USING PEER INSTRUCTION TO TEACH THRESHOLD CONCEPTS IN INTRODUCTORY STATISTICS

Karol P. Binkowski and Nicholas M.K. Tse
Macquarie University, Australia
karol.binkowski@mq.edu.au

Teaching threshold concepts for introductory statistics is challenging in large lecture settings. Peer Instruction (PI) has been promoted as an effective method to increase classroom engagement and support active learning. In our context, the first-year statistics large enrolment course at the undergraduate level is offered to a large cohort of students pursuing a Bachelor of Science and various non-statistical majors. We adopted the PI method during in-theatre hybrid Q&A live lecture sessions. Our data indicate students preferred the PI format over traditional lectures, and live lecture attendance increased compared to previous years. We also explored the relationship between students' perceived PI benefits and academic performance. We observed that with the PI method, lecture attendance doubled compared to prior years. We investigated students' perceived benefits from PI and academic performance within this unit. These findings in the broader context of teaching statistics demonstrate that simple, interactive methods like PI can significantly boost student engagement and performance. This supports the notion that active learning strategies are effective in helping students understand and retain material, especially in large, diverse cohorts.

INTRODUCTION

The study aimed to address the challenge of teaching threshold concepts in introductory statistics to a large cohort of undergraduate students by implementing Peer Instruction (PI) in live lecture sessions. Moreover, the study was prompted by the observed low attendance and lack of engagement in live lectures, highlighting the necessity to improve student participation and comprehension in the teaching of introductory statistics to a diverse undergraduate cohort, ultimately leading to the adoption of PI as a strategy to enhance classroom engagement and support active learning. We considered the following research questions. What are the preferences of students regarding the PI approach in live lectures? How has PI been observed to enhance engagement in live lectures? Are students' perceived benefits from the PI format associated with improved performance? How has PI contributed to students' understanding of threshold concepts?

BACKGROUND

Threshold concepts

Threshold concepts are those ideas that bind a subject together and are fundamental to the thinking and practice within a discipline (Beitelmal, 2022). Helping students understand these concepts is imperative so they can apply the knowledge in their field of study. These chosen questions are, on the surface, easily misunderstood and misinterpreted and are designed to provide a generalised and yet relatable abstraction to core statistical concepts.

In our context, the threshold concepts include understanding variability in data, sampling methods for inference, different distributions and shapes of the normal distribution, hypothesis testing, constructing confidence intervals, and interpreting statistical results.

Peer Instruction

Peer Instruction, also known as Mazur's sequence, is a teaching method that enhances lectures by engaging students actively through short conceptual questions to identify and correct common misconceptions (Crouch & Mazur, 2001; Mazur, 1997). The term "peer" in PI refers to the interaction among students or peers during the learning process. Recent studies by Ibrahim and Geller (2023) have explored the benefits of peer discussion in multimedia learning environments, finding that PI can significantly enhance course assessment performance. Similarly, Rayens and Ellis (2018) have reported on the transition from face-to-face to online student-centred classrooms, highlighting the crucial role of active learning in maintaining student engagement in virtual settings. Al-Haddad et al. (2024) introduced

In: Kaplan, J. & Luebke, K. (Ed.) (2024). *Connecting data and people for inclusive statistics and data science education. Proceedings of the Roundtable conference of the International Association for Statistics Education (IASE), July 2024, Auckland, New Zealand.*

the Technology-Enhanced Supportive Instruction (TSI) model, which integrates technology and supportive feedback to engage students in statistics education. Although our approach adapts to a hybrid environment combining face-to-face and online elements, it focuses primarily on face-to-face interaction. Similar findings have been reported by Arjomandi et al. (2023), who investigated the effects of real-time interactive student polling on academic performance. Their study revealed that such polling did not impact final grades or failure rates, but it did increase student engagement and tutorial attendance. Corbo and Sasaki (2021) examined active learning methodologies in an introductory statistics course and found that although student performance improved, dropout rates remained unaffected. This underscores the complexity of enhancing student retention alongside performance through active learning techniques.

Teaching context

The Faculty of Science and Engineering offers a generalist degree, *Bachelor of Science*, with several majors run by departments and schools across the university. With an annual enrolment of approximately 1200 students in each of the two sessions, the Introductory Statistics unit comprises students enrolled in various predominantly non-statistical, STEM or non-STEM disciplines. Binkowski (2023) provides more background on the unit, including the educational philosophy of Mastery Learning that underpins the assessment structure and analyses students' performance across various degree courses. Binkowski and Tse (2023) explored the relationship between groupings of degree disciplines and students' preferred learning modality, such as face-to-face, online or hybrid.

The Introductory Statistics unit is structured into five topics, each spanning two weeks. Each topic is assessed by three individualised quizzes that sum to 100 marks each, ranking from pass-level questions to two more difficulty levels. All quizzes are graded automatically. The university grade categories are Fail (F, <50%), Pass (P, 50-64), Credit (CR, 65-74), Distinction (D, 75-84), High Distinction (HD, 85-100), and Fail Hurdle (FH, reduces final mark to 49).

Students engaged in three hours of instructional contact per week, including a one-hour lecture and two small group sessions conducted on-campus or online. These small group sessions, accommodating up to 28 students per class, reinforced conceptual understanding. In contrast, practical computer lab classes focused on applying statistical methods using Excel data sets. Our course adopted a hybrid teaching format, integrating pre-recorded content with live online Q&A sessions to enhance the learning experience. The lecture sessions, attended by up to 500 students, were conducted by a lecturer. In Session 1 of 2024, we introduced Peer Instruction as a strategy for teaching, incorporating role-play discussions, pre-recorded lectures, and self-paced quizzes. The outcomes of this approach are discussed in this paper.

Adapting Peer Instruction for large, hybrid, diverse cohorts

Despite the established benefits of PI in various educational settings, there needs to be a greater understanding of its effectiveness in an environment similar to ours. The hybrid learning environment combines face-to-face and online elements, an ample theatre space of up to 500 seats, and diverse cohorts. This knowledge gap prompted us to explore how PI could be adapted and implemented in our Introductory Statistics service unit to enhance student engagement and performance. Our study investigated the impact of PI on student learning outcomes, engagement levels, and overall course satisfaction within this hybrid context.

METHODOLOGY

Participation in the Peer Instruction live lecture session

Out of the 1,200 students enrolled, 763 undergraduate and internal students were eligible to attend the live lecture sessions held in a theatre with a capacity of 500 seats. Students could register for in-person attendance or join via an online synchronous stream that included a video of the lecture screen and stage. Since shifting to an entirely online format in 2022, physical lecture attendance had ceased to be mandatory. Post-COVID-19 lockdowns, only a fraction of students participated in the one-hour Q&A sessions. For instance, attendance in session 2 of 2022 dwindled from a packed theatre in Week 1 to just 70 students by Week 5. In response, several active learning enhancements were introduced in 2023 to increase lecture appeal. Five lecture slides looped continuously before each session (e.g. from 9:55 AM to 10:05 AM), accompanied by cheerful music to welcome students as they entered the theatre. Each lecture featured peer discussions facilitated by lecturers using tablets to provide hints and assist with problem-solving. Surveys were occasionally conducted during these sessions to gather data for Excel demonstrations on sampling distributions. Industry guest speakers were invited to integrate real-world insights into the weekly topics. Theoretical explanations were restricted to 5 minutes within a tightly scripted schedule. These changes improved engagement, stabilising attendance at around 120 by Week 5. In Session 1 of 2024, we shifted from peer discussions to PI, which led to a significant increase in live lecture participation, doubling the figures from the previous year.

Peer Instruction questions

To create effective PI questions, we analysed students' frequent mistakes by reviewing past tests and noting common errors. Students were required to prepare by reading weekly lecture notes and watching pre-recorded lectures. During class, the focus shifted to verifying their understanding through interactive questioning. We designed PI questions and encoded them into Microsoft Forms quizzes that could be accessed on mobile phones via QR codes. Such questions included the following,

- Suppose your statistics instructor says you scored 70 on an exam, and the class mean/average was 74. You should hope that the standard deviation of exam scores was Small/Large.
- Choose the most representative sample of the target population, followed by a description of four choices of the target population and the sample.
- Simpson's Paradox: analyse associations between smoking, age, and age mortality.
- A plot of skewed histogram is followed by a question on what is larger, median, or mean.
- Given a confidence interval, do you reject the hypothesised value of a parameter?
- Question about refutable claim as a precursor to the null hypothesis.

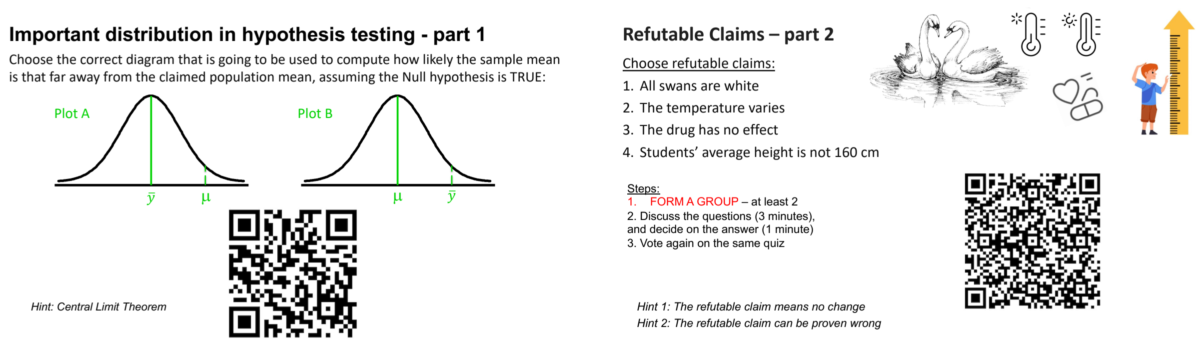

Figure 1. Examples of slides with a PI question, multiple choice, and QR code that leads to Microsoft Form, where the quiz is provided to students and displayed on a projector.

In a typical one-hour session, we incorporated 3 to 4 PI sequences of questions. The sequence introduced the concept, followed by a pre-assessment using QR codes that students answered anonymously using their phones (see Figure 1). This led to a peer discussion during which students exchanged views. The lecturer circulated, offering hints and facilitating the conversation. A re-assessment followed, during which students answered a similar question to measure learning progress. The session concluded with the lecturer explaining the concepts in detail, addressing emerging misconceptions, and reinforcing key points. Students were encouraged to assume roles like advocate, contrarian, or observer to foster discussion, and to use third-person phrases like "Alice says" or "Bob says" to facilitate participation.

Data

Between weeks 1 and 4, at the end of each lecture, we invited all students to participate in a survey by email. The dataset used for analysis was derived from a combination of grade book records and survey responses, necessitating intricate data matching and the removal of personal identifiers such as student IDs, names, and email addresses. Additionally, to maintain analytic clarity, the dataset excluded students who did not participate in the PI lecture sessions or did not respond to the survey.

This exclusion introduced a selection bias, potentially omitting a subset of students whose data could provide valuable insights into the overall effectiveness of the PI method. The Macquarie University Human Research Ethics Committee has approved the study (HREA-16720). The dataset was de-identified before the analysis was performed to ensure compliance with ethical and privacy standards.

Students' questionnaire and profile

Using a Learning Management System (LMS), we built a profile of each student's participation in the peer discussion. Each week after the lecture session, we asked students to respond to the following questions:

1. Have you attended the lecture in Week #? [YES/NO]
2. What is your discipline of study or the name of your degree? [TEXT INPUT]
3. How familiar are you with the topic in Week # [rank: 0, 1, 2]
4. How the lecture improved your understanding of the topic in Week # [rank: 0, 1, 2]

The ranks in question 3 represent: Not familiar at all, Have a general understanding, Have a great understanding, and in question 4 represent: Did not help me, somewhat helped me, significantly helped me. We have run the survey in weeks 1 to 4.

RESULTS

Possible increase in student engagement with PI

The most noticeable outcome of the PI approach was the marked increase in lecture attendance. Attendance data, derived from responses to PI questions, indicated a sequential decrease in student engagement over the five weeks, with response totals of 500, 460, 380, 300, and 200, respectively (see Figure 2). Although there was a decrease, the rate of change was much slower than in the previous years' offering. In other words, the retention rate of student participation in the lectures improved. This pattern of engagement was observed within a hybrid framework of synchronous sessions, incorporating both in-person and online modalities. Such high levels of student participation had yet to be recorded for this unit since the removal of the mandatory two-hour lecture in 2019, before the COVID-19 pandemic.

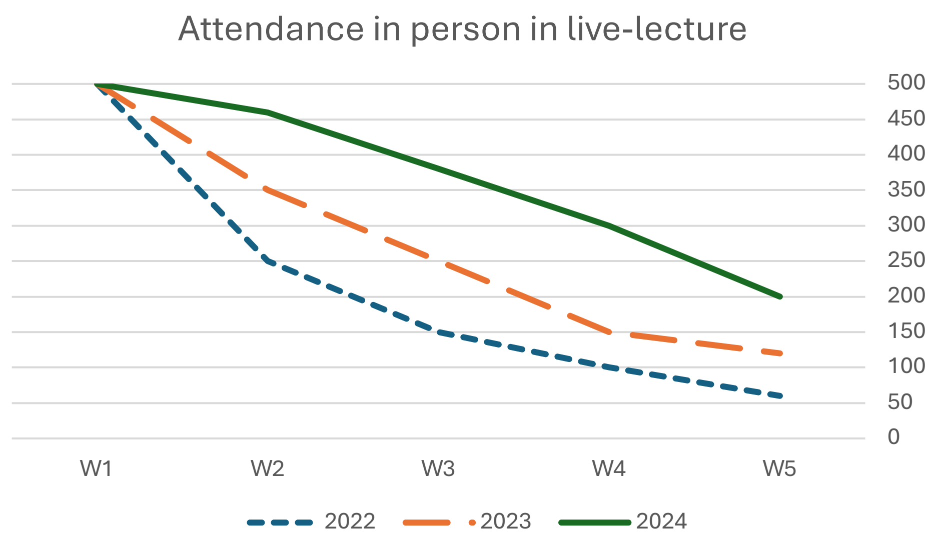

Figure 2. The attendance in live lectures across the five weeks for 2022, 2023 and 2024. They were derived based on the number of responses to online quizzes during the lecture.

Potential improvement in self-efficacy through PI

Further insights were gained by analysing students' self-assessments regarding their comprehension of weekly topics, juxtaposed with their actual performance on quizzes corresponding to these topics. The data revealed a high level of perceived benefit from the PI sessions, with self-reported satisfaction rates of 97%, 96%, 98%, and 94% over the four weeks, as shown in Table 1. Interestingly, while 41% and 40% of students reported a good understanding of the topics in weeks 1 and 2, this percentage notably declined to 28% and 26% in weeks 3 and 4. Despite this decline in self-reported understanding, more students felt that the PI sessions were highly beneficial to their learning process. This shift suggests a complex interaction between perceived knowledge and the perceived effectiveness of the PI approach.

We noted that the cohort is very optimistic, as their belief in their understanding of the topics is very high each week, e.g. 92%, 94%, 95%, and 79%, respectively. The drop in week four might be related to the timing of the first assessment task.

Table 1. Distribution of students' self-reporting of comprehension of weekly topics (Grasp: 1 – not familiar, 1 – general understanding, 2 – good understanding) and the perceived benefit derived from attending PI lecture sessions, categorised by rankings 0 – indifferent, 1 – some benefit, and 2 – great benefit.

Grasp	0 Indifferent	1 Some benefit	2 Great benefit	Total	0 Indifferent	1 Some benefit	2 Great benefit	Total
	Week 1				**Week 2**			
0 Not familiar	1.4%	4.3%	2.9%	9%	0.0%	5.0%	0.8%	6%
1 General	1.8%	40.7%	17.1%	60%	1.7%	39.5%	11.8%	53%
2 Good	0.0%	17.5%	14.3%	32%	2.5%	24.4%	14.3%	41%
Total	**3.0%**	**63.0%**	**34.0%**	**n=280**	**4.0%**	**69.0%**	**27.0%**	**n=119**
	Week 3				**Week 4**			
0 Not familiar	0.0%	4.7%	1.2%	6%	5.3%	12.3%	3.5%	21%
1 General	1.2%	28.2%	21.2%	51%	0.0%	26.3%	17.5%	44%
2 Good	1.2%	18.8%	23.5%	44%	0.0%	22.8%	12.3%	35%
Total	**2.0%**	**52.0%**	**46.0%**	**n=85**	**5.0%**	**61.0%**	**33.0%**	**n=57**

Students self-reported benefit and their marks

According to Figure 3, students who perceived no benefit from the PI format of the lecture, labelled as indifferent, would score the least across all five module tests, except Week 3. Interestingly, students who perceived the PI format as somewhat beneficial achieved higher average marks than those who considered it greatly beneficial. The number of participants who took surveys were 271, 102, 77, and 49 (discrepancy with Table 1 due to some students who did not complete the quizzes). We also want to emphasise that the groups indifferent consist of only several students each week and are tiny fractions, as reported in Table 1.

DISCUSSION

A preliminary analysis of Peer Instruction's effectiveness

The PI pedagogical approach was implemented during the main lecture session, which accounts for one of the three weekly contact hours with students. This strategy significantly boosted attendance to levels not seen before the COVID-19 pandemic. Most students reported benefiting from the PI sessions, indicated by high satisfaction rates and positive self-assessments of understanding. The increasing trend in student performance across the weeks may be attributed to heightened engagement with the material. This engagement could lead to better performance, or the trend might result from self-selection bias, where diminishing response rates leave only the more motivated and engaged students. These students, who are more inclined to participate and report the PI format positively, tend to outperform their peers who are less interested in the material.

Deslauriers et al. (2019) reported that students in active classrooms learned more than their peers in passive environments despite having a lower perception of their learning outcomes. This finding aligns with our study's result, indicating that students who viewed the PI format as moderately beneficial achieved higher average marks than those who perceived it as highly beneficial.

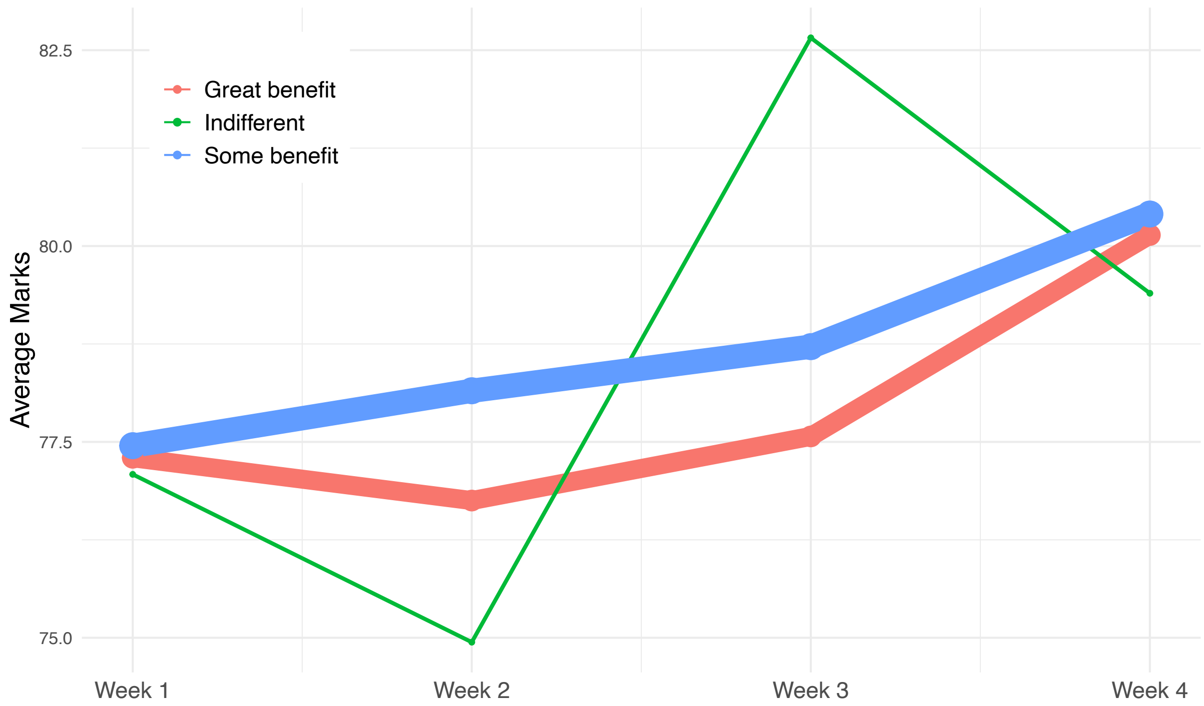

Figure 3. The average students' marks for each of the four surveyed weeks, categorised by their self-reported benefit from the PI approach. The width of the line indicates the proportion of the average proportion of responses (Indifferent 3.5%, Great benefit 35%, and Some benefit 61.5%).

In this exploratory study, we aimed to answer whether PI is an effective method in helping learners shape their understanding of threshold concepts. Our analysis suggests a complex interplay between student satisfaction, subjective learning beliefs, and actual performance as the coursework progressed. Challenges like survey fatigue, technical issues with QR codes, and some students' preference for traditional lectures also affected the study's outcomes. Therefore, while initial results are encouraging, they are preliminary.

Considerations and limitations

This study faces several limitations that may impact the interpretation of its findings. First, a segment of the student population preferred traditional lecture formats, perceiving the PI approach as overly intrusive. Additionally, technical barriers were noted, including difficulties with QR code access, compounded by not all students possessing mobile phones to participate in class-based self-assessments. We did not have a method to record the true lecture attendance, and instead, we used the number of responses to online quizzes during the lecture. Survey fatigue also posed a significant challenge, likely leading to inaccuracies in self-reporting and diminished participation as the semester progressed. As students complete the form voluntarily, this is a potential source of bias for those who decided to complete the form each week. Survey response started at n=280 and decreased to n=57 (see Table 1). Lastly, the surveys were conducted post-lecture via email, introducing a delay that could misrepresent the actual engagement during the lectures, as evidenced by a decline in attendance not mirrored by the survey response rates. These factors collectively suggest that the study's current conclusions should be considered preliminary.

CONCLUSIONS

We considered a creative way to support learners within under-resourced contexts, particularly students having trouble with challenging abstract concepts or first-in-family students who needed more time and opportunities to engage with the concepts. We define these students as those who lack a solid foundation in mathematics due to insufficient instruction, inadequate resources, or other educational disruptions.

We also presented a refreshed PI method using QR codes and active learning elements in a modern learning environment. We plan to refine and enhance PI implementation in Session 2 of 2024, with further data collection and analysis to assess its efficacy more definitively.

In future studies, we plan to explicitly address the fourth research question on how PI helps to improve understanding of threshold concepts. We will closely align the live-lecture quizzes with the graded assessment tasks in time to better capture the impact of PI on students' learning.

REFERENCES

Al-Haddad, S., Chick, N., & Safi, F. (2024). Teaching statistics: A Technology-Enhanced Supportive Instruction (TSI) model during the Covid-19 pandemic and beyond. *Journal of Statistics and Data Science Education*, 32(2), 129–142. https://doi.org/10.1080/26939169.2024.2315939

Arjomandi, A., Paloyo, A. R., & Suardi, S. (2023). Active learning and academic performance: The case of real-time interactive student polling. *Statistics Education Research Journal*, *22*(1). https://doi.org/10.52041/serj.v22i1.122

Beitelmal WH, Littlejohn R, Okonkwo PC, Hassan IU, Barhoumi EM, Khozaei F, Hassan AM, & Alkaaf KA. Threshold concepts theory in higher education—Introductory statistics courses as an example. *Education Sciences*, *12*(11), 748. https://doi.org/10.3390/educsci12110748

Binkowski, K. P. (2023). Mastery of learning — does it make a difference to students' online engagement and performance in a first-year statistics unit? *Proceedings of the Satellite conference of the International Association for Statistical Education (IASE) 2023, Fostering Learning of Statistics and Data Science.* International Association for Statistical Education.

Binkowski, K. P., & Tse, N.M.K. (2023). What learning activities and resources do the students prefer for their learning in a large first-year enrolment statistics service unit? *Proceedings of the 11th Australian Conference on Teaching Statistics (OZCOTS) 2023.*

Corbo, A. R., & Sasaki, D. G. G. (2021). Using active learning methodologies in an introductory statistics course. *Statistics Education Research Journal, 20*(2). https://doi.org/10.52041/serj.v20i2.113

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

Deslauriers, L., McCarty, L. S., Miller, K., Callaghan, K., & Kestin, G. (2019). Measuring actual learning versus feeling of learning in response to being actively engaged in the classroom. *Proceedings of the National Academy of Sciences - PNAS*, *116*(39), 19251-19257. https://doi.org/10.1073/pnas.1821936116

Ibrahim, D., & Geller, E. H. (2023). Peer Discussion in Multimedia Learning Environments. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *45*. https://escholarship.org/uc/item/5xf6v732

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

Rayens, W., & Ellis, A. (2018). Creating a student-centered learning environment online. *Journal of Statistics Education*, *26*(2), 92–102. https://doi.org/10.1080/10691898.2018.1475205