters, and the tool was updated over the duration of the study, but the structure of the Peer Instruction activities and the group assignment methods remained unchanged.

4.1 Normalized Change

To measure the extent of student learning after doing PI, we calculated the normalized change (NC). It measures the amount of learning as a fraction of available learning [29] and is commonly used in Peer Instruction research [25, 35]. It applies the following formula:

$$NC = \frac{V_2CorrectRate - V_1CorrectRate}{1 - V_1CorrectRate}$$

For example, if 50% of the students correctly answer in the first vote and 80% of the students correctly answer in the second vote, this is a 30% learning gain, but NC divides this 30% by the percent who could improve on the second vote. That is, $(80\% - 50\%)/(1 - 50\%)$ is 60%, so that question would have a 60% normalized change.

4.2 Analysis of Student Log Data

To determine if there was a significant difference in learning from verbal and text-chat discussions, we conducted a Welch's t-test on the average normalized change. This choice was due to the number of verbal and text-chat discussion questions used at each of the institutions being unequal. The normality assumption was checked using the Shapiro-Wilk normality test and was satisfied for each dataset. Since we tested the text-chat mode and verbal discussion mode on different PI questions, we also need to ensure that the average difficulties of the questions between the two modes were not significantly different. We conducted two sample t-tests to compare the percent correct in the first vote between the text-chat questions and verbal discussion questions. None of the datasets had statistically significant differences between the percent correct in the first vote, which indicates the groups were comparable.

Table 1 shows the normalized change for in-lecture verbal discussion questions and text-chat questions, as well as the t-test results. The results show that the learners had statistically significantly higher normalized change on the questions that were discussed with text-chat than with verbal discussion with nearby peers across all institutions and semesters. Note that the data from institution 4

Survey question	Institution 1 (n = 157)	Institution 2 (n = 140)	Institution 3 (n = 521)	Institution 4 (n = 30)
Prefer verbal discussion over text-chat	86%	56%	80%	97%
Find the Peer Instruction questions too easy	7%	6%	16%	13%
Verbal discussion is valuable for my learning	80%	72%	75%	97%
Text-chat is valuable for my learning	29%	49%	39%	27%
Actually discussed the PI question most of the time during verbal discussion	75%	72%	76%	87%
Actually discussed the PI question most of the time during text-chat	38%	52%	45%	50%
PI questions helped me pay attention compared to traditional lectures	63%	63%	61%	70%
I recommend other instructors use PI	60%	57%	59%	80%
Having someone in the group who knows the correct answer is necessary in order to make the discussion productive.	56%	45%	48%	60%

Table 2: The Percentage of Students who Chose ‘Agree’ or ‘Strongly Agree’ for Each Statement on a Five-Point Likert Scale

is not included in the statistical analysis, due to the sample sizes being too small (7 verbal discussion questions and 2 text-chat discussions). Overall, the results provide evidence that text-chat PI discussions with assigned peers to maximize the number of peers with different answers is more effective than traditional PI verbal discussions.

5 Student Survey

To address RQ2, we conducted student surveys on their experiences with Peer Instruction near the end of the semester. The survey included questions from a prior student attitudinal survey about Peer Instruction [18, 21]. Students were asked to answer the questions on a five-point Likert scale, with "neutral" as an option. To understand which mode of Peer Instruction students prefer, the survey additionally included the question, *"Which type of discussion do you prefer?"* Lastly, we also included short-answer questions in the survey asking for the perceived advantages and disadvantages of both the in-lecture verbal discussion and text-chat mode. The survey was distributed across the four institutions. All students were instructed to complete the survey as part of their class activities, but only the data from students who consented to have their survey responses be included in the research were included in this study. In total, 848 students completed the surveys and consented to participate. A link to the student survey questions will be provided upon request.

5.1 Quantitative Survey Results

In general, most students were positive about both Peer Instruction and the free tool across all four institutions, as shown in Table 2. Between 61% (at institution 3) and 70% (institution 4) of the students agreed or strongly agreed to the statement that Peer Instruction questions helped them pay attention in the course. Between 57% and 80% of the students would recommend that other instructors use PI. At most 16% of the students found the PI questions too easy, which is also an encouraging outcome, since PI questions are more effective if they are challenging.

When asked *"I prefer to talk to the people around me about a Peer Instruction question rather than use the chat in the e-book"*, more than 80% of the students agreed at institutions 1, 3, and 4, while 56% of the students agreed at institution 2. We asked more specifically whether the students believed that verbal discussion and text-chat are valuable for their learning. Across the institutions, students generally agree (72%-97%) that verbal discussion is valuable for their learning, while relatively fewer feel that way about the text-chat (27%-49%). Students also reported that their discussions were more on-topic during verbal discussion (72%-87%) than during the text-chat (38%-52%).

These findings are interesting because text-chat discussions are associated with greater normalized change in learning gains, but students perceived the text-chat as less valuable for their learning. There are also differences in responses between the institutions. Students from institutions 2 and 3 were less positive about the effectiveness of the verbal discussions and more positive about the effectiveness of the text-chat, compared to students from institutions 1 and 4. We hypothesize that this could be due to class sizes and classroom layouts, as students from Institutions 1 and 4 had smaller class sizes and sat around tables to form groups, while students from Institutions 2 and 3 attended larger classes in traditional lecture halls.

5.2 Thematic Analysis of Short Answers

5.2.1 Analysis. To gain further insight on the reasoning behind student preferences, we conducted coding reliability thematic analysis on the short answers [5]. The analysis used a codebook developed for semi-structured interviews with students’ experiences using Peer+ from a parallel study. One author developed the original codebook by open-coding student interviews to generate the initial codes. The author then iterated the codebook with a graduate student research assistant by each independently coding the interviews and routinely meeting to discuss code definitions and resolve disagreements. The themes captured the student’s reasons for preferring one of the discussion modes, including “ease of articulating thoughts”, “embarrassment of making mistakes”, and “having in-depth discussions”. They reached an inter-rater agreement of 0.826 (Krippendorff’s alpha) for all codes by testing across a quarter of the interview dataset, which indicates moderate agreement and is adequate for analysis. Since the survey question about the student’s discussion-mode preference is the same as the interview question, the codebook was adapted to analyze the survey responses.

5.2.2 Findings. The most prominent reasons students gave for preferring the verbal discussion over the text-chat is that it was easier to communicate their thoughts (n=285). Students find it easier to physically point at code and explain their thoughts verbally, compared to typing out an explanation. As one student put it: *"I'm slow at typing so it's hard for me to put my thoughts in a short amount of time."* This reason was most often cited at institution 4, where the classroom layout facilitates small group discussions.

Students had mixed responses on which mode provided better social interaction. With 171 students preferring the social aspects of verbal discussions, since they can talk with familiar peers (n=102), see face-to-face social cues (n=43), and be more engaged (n=26). In contrast, 106 students preferred the social interactions during a text-chat, because it was less awkward if they did not know their peers (n=53) and the anonymity of the text-chat makes them less embarrassed to make mistakes (n=52). This shows that there is value in providing students with multiple modalities to interact with their peers. This is especially the case in larger classes, as students at institution 2 mentioned the benefit of text-chat anonymity the most.

Relatively more students perceived verbal discussions as more conducive to learning than text-chat. In addition to the ease of communication, 92 students also find their verbal discussions to be more in-depth. Additionally, 116 students mentioned that during text-chat, some peers would not contribute much to the discussion or not say anything at all. Reports of a similar lack of participation during verbal discussion are much rarer (n=8). On the other hand, 54 students said the text-chat allows them to discuss with peers who they don't usually interact with, which could expose them to more diverse ways of thinking. One student noted that *"honestly in person chat a lot of the time it would be my friends so we'd all have chosen the same thing already or they weren't paying attention."*

6 Discussion

6.1 Effects of Text-Chat on Learning

The multi-institutional log file analysis found that the text-chat mode that assigns students to maximize the number of groups with members with different answers had a statistically significantly higher normalized change than verbal discussions with nearby peers. We also found evidence to support this hypothesis in the student surveys. Some students mentioned that it was more likely that the group had the same answers during verbal discussions and that they learned more when the group had different answers, since they were exposed to other perspectives and ways of thinking. This finding aligns with Vygotsky's theory of social constructivism [31] in that it predicts students will learn more by interacting with more knowledgeable others.

6.2 Student Preference of Discussion Modes

The student survey results provided evidence that most students at the four institutions prefer verbal discussion with peers versus the text-chat. However, this ranged from 56% preferring verbal at institution 2 (larger class size, lecture hall) to 97% at institution 4 (smaller class size, students seated around tables). Reasons for preferring the verbal discussion to the text-chat included: it is easier to communicate verbally, more social, and results in more in-depth discussions. Reasons for preferring the text-chat include feeling less embarrassed about making mistakes, feeling less awkward, and being exposed to different answers and approaches. One common complaint was that group members did not respond in the text-chat. However, this may have been due to using a default group size of two at institution 1 during fall 2022 and winter 2023.

The students' preference for verbal discussion, despite the evidence suggesting text-chat discussion can be more effective, has important implications. Since negative feedback from students is a major instructor concern [4], improving the user experience is critical. The current chat interface is relatively basic, relying solely on text to communicate. This can cause students to become less engaged in discussion. Future work could explore a verbal chat with assigned group members and a shared annotation space to improve communication.

7 Limitations and Future Work

Although we found evidence that text-chat led to a significantly higher normalized change than the verbal discussion, we did not test the two modes using an A/B test study design. We plan to do this in winter 2026. Additionally, despite being multi-institutional, this study only researched learning contexts within the US. The discussion modes of PI included in this study might be received differently across different learning environments. Future work could test PI modalities in institutions outside the US.

8 Conclusion

Decades of research on Peer Instruction have shown that it improves student engagement, learning, and retention. However, most of this research used the traditional in-person verbal discussion with nearby peers. To possibly improve learning outcomes and support remote synchronous lecture, we created a new tool that supports two discussion modalities for Peer Instruction questions: verbal discussion with nearby peers, and a text-chat mode that assigns students to groups in order to maximize the number of groups that have members with different answers. We found that most students value Peer Instruction, which is consistent with previous work. Most students preferred the verbal discussion to the text-chat. However, the text-chat led to significantly higher normalized change than the verbal discussion with nearby peers. The text-chat also has the advantage that it can be used even when students are remote. Since most students preferred verbal discussion with nearby peers, we plan to test a virtual verbal chat room with assigned groups and a shared annotation space. Our goal is to encourage more instructors to use Peer Instruction during lecture by making the tool freely available and by adding unique features that provide instructors and students with more options.

Acknowledgments

This material is based upon work supported by the National Science Foundation under Awards 2043207 and 2336755. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. We also thank Denys Zhytkov, Shefali Patel, Ian Kennedy, Alexander Cetinel, and Chretien Nishimwe for their contributions.

References

[1] Sasha Barab and Kurt Squire. 2004. Design-based research: Putting a stake in the ground. *The journal of the learning sciences* 13, 1 (2004), 1–14.

[2] Sameer Bhatnagar, Nathaniel Lasry, Michel Desmarais, and Elizabeth Charles. 2016. DALITE: Asynchronous Peer Instruction for MOOCs. In *Adaptive and Adaptable Learning (Lecture Notes in Computer Science)*, Katrien Verbert, Mike Sharples, and Tomaž Klobučar (Eds.). Springer International Publishing, Cham, 505–508. doi:10.1007/978-3-319-45153-4_50

[3] Dennis Bouvier, Ellie Lovellette, John Matta, Jing Bai, Jacqueline Chetty, Stan Kurkovsky, and Jia Wan. 2019. Factors Affecting the Adoption of Peer Instruction in Computing Courses. In *Proceedings of the Working Group Reports on Global Computing Education*. 1–25.

[4] Dennis Bouvier, Ellie Lovellette, John Matta, Jing Bai, Jacqueline Chetty, Stan Kurkovsky, and Jia Wan. 2019. Factors Affecting the Adoption of Peer Instruction in Computing Courses. In *Proceedings of the Working Group Reports on Global Computing Education (CompEd-WGR '19)*. Association for Computing Machinery, New York, NY, USA, 1–25. doi:10.1145/3372262.3375396

[5] Virginia Braun and Victoria Clarke. 2023. Toward good practice in thematic analysis: Avoiding common problems and be(com)ing a knowing researcher. *International Journal of Transgender Health* 24, 1 (Jan. 2023), 1–6. doi:10.1080/26895269.2022.2129597 Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/26895269.2022.2129597.

[6] Michelene TH Chi. 2021. Translating a theory of active learning: An attempt to close the research-practice gap in education. *Topics in Cognitive Science* 13, 3 (2021), 441–463.

[7] Michelene TH Chi and Ruth Wylie. 2014. The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational psychologist* 49, 4 (2014), 219–243.

[8] Chih-Yueh Chou and Pin-Hsun Lin. 2015. Promoting discussion in peer instruction: Discussion partner assignment and accountability scoring mechanisms. *British Journal of Educational Technology* 46, 4 (July 2015), 839–847. doi:10.1111/bjet.12178 Publisher: John Wiley & Sons, Ltd.

[9] Catherine H Crouch and Eric Mazur. 2001. Peer instruction: Ten years of experience and results. *American journal of physics* 69, 9 (2001), 970–977.

[10] Catherine H Crouch, Jessica Watkins, Adam P Fagen, and Eric Mazur. 2007. Peer instruction: Engaging students one-on-one, all at once. *Research-based reform of university physics* 1, 1 (2007), 40–95.

[11] Pranita Deshpande, Cynthia B. Lee, and Irfan Ahmed. 2019. Evaluation of Peer Instruction for Cybersecurity Education. In *Proceedings of the 50th ACM Technical Symposium on Computer Science Education (SIGCSE '19)*. Association for Computing Machinery, New York, NY, USA, 720–725. doi:10.1145/3287324.3287403

[12] Barbara J Ericson and Bradley N Miller. 2020. Runestone: A Platform for Free, Online, and Interactive Ebooks. In *Proceedings of the 51st ACM Technical Symposium on Computer Science Education*. 1012–1018.

[13] Bhuvana Gopal and Stephen Cooper. 2021. Peer Instruction in Software Testing and Continuous Integration. In *Proceedings of the 52nd ACM Technical Symposium on Computer Science Education (SIGCSE '21)*. Association for Computing Machinery, New York, NY, USA, 548–554. doi:10.1145/3408877.3432404

[14] Bhuvaneswari Gopal and Stephen Cooper. 2022. Peer instruction in online software testing and continuous integration: a replication study. In *Proceedings of the ACM/IEEE 44th International Conference on Software Engineering: Software Engineering Education and Training (ICSE-SEET '22)*. Association for Computing Machinery, New York, NY, USA, 199–204. doi:10.1145/3510456.3514168

[15] Geoffrey L. Herman and Sushmita Azad. 2020. A Comparison of Peer Instruction and Collaborative Problem Solving in a Computer Architecture Course. In *Proceedings of the 51st ACM Technical Symposium on Computer Science Education (SIGCSE '20)*. Association for Computing Machinery, New York, NY, USA, 461–467. doi:10.1145/3328778.3366819

[16] David Hestenes, Malcolm Wells, and Gregg Swackhamer. 1992. Force concept inventory. *The physics teacher* 30, 3 (1992), 141–158.

[17] Jennifer K Knight and William B Wood. 2005. Teaching more by lecturing less. *Cell biology education* 4, 4 (2005), 298–310.

[18] Cynthia Bailey Lee, Saturnino Garcia, and Leo Porter. 2013. Can peer instruction be effective in upper-division computer science courses? *ACM Transactions on Computing Education (TOCE)* 13, 3 (2013), 1–22.

[19] Leo Porter, Cynthia Bailey Lee, and Beth Simon. 2013. Halving fail rates using peer instruction: a study of four computer science courses. In *Proceeding of the 44th ACM technical symposium on Computer science education*. 177–182.

[20] Leo Porter, Cynthia Bailey Lee, Beth Simon, Quintin Cutts, and Daniel Zingaro. 2011. Experience report: a multi-classroom report on the value of peer instruction. In *Proceedings of the 16th annual joint conference on Innovation and technology in computer science education*. 138–142.

[21] Leo Porter, Cynthia Bailey Lee, Beth Simon, and Daniel Zingaro. 2011. Peer instruction: do students really learn from peer discussion in computing?. In *Proceedings of the seventh international workshop on Computing education research*. 45–52.

[22] Leo Porter, Dennis Bouvier, Quintin Cutts, Scott Grissom, Cynthia Lee, Robert McCartney, Daniel Zingaro, and Beth Simon. 2016. A multi-institutional study of peer instruction in introductory computing. In *Proceedings of the 47th ACM Technical Symposium on Computing Science Education*. 358–363.

[23] Leo Porter, Saturnino Garcia, John Glick, Andrew Matusiewicz, and Cynthia Taylor. 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*. 129–134.

[24] Leo Porter and Beth Simon. 2013. Retaining nearly one-third more majors with a trio of instructional best practices in CS1. In *Proceeding of the 44th ACM technical symposium on Computer science education*. 165–170.

[25] Leo Porter, Daniel Zingaro, and Raymond Lister. 2014. Predicting student success using fine grain clicker data. In *Proceedings of the tenth annual conference on International computing education research*. 51–58.

[26] Beth Simon, Michael Kohanfars, Jeff Lee, Karen Tamayo, and Quintin Cutts. 2010. Experience report: peer instruction in introductory computing. In *Proceedings of the 41st ACM technical symposium on Computer science education*. 341–345.

[27] Beth Simon, Julian Parris, and Jaime Spacco. 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*. 41–46.

[28] Michelle K Smith, William B Wood, Wendy K Adams, Carl Wieman, Jennifer K Knight, Nancy Guild, and Tin Tin Su. 2009. Why peer discussion improves student performance on in-class concept questions. *Science* 323, 5910 (2009), 122–124.

[29] Michelle K Smith, William B Wood, Ken Krauter, and Jennifer K Knight. 2011. Combining peer discussion with instructor explanation increases student learning from in-class concept questions. *CBE—Life Sciences Education* 10, 1 (2011), 55–63.

[30] Xiaohang Tang, Xi Chen, Sam Wong, and Yan Chen. 2023. VizPI: A Real-Time Visualization Tool for Enhancing Peer Instruction in Large-Scale Programming Lectures. In *Adjunct Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology (UIST '23 Adjunct)*. Association for Computing Machinery, New York, NY, USA, 1–3. doi:10.1145/3586182.3616632

[31] Lev Semenovich Vygotsky. 1980. *Mind in society: The development of higher psychological processes*. Harvard university press.

[32] Barry J Wadsworth. 1996. *Piaget's theory of cognitive and affective development: Foundations of constructivism.* Longman Publishing.

[33] Feng Wang and Michael J Hannafin. 2005. Design-based research and technology-enhanced learning environments. *Educational technology research and development* 53, 4 (2005), 5–23.

[34] Daniel Zingaro. 2010. Experience report: Peer instruction in remedial computer science. In *EdMedia+ Innovate Learning*. Association for the Advancement of Computing in Education (AACE), 5030–5035.

[35] Daniel Zingaro and Leo Porter. 2014. Peer Instruction in computing: The value of instructor intervention. *Computers & Education* 71 (2014), 87–96.