Brief Report on Peer evaluation system

Harnessing Peer Power: A Novel Approach to Auto Answer Evaluation in Education

In large classes and MOOCs, the traditional method of manual evaluation by teachers and TAs can be overwhelming due to the sheer volume of student responses. This often leads to delays in teaching / Quality of teaching, which can hinder student learning and engagement. To address this challenge, we propose a peer evaluation system for auto answer evaluation that distributes the evaluation workload among students.

Key Components of the Peer Evaluation System:

1. **Auto Evaluation Mechanism:** Develop an automated system that can evaluate student answers based on predefined criteria. This system should be capable of handling various types of questions, including multiple-choice, short answer, and essay questions.

2. **Peer Evaluation Framework:** Implement a peer evaluation framework where students are assigned a set of answers from their peers to evaluate based on the same criteria used by the automated system. Each answer should be evaluated by multiple peers to ensure fairness and accuracy.

3. **Incentive Structure:** Provide incentives for students to participate in the peer evaluation process. This could include bonus points, recognition, or other rewards that motivate students to take the task seriously.

4. **Quality Control Mechanisms:** Implement quality control mechanisms to ensure the accuracy and reliability of peer evaluations. This could include cross-checking evaluations by multiple peers and allowing students to appeal the evaluations.

5. **Feedback Loop:** Establish a feedback loop where students receive feedback on their evaluations and have the opportunity to learn from their peers' feedback on their own answers. This promotes a culture of continuous improvement and peer learning.

Benefits of the Peer Evaluation System:

1. **Reduced Workload:** By distributing the evaluation workload among students, the burden on teachers and TAs is significantly reduced, allowing them to focus on other aspects of teaching and learning.

2. **Faster Feedback:** The peer evaluation system can provide faster feedback to students, enabling them to identify and address their mistakes more quickly, leading to improved learning outcomes.

3. **Promotes Critical Thinking:** Evaluating their peers' answers can help students develop critical thinking and analytical skills, as they are required to assess the quality of the answers based on predefined criteria.

4. **Encourages Engagement:** The incentive structure and peer interaction involved in the evaluation process can increase student engagement and motivation, leading to a more active learning environment.

Overall, implementing a peer evaluation system for auto answer evaluation in large classes and MOOCs can lead to more efficient and effective evaluation processes, benefiting both students and teachers.

1. Introduction:

Traditional teaching methods often involve passive or active learning, which can lead to an increased workload for teachers and TAs, particularly in large classes with more than 100 students. To alleviate this burden, we propose a hierarchical teaching approach that distributes the workload evenly among teachers, TAs, super peers, and peers (students). This approach not only reduces the workload on instructors but also promotes collaborative learning among students.

Learning Management Systems (LMS) are widely used for assignments, quizzes, and short-answer questions in educational settings. However, these systems typically require the presence of teachers or TAs for evaluation, limiting their effectiveness as true auto evaluation systems. Peer review and feedback systems have been recognized in literature as effective methods for reducing the evaluation burden on instructors. In this paper, we focus on the implementation of a peer evaluation system for auto answer evaluation.

In our proposed peer evaluation system, students act as evaluators for their peers' answer scripts. They receive an answer scheme from the teacher or TA and evaluate the answer scripts accordingly. The evaluation process is autonomous, meaning that students do not know who evaluated their answers, and vice versa. Additionally, each answer sheet is evaluated multiple times by different peers to ensure accuracy and fairness in evaluation.

The peer evaluation system presented in this paper offers a novel approach to auto answer evaluation in large classes and MOOCs. By leveraging the collective intelligence of students, this system effectively distributes the evaluation workload, reducing the burden on teachers and TAs. However in this report we have taken the help of both Teacher's and Teaching Assistants to compare and justify the results obtained. Further research could focus on implementing and evaluating this system in real-world educational settings to assess its effectiveness and scalability.

2. Literature Review:

2.1 LearnEval Peer Assessment Platform: Iterative Development Process and Evaluation (Gabriel Badea and Elvira Popescu, Senior Member, IEEE):

"LearnEval Peer Assessment Platform: Iterative Development Process and Evaluation" by Gabriel Badea and Elvira Popescu, Senior Member, IEEE, provides a comprehensive overview of the development, features, and evaluation of LearnEval, a peer assessment platform designed for educational settings. The paper outlines the iterative development process that was undertaken to create the platform, emphasizing the various stages and methodologies utilized to enhance its functionality and user-friendliness.

The initial design of LearnEval aimed to simplify the peer assessment process, ensuring fairness, anonymity, and reliability in evaluations. Challenges encountered during development, such as algorithm refinement for automatic answer evaluation and feedback generation, are discussed. Integration of machine learning algorithms into the platform is highlighted as a key strategy to improve its capabilities.

The paper also delves into the evaluation of LearnEval, detailing the methodologies used to assess its effectiveness. This includes user surveys, performance metrics analysis, and comparative studies with existing peer assessment tools. Results from these evaluations indicate that LearnEval has a positive impact on student learning outcomes and can streamline the assessment process in educational settings.

Overall, the paper presents LearnEval as a robust and innovative peer assessment platform. Its iterative development approach and thorough evaluation process contribute valuable insights into the design, implementation, and evaluation of similar platforms in educational environments.

2.2 Incentive Design in Peer Review: Rating and Repeated Endogenous Matching (Yuanzhang Xiao, Florian Dorfler, and Mihaela van der Schaar, Members/Fellows, IEEE):

The paper "Incentive Design in Peer Review: Rating and Repeated Endogenous Matching" by Yuanzhang Xiao, Florian Dorfler, and Mihaela van der Schaar, published in IEEE Transactions on Network Science and Engineering, addresses the challenges of traditional peer review systems and proposes a novel incentive mechanism to improve the quality and integrity of the review process.

The authors highlight the critical role of peer review in academic publishing and other domains for ensuring the validity and quality of submissions. However, they note several shortcomings in traditional peer review, including biased reviews, strategic behavior by reviewers, and a lack of sufficient incentives for reviewers to provide thoughtful and accurate evaluations.

To address these challenges, the authors propose a new incentive mechanism that combines two key components: rating and repeated endogenous matching. In their system, reviewers are not only asked to provide a binary decision (accept/reject) for a submission but also to rate the submission on a continuous scale. Reviewers are then matched repeatedly with the same submission over multiple rounds, allowing them to update their ratings based on feedback from other reviewers. This repeated matching incentivizes reviewers to provide accurate and thoughtful evaluations, as their reputations are tied to the quality of their ratings.

The paper discusses the theoretical properties of the proposed mechanism, including its convergence properties and its ability to incentivize truthful reporting. The authors also conduct simulations to demonstrate the effectiveness of their approach compared to existing mechanisms. The results show that the proposed mechanism improves the quality of reviews and incentivizes honest behavior among reviewers.

Overall, the paper provides a comprehensive framework for designing incentives in peer review systems, aiming to enhance the quality and integrity of the review process in academic publishing and beyond.

2.3 A Framework for Teaching Evaluation (Divya Nalla, Department of Computer Science, Nalla Malla Reddy Engineering College):

The paper proposes a comprehensive framework for evaluating teaching effectiveness, which could be beneficial for educators, administrators, and policymakers. The framework likely includes various components such as:

1. Objective Setting: Defining clear objectives for teaching evaluation, ensuring alignment with the educational goals of the institution or program.
2. Metrics Selection: Identifying relevant metrics or criteria for evaluating teaching effectiveness, which could include student performance, student feedback, peer evaluations, and self-assessment by instructors.
3. Data Collection: Describing methods for collecting data on teaching effectiveness, such as surveys, classroom observations, and analysis of student outcomes.
4. Data Analysis: Providing a systematic approach for analyzing the collected data to assess teaching effectiveness objectively.
5. Feedback Mechanism: Implementing a feedback mechanism to communicate evaluation results to instructors, allowing them to improve their teaching practices.
6. Continuous Improvement: Emphasizing the importance of continuous improvement in teaching based on evaluation results and feedback.
7. Implementation Guidelines: Offering guidelines for implementing the framework in educational institutions, including training for evaluators and instructors.

Overall, the framework is likely designed to enhance the quality of teaching and learning by providing a structured approach to evaluating and improving teaching effectiveness.

2.4 Learning Management System User Guide: Self & Peer Assessment (University of Melbourne): The paper "Learning Management System User Guide: Self & Peer Assessment (University of Melbourne)" likely provides a comprehensive guide on how to use the self and peer assessment features within the University of Melbourne's Learning Management System (LMS). A typical summary might include the following key points:

1. Introduction to Self & Peer Assessment: The guide starts with an overview of the importance and benefits of self and peer assessment in higher education, highlighting its role in promoting deeper learning, critical thinking, and feedback skills among students.
2. Accessing the Self & Peer Assessment Tool: It explains how students and instructors can access the self and peer assessment tool within the University of Melbourne's LMS. This includes step-by-step instructions on navigating to the tool and accessing assessment activities.
3. Creating Assessment Activities: The guide provides instructions on how instructors can create self and peer assessment activities within the LMS. This includes setting assessment criteria, deadlines, and instructions for students.

4. Participating in Self & Peer Assessment: For students, the guide explains how to participate in self and peer assessment activities. This includes instructions on how to submit their own work for assessment, how to evaluate peer submissions, and how to provide constructive feedback.
5. Viewing Assessment Results: Both students and instructors may be provided guidance on how to view assessment results within the LMS. This could include accessing feedback from peers, viewing grades, and reviewing assessment reports.
6. Best Practices and Tips: The guide may also include best practices and tips for both students and instructors to make the most of the self and peer assessment process. This could include advice on giving and receiving feedback, maintaining academic integrity, and using assessment results for learning improvement.
7. Troubleshooting and Support: Lastly, the guide provides information on troubleshooting common issues that may arise during the self and peer assessment process. It also includes information on where to seek further support, such as contacting IT support or academic staff.

Overall, the guide is designed to provide comprehensive support to both students and instructors using the self and peer assessment features within the University of Melbourne's LMS, aiming to enhance the effectiveness of assessment practices and promote student learning outcomes.

2.5 A Graph Analysis Method to Improve Peer Grading Accuracy for Blended Teaching Courses (Xing Du, Xingya Wang, and Yan Ma):

The paper by Xing Du, Xingya Wang, and Yan Ma presents a comprehensive method for enhancing the accuracy of peer grading in blended teaching courses, focusing on the challenges of bias, inconsistency, and lack of expertise among peers. The authors begin by highlighting the importance of peer grading in large-scale courses for providing timely and personalized feedback to students, as well as promoting a deeper understanding of course material through the evaluation of peers' work.

To address these challenges, the authors propose a novel approach based on graph analysis, which they argue can effectively model the complex relationships between students and their grading behaviors. The method involves several key steps, starting with data preprocessing to clean and standardize the grading data. This includes identifying and correcting errors in grading, as well as normalizing scores to account for differences in grading standards among peers.

Next, the authors describe the process of constructing the grading graph, which involves representing students as nodes and grading relationships as edges. The graph is constructed based on various factors, including the similarity of grading patterns between students and the reliability of graders. This allows for the identification of clusters of students with similar grading behaviors, as well as the detection of outliers who may exhibit biased or inconsistent grading.

The paper also discusses the use of graph analysis techniques to analyze the grading graph and improve grading accuracy. This includes identifying and removing biased or unreliable graders, as well as adjusting grades based on the consensus of multiple graders. The authors demonstrate the effectiveness of their approach through experimental results, which show a significant improvement in grading accuracy compared to traditional peer grading methods.

Overall, the paper provides a detailed and innovative approach to improving peer grading accuracy in blended teaching courses. By leveraging graph analysis techniques, the proposed method offers a promising solution to the challenges of bias, inconsistency, and lack of expertise in peer grading, ultimately enhancing the quality of feedback and learning outcomes for students.

2.6 Fostering Peer Learning With a Game-Theoretical Approach in a Blended Learning Environment Seyede Fatemeh Noorani , Mohammad Hossein Manshaei , Mohammad Ali Montazeri, and Behnaz Omoomi.

"Fostering Peer Learning With a Game-Theoretical Approach in a Blended Learning Environment" by Seyede Fatemeh Noorani, Mohammad Hossein Manshaei, Mohammad Ali Montazeri, and Behnaz Omoomi explores the use of a game-theoretical approach to promote peer learning in a blended learning environment. The paper discusses the challenges of traditional learning methods and proposes a novel approach that combines game theory with blended learning techniques to enhance student engagement and collaboration.

The authors first highlight the limitations of traditional learning methods, such as passive learning and lack of student interaction, which can hinder the effectiveness of the learning process. They then introduce the concept of blended learning, which combines traditional face-to-face teaching with online learning activities, as a potential solution to these challenges.

The paper proposes a game-theoretical framework called "Peer-Incentive-Based Learning (PIBL)" to incentivize peer learning in a blended learning environment. The

PIBL framework incorporates game elements, such as rewards and competitions, to motivate students to actively participate in learning activities and collaborate with their peers. The authors discuss how the PIBL framework can be implemented in practice, including the design of learning activities, assessment methods, and feedback mechanisms.

Furthermore, the paper presents a case study where the PIBL framework was applied in a real-world educational setting. The results of the case study demonstrate the effectiveness of the PIBL framework in promoting peer learning, improving student engagement, and enhancing learning outcomes.

Overall, the paper provides valuable insights into the potential of game-theoretical approaches to enhance peer learning in blended learning environments. The PIBL framework offers a novel and effective strategy to foster collaboration and engagement among students, ultimately improving the overall learning experience.

3. Summary and Shortcoming of above paper's:

1. LearnEval Peer Assessment Platform: Iterative Development Process and Evaluation:
    - Summary: The paper outlines the iterative development process of LearnEval, a peer assessment platform, emphasizing its features and evaluation methodologies. It aimed to simplify peer assessment while ensuring fairness, anonymity, and reliability in evaluations. Integration of machine learning algorithms was highlighted for enhancing its capabilities.
    - Evaluation: The paper provides a comprehensive overview of the platform's development and evaluation, showcasing its positive impact on student learning outcomes and the assessment process. However, specific metrics and comparative studies with other platforms could enhance the evaluation's depth.
2. Incentive Design in Peer Review: Rating and Repeated Endogenous Matching:
    - Summary: The paper proposes a novel incentive mechanism for peer review systems, combining rating and repeated endogenous matching to improve review quality and integrity. It addresses challenges in traditional peer review, such as biased reviews and lack of reviewer incentives.
    - Evaluation: The paper presents a well-thought-out mechanism with theoretical analysis and simulations. However, real-world implementation and validation in academic publishing settings could provide more concrete evidence of its effectiveness.
3. A Framework for Teaching Evaluation:

- Summary: The paper proposes a framework for evaluating teaching effectiveness, covering objective setting, metrics selection, data collection, analysis, feedback mechanisms, continuous improvement, and implementation guidelines.
- Evaluation: The framework offers a structured approach to teaching evaluation but lacks specific details on metrics selection, data analysis methodologies, and implementation guidelines. Real-world application and validation would enhance its practicality.

4. Learning Management System User Guide: Self & Peer Assessment:
    - Summary: The guide provides instructions on using self and peer assessment features within the University of Melbourne's Learning Management System, covering accessing the tool, creating assessment activities, participating in assessments, viewing results, best practices, and troubleshooting.
    - Evaluation: The guide offers comprehensive support for users but could benefit from more visual aids and examples to enhance usability. It could also include case studies or user testimonials to illustrate its effectiveness.
5. A Graph Analysis Method to Improve Peer Grading Accuracy for Blended Teaching Courses:
    - Summary: The paper proposes a graph analysis method to enhance peer grading accuracy in blended teaching courses, addressing challenges of bias, inconsistency, and lack of expertise among peers.
    - Evaluation: The method presents a novel approach but lacks detailed discussion on practical implementation and scalability. Real-world validation and comparison with existing methods would strengthen its credibility.
6. Fostering Peer Learning With a Game-Theoretical Approach in a Blended Learning Environment:
    - Summary: The paper introduces a game-theoretical framework, Peer-Incentive-Based Learning (PIBL), to promote peer learning in a blended learning environment, combining game elements with online learning activities.
    - Evaluation: The framework offers a creative approach to enhance peer learning, but its effectiveness could be further validated through more extensive case studies and comparisons with traditional methods.

In conclusion, while these papers introduce innovative approaches and frameworks to enhance various aspects of education, including peer assessment, peer review, teaching evaluation, and peer learning, they could benefit from more robust evaluations, practical implementations, and comparisons with existing methods to validate their effectiveness in real-world educational settings.

4. Proposed Peer Evaluation System (Annexure-I for step-by-step implementation)

The objective of this proposed system is to enhance the auto answer evaluation process in large classes and Massive Open Online Courses (MOOCs) by leveraging peer evaluation. This system aims to improve the efficiency, accuracy, and fairness of evaluations while incentivizing student participation and learning engagement.

Key Features:

1. Criteria Alignment:
    - Students evaluate their peers' answers based on the same criteria used by the automated system.
    - Criteria are clearly defined and provided to students to ensure consistency and fairness in evaluations.
2. Peer Evaluation Groups:
    - Each answer is evaluated by multiple peers to ensure diverse perspectives and reduce bias.
    - Peers are assigned randomly to evaluate answers, ensuring a fair distribution of evaluations.
3. Quality Control Mechanisms:
    - To maintain the reliability of evaluations, a subset of answers is selected for double-blind evaluation by both peers and TAs.
    - A consensus mechanism is employed to resolve discrepancies between peer evaluations and ensure accuracy.
4. Incentivization:
    - Peers are incentivized with health points, which are proportional to the marks assigned by the automated system.
    - Health points can be redeemed for various benefits, such as bonus marks, exemptions from certain assignments, or priority registration for courses.
5. User-Friendly Interface:
    - The system features a user-friendly interface that guides students through the evaluation process.
    - It provides clear instructions, examples, and rubrics to help students understand the evaluation criteria.
6. Timely Feedback:
    - The system ensures timely feedback to students by setting deadlines for evaluations and providing automated reminders.
    - Feedback is provided anonymously to maintain confidentiality and encourage honest evaluations.

7. Scalability:
    - The system is designed to be scalable, allowing it to handle a large number of evaluations simultaneously.
    - It can be easily integrated into existing learning management systems (LMS) or MOOC platforms.

Workflow:

1. Assignment Submission:
    - Students submit their answers to assignments or quizzes through the online platform.
2. Peer Evaluation Assignment:
    - Peers are assigned a subset of answers to evaluate based on the criteria provided.
3. Quality Control Evaluation:
    - A subset of answers is selected for double-blind evaluation by both peers and TAs.
    - Discrepancies between peer evaluations and TA evaluations are resolved through a consensus mechanism.
4. Health Points Allocation:
    - Peers are awarded health points based on the accuracy and consistency of their evaluations.
    - Health points are credited to the peer's account and can be redeemed for benefits as mentioned earlier.
5. Feedback and Results:
    - Students receive feedback on their submissions, including peer evaluations and TA feedback.
    - They can use this feedback to improve their future submissions and enhance their learning experience.

The proposed peer evaluation system for auto answer evaluation offers a comprehensive framework to enhance the evaluation process in large classes and MOOCs. By leveraging peer evaluation, the system aims to improve the efficiency, accuracy, and fairness of evaluations while incentivizing student participation and engagement.oceaoceaocea

5. Methodology:

5.1 Participants

A cohort of 150 students enrolled in a specific course participated in the study. Among them, 132 students actively participated in the examination.

5.2 Examination Structure

The examination consisted of 15 questions of various types, designed to comprehensively assess the students' understanding of the course material.

5.3 Answer Sheet Handling

Following the examination, the answer sheets were scanned and stored in a Google Drive folder for further analysis and evaluation.

5.4 Anonymization and Distribution

To ensure fairness and impartiality, each of the 132 answer sheet copies was assigned a unique identification number. A meticulously planned script was then utilized to distribute these copies among students and Teaching Assistants (TAs) in seven groups labeled as G1, G2, ..., G7.

5.5 Evaluation Process

Each question on every test copy was evaluated by a minimum of 3 students and one TA to ensure comprehensive evaluation. Additionally, a unique cross-evaluation system was implemented, whereby different groups were assigned to evaluate the test copies of another group to mitigate potential bias in the evaluation process. Detailed marking schemes with solutions were provided to every student and TA to standardize the evaluation process and maintain consistency.

6. Results:

Peer evaluation is a method for enhancing the accuracy and fairness of assessments in educational settings. In this study, we present a detailed analysis of a peer evaluation system implemented in a university course to evaluate student responses to examination questions. The system involved 132 students, each of whom had their responses evaluated by peers and Teaching Assistants (TAs) to ensure question-wise accuracy. Anonymity was maintained throughout the evaluation process, with each student's entry linked to a unique ID.

The peer evaluation system was implemented as follows:

1. Anonymization: Each student's entry number was linked to a unique ID to ensure anonymity throughout the evaluation process.
2. Peer Evaluation: Every question on every test copy was evaluated by a minimum of three students and one TA to ensure fairness and accuracy.
3. Cross-Evaluation: Different groups were assigned to evaluate the test copies of another group to mitigate potential bias in evaluation.
4. Detailed Marking Scheme: A detailed marking scheme with solutions was provided to every student and TA to standardize the evaluation process.

The results of the analysis indicate that the peer evaluation system was highly effective in ensuring question-wise accuracy. By comparing the evaluations done by peers and TAs, we were able to identify areas where the auto answer evaluation system may need improvement. The anonymization process also helped in maintaining the integrity of the evaluation process.

Unique id	Evaluated by	Q1	Q2	Q3		Q4	Q5	Q6		Q7	Q8	Q9		Q10	Q11	Q12		Q13	Q14	Q15		Grand Total
487	Peer 1	0	0	3	3	0	10	0	10	0	0	5	5	0	0	2	2	0	0	3	3	23
	Peer 2	0	0	3	3	0	10	0	10	0	0	8	8	0	0	0	0	0	2	3	5	26
	Peer 3	0	0	3	3	0	10	0	10	0	0	5	5	0	4	0	4	0	2	3	5	27
	TA	0	0	10	10	0	10	0	10	0	0	5	5	0	0	0	0	2	0	3	5	30

Table-1

Evaluation process is done in such a way that different groups are evaluating answer scripts of another group as shown in Table-II

Name of Peer	link to drive	Test Script for Checking	Group for which checking is done	Q NO	Q NO	Q NO	link to google sheet	SHEET NAME
Vrushank Ahire	Please refer Annexure-1	G7	G1	Q1	Q2	Q3	Please refer Annexure-1	T5
Aryan Verma	Please refer Annexure-1	G7	G1	Q4	Q5	Q6	Please refer Annexure-1	T6
Manish Kumar	Please refer Annexure-1	G7	G1	Q7	Q8	Q9	Please refer Annexure-1	T9
Pratik Kumar	Please refer Annexure-1	G7	G1	Q10	Q11	Q12	Please refer Annexure-1	T10
T.Shayan	Please refer Annexure-1	G7	G1	Q13	Q14	Q15	Please refer Annexure-1	T11

Ajmeera Vasuram	Please refer Annexure-1	G7	S1	Q1	Q2	Q3	Please refer Annexure-1	T13
Arav Bhargava	Please refer Annexure-1	G7	S1	Q4	Q5	Q6	Please refer Annexure-1	T14
Prashant Kumar	Please refer Annexure-1	G7	S1	Q7	Q8	Q9	Please refer Annexure-1	T16
Animan Naskar	Please refer Annexure-1	G7	S1	Q10	Q11	Q12	Please refer Annexure-1	T17
Hartik Arora	Please refer Annexure-1	G7	S1	Q13	Q14	Q15	Please refer Annexure-1	T18
Jyoti	Please refer Annexure-1	G7	S1	Q1	Q2	Q3	Please refer Annexure-1	T19
Pranav Menon	Please refer Annexure-1	G7	S1	Q4	Q5	Q6	Please refer Annexure-1	T20
Jitender	Please refer Annexure-1	G7	S1	Q7	Q8	Q9	Please refer Annexure-1	T25
Vanjivaka Sairam	Please refer Annexure-1	G7	S1	Q10	Q11	Q12	Please refer Annexure-1	T26
Aadit Mahajan	Please refer Annexure-1	G7	S1	Q13	Q14	Q15	Please refer Annexure-1	T28

Table-II

Complete data of 132 students appeared in the exam with Peer average and SD:

	STU 1	STU 2	STU 3	STU 4	STU 5	STU 6	STU 7	STU 8	STU 9	STU 10	STU 11	STU 12	STU 13	STU 14	STU 15	STU 16	STU 17	STU 18	STU 19	STU 20	STU 21	STU 22
P1	23	45	90	93	71	33.5	52	108.5	97.5	127	109	60.5	90.5	105.5	108	102.5	95	67.5	59.5	115.5	119	130.5
P2	26	47	94	91	69.5	40	59	121	98.5	128	129	71.5	96.5	114.5	105	123	113	77	65.5	103	100.5	108.5
P3	27	49.5	102.5	100	78	49.5	55	135.5	106.5	127	125.5	82	97	117	106	117	111	93	60	90	95.5	109.5
TA	30	46	99.5	93	71	54	60	120	97	114	108.5	69	87.5	109	110.5	111	98	67	64	114	111.5	114.5
STA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
PEER AVERAGE	25.33333333	47.16666667	95.5	94.66666667	72.83333333	41	55.33333333	121.6666667	100.8333333	127.3333333	121.1666667	71.33333333	94.66666667	112.3333333	106.3333333	114.1666667	106.3333333	79.16666667	61.66666667	102.8333333	105	116.1666667
Standard Deviation	2.886751346	1.931105038	5.582711408	3.947573094	3.816084381	9.242113755	4.787135539	11.06420655	4.460474584	6.683312552	10.77806414	8.855318553	4.643543905	5.212165257	2.428133714	8.75	9.069178574	12.15438876	2.958039892	11.81365735	10.61739924	10.17758976

	STU 23	STU 24	STU 25	STU 26	STU 27	STU 28	STU 29	STU 30	STU 31	STU 32	STU 33	STU 34	STU 35	STU 36	STU 37	STU 38	STU 39	STU 40	STU 41	STU 42	STU 43	STU 44
P1	73.5	116	108.5	129	96	113.5	111	77	93.5	93	104	122.5	111.5	91.5	58	116.5	67	107.5	111.5	87.5	102.5	99
P2	58.5	100.5	95.5	103	86.5	92	84.5	71	85.5	79.5	103	97	105	69	37.5	100	49.5	83	81	64	76	90
P3	60	97	89	91.5	76.5	83	80	56.5	76.5	75.5	82.5	100.5	92	72.5	37	89.5	65.5	111.5	109.5	83	96	110
TA	75	101.5	104.5	111	91.5	97.5	84.5	82.5	97	79.5	94.5	108	102	84.5	56.5	90	65.5	101.5	93.5	71.5	88	94
STA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
PEER AVERAGE	64	104.5	97.66666667	107.8333333	86.33333333	96.16666667	91.83333333	68.16666667	85.16666667	82.66666667	96.5	106.6666667	102.8333333	77.66666667	44.16666667	102	60.66666667	100.6666667	100.6666667	78.16666667	91.5	99.66666667
Standard Deviation	8.703447593	8.391463917	8.797490172	15.76586503	8.370334521	12.81275406	14.15980226	11.19895828	9.123002064	7.652013933	9.958246164	11.3063404	8.117624858	10.45924631	11.56503351	12.62933094	8.280247581	12.60538906	14.38387407	10.7160316	11.41179945	8.655441448

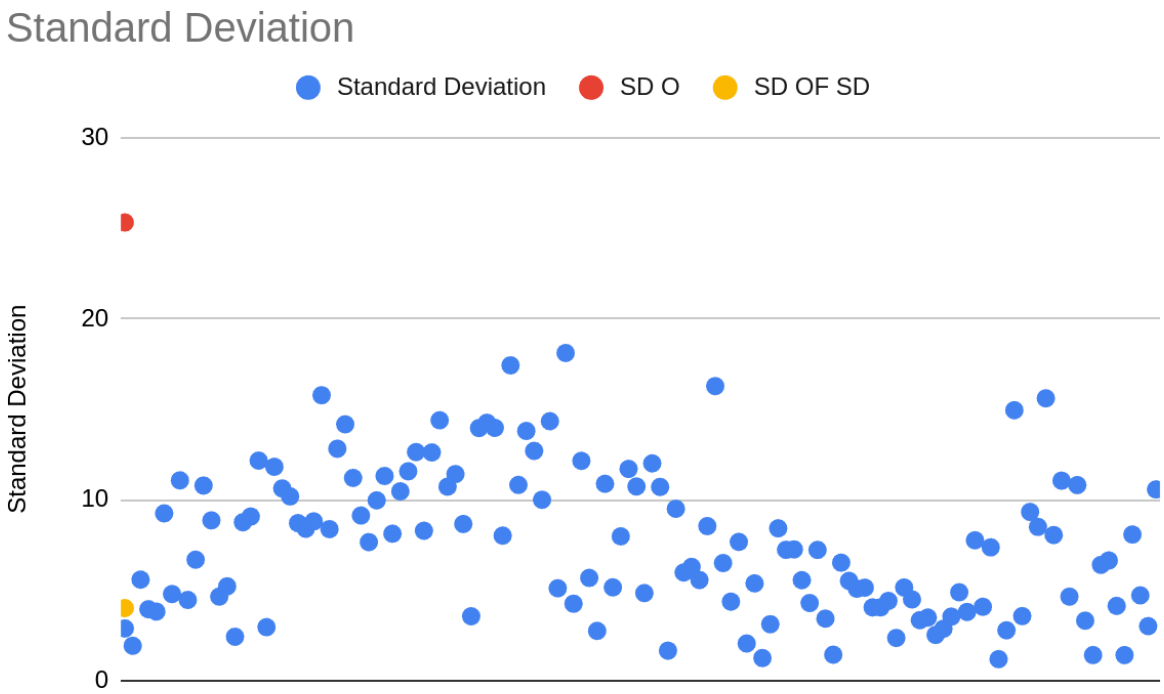
	STU 45	STU 46	STU 47	STU 48	STU 49	STU 50	STU 51	STU 52	STU 53	STU 54	STU 55	STU 56	STU 57	STU 58	STU 59	STU 60	STU 61	STU 62	STU 63	STU 64	STU 65	STU 66
P1	31	78.5	127.5	123	69.5	116.5	98.5	116	112.5	101	123	90	136	114	122.5	101.5	119	78	69.5	85.5	137	96
P2	29	63	97.5	99	52.5	87.5	78.5	86.5	85.5	86	91	81.5	94	106	93	88	112.5	61	57.5	66.5	110.5	73
P3	32	97	125.5	132	69	114.5	99.5	115.5	112	97.5	119.5	91.5	127	106.5	111	97.5	114.5	84	66.5	74	115	94.5
TA	24	77	109	116	61	83.5	82.5	105.5	100.5	79.5	111.5	82.5	116.5	104.5	109	96.5	114.5	84	63	78.5	117.5	83.5
STA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
PEER AVERAGE	30.66666667	79.5	116.8333333	118	63.66666667	106.1666667	92.16666667	106	103.3333333	94.83333333	111.1666667	87.66666667	119	108.8333333	108.8333333	95.66666667	115.3333333	74.33333333	64.5	75.33333333	120.8333333	87.83333333
Standard Deviation	3.559026084	13.95453929	14.24415085	13.96424004	8.010409894	17.4164673	10.81280106	13.792359	12.69104015	9.991663192	14.33236431	5.105144464	18.0988878	4.252450274	12.14066857	5.677073777	2.75	10.87428159	5.153882032	7.97260936	11.6975781	10.72769003

	STU 67	STU 68	STU 69	STU 70	STU 71	STU 72	STU 73	STU 74	STU 75	STU 76	STU 77	STU 78	STU 79	STU 80	STU 81	STU 82	STU 83	STU 84	STU 85	STU 86	STU 87	STU 88
P1	107	75.5	122.5	93.5	122.5	106.5	122	116.5	118	107	111.5	110.5	107	63	45	11.5	77	83	129	85	95	91.5
P2	103	46.5	100	93.5	112	93.5	109.5	105.5	98.5	67.5	100	115	108.5	63.5	45	14.5	80.5	83.5	127	70	95.5	87
P3	113.5	57	121	95	114	104.5	116	117.5	114.5	86	115	121	115	67.5	55	12.5	78	95	129	85.5	107	96.5
TA	103.5	58.5	108.5	91	99.5	98	108.5	115.5	108.5	91	111	117	96.5	65.5	53.5	13	73	74.5	114	82	98.5	88
STA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
PEER AVERAGE	107.8333333	59.66666667	114.5	94	116.1666667	101.5	115.8333333	113.1666667	110.3333333	86.83333333	108.8333333	115.5	110.1666667	64.66666667	48.33333333	12.83333333	78.5	87.16666667	128.3333333	80.16666667	99.16666667	91.66666667
Standard Deviation	4.839077047	12.00260388	10.70046720	1.658312395	9.495613022	5.977387947	6.284902545	5.580275773	8.537905676	16.27049989	6.498397238	4.366062299	7.664854858	2.05649378	5.375484474	1.25	3.119161212	8.416254115	7.228416147	7.25	5.552777083	4.291464397

	STU 89	STU 90	STU 91	STU 92	STU 93	STU 94	STU 95	STU 96	STU 97	STU 98	STU 99	STU 100	STU 101	STU 102	STU 103	STU 104	STU 105	STU 106	STU 107	STU 108	STU 109	STU 110
P1	111	106.5	104.5	54	45.5	93	105.5	60	85	65	115.5	74.5	107	65	106.5	110	104	116	108	52	83	121
P2	122.5	99	104	54.5	36.5	97.5	109.5	68.5	86.5	69	120.5	82	105.5	69	106	111	107	108.5	102	55.5	83	116
P3	119.5	99.5	107	68	42	104	111	64.5	84	62.5	117.5	87	115	64.5	101.5	111.5	105	111.25	110.5	48.5	80.5	126
TA	107	102	104	57.5	49.5	93.5	99.5	68.5	93	72.5	115.5	80.5	112.5	71.5	110	106	110.5	108.5	113.5	57	97.5	121
STA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
PEER AVERAGE	117.6666667	101.6666667	105.1666667	58.83333333	41.33333333	98.16666667	108.6666667	64.33333333	85.16666667	65.5	117.8333333	81.16666667	109.1666667	66.16666667	104.6666667	111.8333333	105.3333333	111.9166667	106.8333333	52	82.16666667	121
Standard Deviation	7.22264956	3.4278273	1.436140662	6.519202405	5.513241636	5.082650227	5.137687547	4.049176871	4.049176871	4.406434689	2.362907813	5.14781507	4.490731195	3.341656276	3.488074923	2.528866484	2.868652413	3.537742920	4.881939505	3.799828583	7.756717519	4.082482905

	STU 111	STU 112	STU 113	STU 114	STU 115	STU 116	STU 117	STU 118	STU 119	STU 120	STU 121	STU 122	STU 123	STU 124	STU 125	STU 126	STU 127	STU 128	STU 129	STU 130	STU 131	STU 132
P1	93.5	123.5	76.5	94.5	84	105	121.5	108.5	32.5	123.5	77	114.5	81	22	90.5	103	126	76	85	57.5	46.5	99
P2	103	123	79.5	83.5	76	105	120	112	43	123	73	121.5	78	21	101.5	101	132	75	82.5	58.5	44	107.5
P3	108	124	70	101	87.5	119	138.5	135	46.5	138	81.5	130	81	22	97.5	111.5	125	78	105	67.5	39.5	118.5
TA	110	121.5	72	85	80	98.5	124	98	51.5	111	71	104.5	74	10	87.5	95.5	127.5	75	85.5	64	45	94.5
STA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
PEER AVERAGE	101.5	123.1666667	75	93	83.5	109.6666667	126.6666667	118.5	40.66666667	128.1666667	77.16666667	122	80	21.66666667	96.5	105.1666667	131	76.33333333	97.5	61.16666667	43.33333333	108.3333333
Standard Deviation	7.364045537	1.190238071	2.783882181	14.93876389	3.567790771	9.321793461	8.495096625	15.59580606	8.045443845	11.04618899	4.643543905	10.80412373	3.316624790	1.414213562	6.396613687	6.639528096	4.130677091	1.414213562	8.072587359	4.714791172	3.013858867	10.58231509

Table-III



OBJ OBJ

Plots showing complete evaluation data and it's analysis (as per data obtained in Table-III)

7. Conclusion:

The proposed peer evaluation system represents a significant advancement in addressing the challenges associated with manual evaluation in large classes and Massive Open Online Courses (MOOCs). By harnessing the collective intelligence of students, this system offers a viable solution to distribute the evaluation workload, alleviate the burden on instructors, and expedite feedback delivery to students. Furthermore, it has the potential to enhance student engagement, foster critical thinking skills, and promote a collaborative learning environment.

The effectiveness and scalability of this system make it a promising avenue for further research and implementation in educational settings. Future studies could focus on evaluating the impact of the peer evaluation system on student learning outcomes, instructor workload, and overall course satisfaction. Additionally, exploring ways to optimize the system's design and integration within existing educational technology frameworks could further enhance its utility and adoption.

In conclusion, the proposed peer evaluation system represents a valuable contribution to the field of education technology. Its innovative approach to evaluation not only addresses current challenges but also opens up new possibilities for enhancing the learning experience in large classes and MOOCs. By continuing to refine and explore this system, educators can potentially transform the way assessments are conducted and create more engaging and effective learning environments for students.

8. Future Work:

1. Developing Software: We will develop software with a user-friendly interface to automate the peer evaluation process. The software allows students to easily access and evaluate their peers' answer sheets, provides prompts for feedback, and calculates scores based on predefined criteria.

9. References:

1. *LearnEval Peer Assessment Platform: Iterative Development Process and Evaluation (Gabriel Badea and Elvira Popescu, Senior Member, IEEE).*
2. *Incentive Design in Peer Review: Rating and Repeated Endogenous Matching (Yuanzhang Xiao, Florian Dorfler, and Mihaela van der Schaar, Members/Fellows, IEEE)*
3. *A Framework for Teaching Evaluation (Divya Nalla, Department of Computer Science, Nalla Malla Reddy Engineering College).*
4. *Learning Management System User Guide: Self & Peer Assessment (University of Melbourne).*
5. *A Graph Analysis Method to Improve Peer Grading Accuracy for Blended Teaching Courses (Xing Du, Xingya Wang, and Yan Ma).*
6. *Fostering Peer Learning With a Game-Theoretical Approach in a Blended Learning Environment (Seyede Fatemeh Noorani , Mohammad Hossein Manshaei , Mohammad Ali Montazeri, and Behnaz Omoomi)*

Annexure-1 (Step-by-Step process of running peer evaluation)

INTRODUCTION

The learning management system (LMS) like Moodle, Canvas etc has been used only for assignment checking.

We have designed an Automated Peer Evaluation System with the aim ` to save crucial and important work-hours of the teachers and TAs which can be used in more productive outcomes to improve the education delivered to the students.

This will also enhance the fairness system amongst the students and students will get to know about the feedback of their performance.

The students will also get a hang of the deeper level of the subject while evaluating others. The best way to learn a subject is by teaching it.

Thus the peer evaluation will work in multiple ways to improve the education outcomes of the course.

The System

The System has three components, Question Paper Printing which custom prints the Unique ID of the student on their sheet to ensure that the students get the sheet they have been assigned to. The next part is the Pre-Evaluation system which sets up the peer evaluation groups and sends the details to each student for their specific evaluation by mail. The mail contains the link to the drive in view-only mode and a link to the evaluation spreadsheet in the

Components

The system has three major components, including the Printing Manager, the Pre-Evaluation Script, and the Post-Evaluation Script encompassing the entire testing process. The making of the test paper, collecting them and then evaluating the answer scripts, and then consolidating the results.

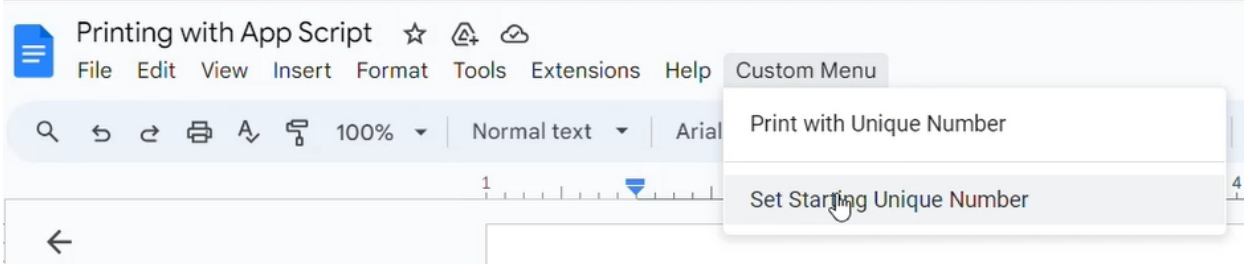
> **The Peer Evaluation System components are intuitive and easy to use. Automation should be about ease of use and functionality.**

QUESTION PAPER PRINTING

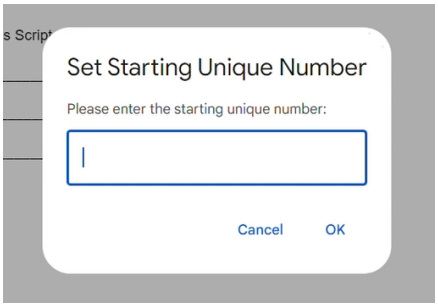
Introduction and Usage

The Question Paper Printing system is hosted on Google Docs and uses add-ons to make the printing process personalized for each candidate to ensure that they get the script that is intended for them. This helps in tracking the candidate throughout the entire evaluation process.

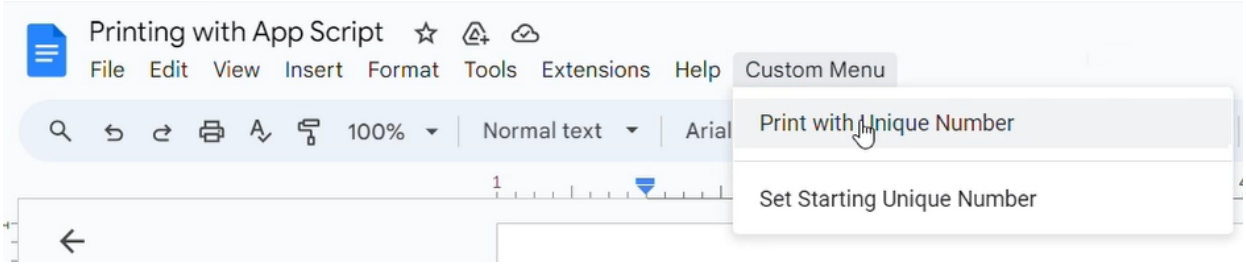
To start, we go to the Google Doc and make the question paper, giving the space required for writing the answers to the questions. Then we go to the ***Custom Menu*** → ***Set Starting Unique Number***



This will open the dialogue box to enter the starting unique number for each student. This will increment automatically and only the first number is required to be entered.



Next, we need to go to the ***Custom Menu*** → ***Print with Unique Number*** to print the question paper sheets. These will be done in batches and each batch will have a different unique ID corresponding to the candidate.



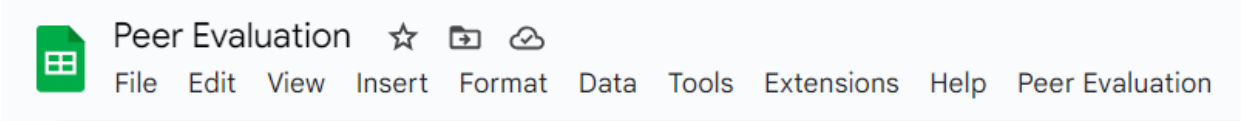
PRE-EVALUATION SYSTEM

Introduction

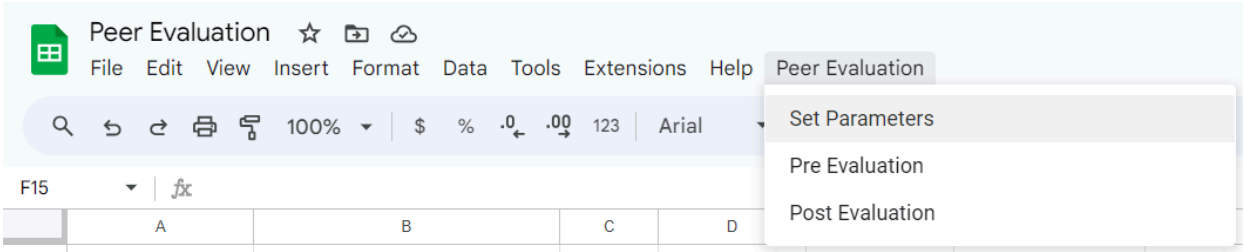
The Pre-Evaluation System script should be run after the test has been taken and the scanned files are in the Source Folder in Google Drive. These scanned files are the answer scripts of the candidates.

Guide to Usage

The system has a custom menu on the Navigation Bar of the Peer Evaluation Google Sheet named Peer Evaluation.

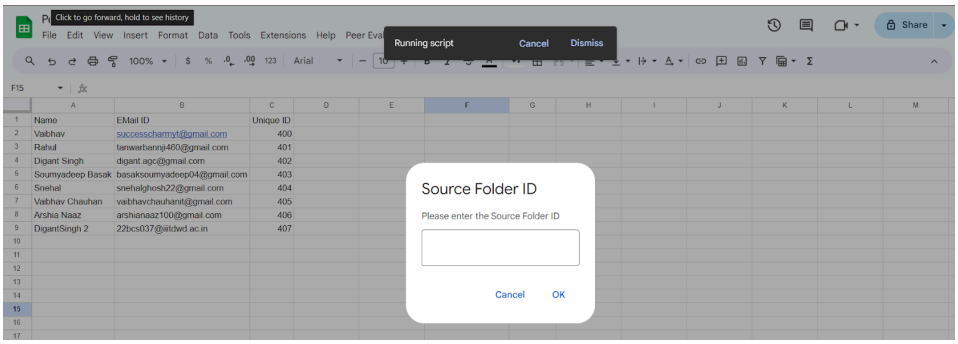


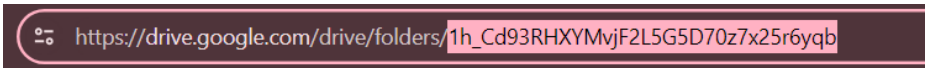
The Peer Evaluation system needs **4 system inputs** which are fed into the system by the ***Peer Evaluation → Set Parameters***



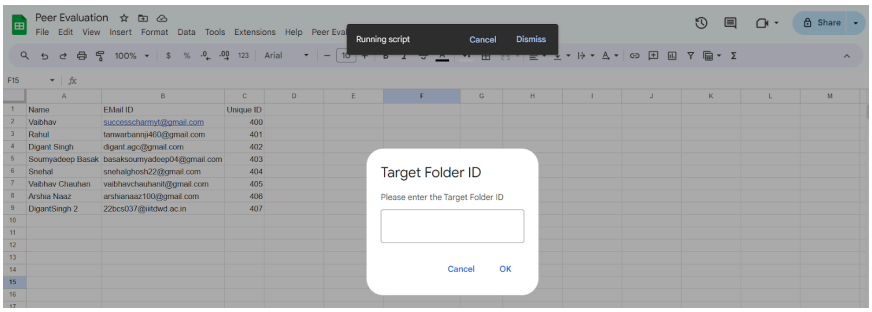
You will get 4 dialogue boxes asking for multiple responses.

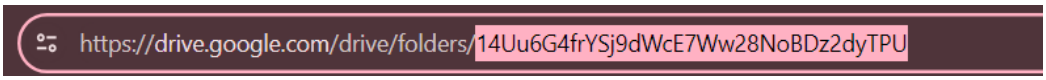
First Step: Source Folder ID



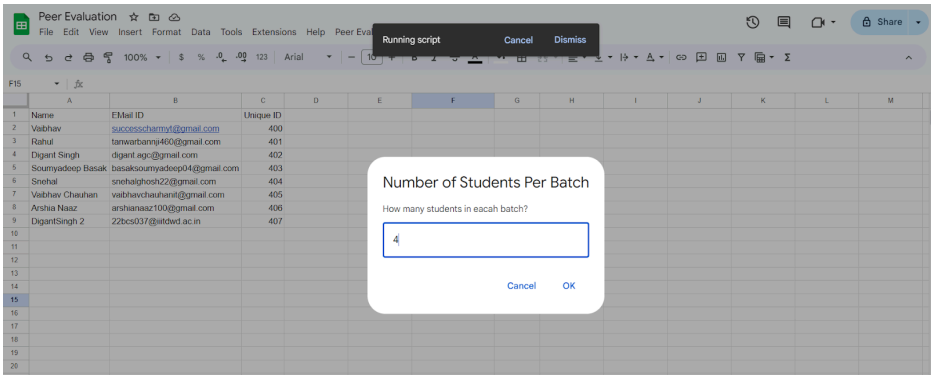


Next Step: Target Folder ID

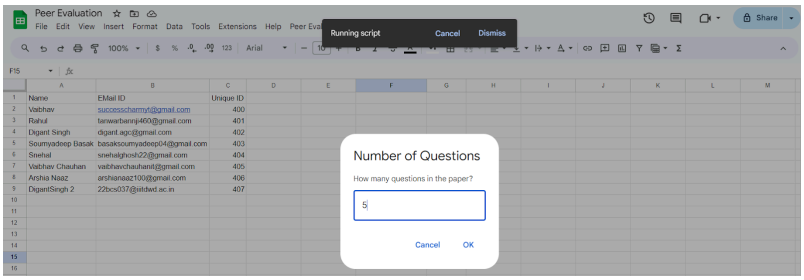




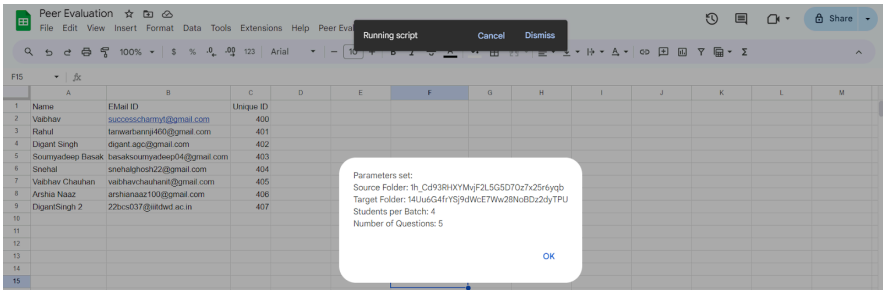
Next Step: Number of Students Per Batch



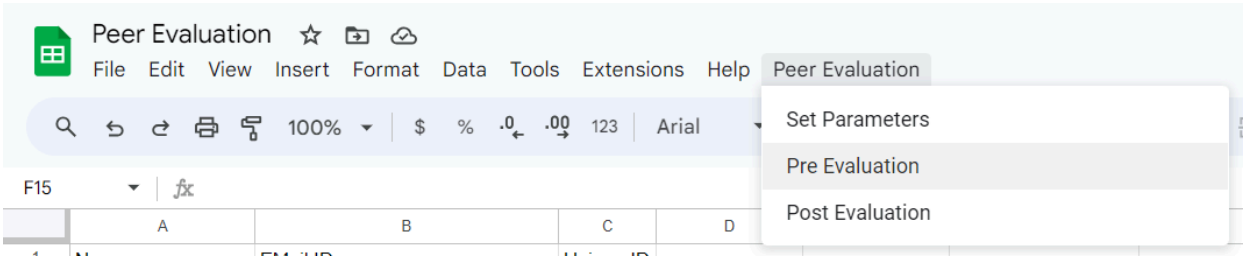
Final Step: Number of Questions



Then the confirmation box shows the options set by the user

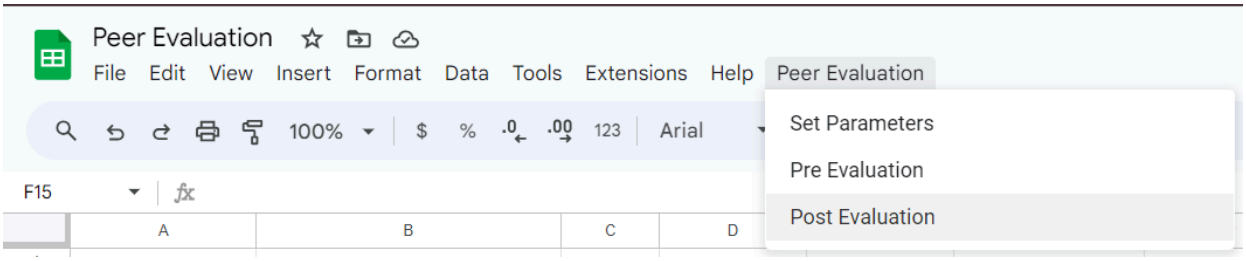


Next we go to the ***Peer Evaluation → Pre Evaluation***



This will run all the scripts for the Pre Evaluation processes. This has simplified the process since the last version where we had to run the script manually. This button type setting helps simplify the process for the system administrators.

Next we go to the ***Peer Evaluation → Post Evaluation.*** This should be run only after all the peers have done their evaluation. This will usually be after some days of the test, although timelines may vary.



This will consolidate the resDebuggingults from all the spreadsheets that are distributed amongst the peers for the evaluation. Once consolidated in one sheet the final results will be calculated and the statistics of the results will be generated for the reference of the administrators.

Debugging

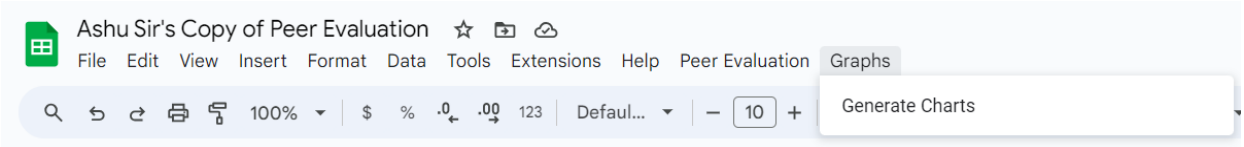
The code in the Google Sheet Apps Script has a lot of error handling enabled by try-catch blocks which have the error codes filled in them. Any error can be fixed or handled by reading these error codes in the execution logs and following them to the script and making necessary corrections. Although, this will seldom be required.

GENERATING GRAPHS

The insightful graphs are to be generated to find out the metrics about the peer evaluation that has been done by the students. These graphs tell us about the standard deviations and other important factors that will help the admins to reach fruitful conclusions.

Guide to Usage

Go to the Google Sheet and find the ***Graphs*** button on the main menu. ***Graphs*** → ***Generate Charts.*** This will generate all the visualizations in the “Evaluations Results” sheet.



REPLICATION

When replicating the sheet and system for another test, the Google Sheet will have to be copied by ***File*** → ***Make a copy*** which will also make a copy of the Apps Script. You will then need to run the ***onOpen*** function from ***Buttons.gs*** Script. This will set up everything as you want for the administrator to start off without any hiccups.