Journal of Artificial Intelligence Research 79 (2024) 1001-1046 Submitted 08/2023; published 03/2024

# Removing Bias and Incentivizing Precision in Peer-grading

**Anujit Chakraborty** CHAKRABORTY@UCDAVIS.EDU
*University of California Davis*

**Jatin Jindal** JATINJINDAL369@GMAIL.COM
*Google (India)*

**Swaprava Nath** SWAPRAVA@CSE.IITB.AC.IN
*Indian Institute of Technology Bombay*

## Abstract

Most peer-evaluation practices rely on the evaluator's goodwill and model them as potentially noisy evaluators. But what if graders are competitive, i.e., enjoy higher utility when their peers get lower scores? We model the setting as a multi-agent incentive design problem and propose a new mechanism, `PEQA`, that incentivizes these agents (peer-graders) through a *score-assignment rule* and a *grading performance score*. `PEQA` is designed in such a way that it makes grader-bias irrelevant and ensures grader-utility to be monotonically increasing with the grading-precision, despite competitiveness. When grading is costly and costs are private information of the individual graders, a modified version of `PEQA` implements the socially optimal grading-choices in equilibrium. Data from our classroom experiments is consistent with our theoretical assumptions and show that `PEQA` outperforms the popular *median* mechanism, which is used in several massive open online courses (MOOCs).

## 1. Introduction

A peer-evaluation process aggregates assessments from peers to judge the quality of submitted work. Scientific communities use peer-evaluation for reviewing the quality of articles and grant proposals (Campanario, 1998). Coursera and EdX, that offer Massive Open Online Courses (MOOCs) to 94 million learners[1], use peer-grading to evaluate submitted assignments. Many in-person classes are also adopting it and its growing popularity can be explained by the following three reasons. First, it simplifies and accelerates the evaluation and grading process. Second, it improves learning outcomes of the participating students (Sadler & Good, 2006). Third, it easily scales to large classes.

There is, however, a scope for skepticism about the accuracy of peer-graded outcomes. A grader might be unmotivated to evaluate diligently when peer-grading is effort-intensive and unincentivized. She might also be biased in her evaluations if she cares about her relative success within peers.[2] This creates perverse incentives for peer-graders. In an anonymous survey that we ran on the students of a reputed technical institute in India, 49% of the

[1] Numbers from Coursera's and EdX's 2020 impact reports.

[2] When students are evaluated on a curve, students naturally care about their relative performance vis-à-vis peers. Even when evaluated on an absolute grading scale, students care about their relative performance due to the role it plays in admission into jobs or higher studies.

549 respondents expected that their fellow students would grade aggressively to reduce the scores of others, and thereby try to improve their relative class-ranking.[3]

We study the problem of incentivizing peer-graders, while allowing for competitive and strategic behavior. In our model, students take an exam[4] and then peer-grade each others' exams. Thus, every student has dual roles: (i) the student role, where she takes an exam that gets evaluated, and (ii) the grader role, where she evaluates others. As is the norm in most MOOC courses that utilize peer-grading, a student's total course-score is the sum of their own exam score (aggregated from peer-reports) and a score based on their peer-grading performance. To model competitive students, we assume that their utility is linearly increasing in their total course-score and linearly decreasing in their peers' total course-scores.

To model strategic grading, we adapt the widely used $\mathbf{PG}_1$ model of Piech et al. (2013) to a strategic environment. Piech et al. (2013) introduce $\mathbf{PG}_1$ as a statistical model of peer-grading and use it for estimating and correcting for grader bias and reliability (inverse of variance) in a large-scale data mining exercise.[5] $\mathbf{PG}_1$ assumes that each paper has a true score and any peer-grader's bias and reliability are drawn randomly from a known distribution. We instead assume that each peer-grader *strategically chooses* the reliability of the independent, noisy signals that they observe about the true score. By choosing a higher reliability, they can observe a more accurate signal. Graders can then decide to add a bias of their choice to their observed signal while reporting their assessment. Graders who care about their relative success within peers might purposefully bias their evaluations. They may also choose to receive less reliable signals.

What is the set of desiderata one could ask for a mechanism in this setup? At a minimum, the mechanism should make bias irrelevant and incentivize reliable grading. Also, the aggregation rule over peer reports should assign a final score that is close to the true score. To simplify, we initially assume that more reliable grading (lower variance) does not come at an extra cost to the peer-grader.

We propose a new mechanism, `Peer Evaluation with Quality Assurance` (`PEQA`), that ensures that (Theorem 1):

- ▷ Assigned scores and grader's utility are *bias-insensitive* (Definition 2). Thus, graders have no incentive to introduce a bias, and bias does not affect the grading process.
- ▷ Choosing higher grading reliability ensures monotonically higher utility to the grader, despite her competitiveness. This holds for all actions of her co-graders. (*reliability monotonicity*, Definition 3).

How should one aggregate the peer reports to ensure accuracy of assigned scores? A candidate score-assignment function is one that minimizes the expected squared error, i.e, the squared distance between the assigned score and the true score of a paper. In Equation (30) of Appendix C, we show that under the distribution-assumptions of Piech et al. (2013), a reliability-weighted average of the *de-biased* peer reports minimizes squared error.

[3] Ideally, we would also ask students if they themselves would do the same while grading peers, but students are likely to under-report such activity.

[4] We use the terms *exam*, *paper*, and *answerscript* interchangeably in this paper.

[5] The authors mention: "we present the largest peer grading networks analysed to date with over 63,000 peer grades. ... we present, in order of increasing complexity, three statistical models that we have found to be particularly effective". $\mathbf{PG}_1$ is the first one of this three models.

PEQA's score-assignment function closely approximates the squared-error minimizer (see Appendix D), while uniquely and flexibly satisfying the monotonic relation between utility and reliability (Theorem 2).

In Section 6, we address if PEQA satisfies the more ambitious desiderata of implementing a "preferred level" of grading among competitive peer-graders, while accounting for the cost of grading reliably. We assume that students face an additional disutility (cost) from grading that increases with their reliability. How much effort should one ask students to exert? Reliability is desirable, but it might be prohibitively costly for students to spend all their time on grading! We define the *net student welfare* (Equation (11)) from the game as the difference between the social benefit and the aggregate cost of reliability. Under this setup, we show that:

- ▷ A modified version of PEQA implements Nash equilibria of the peer-grading game (with private costs) in which graders grade at the welfare-optimal level of reliability (Theorem 4).
- ▷ The modified PEQA maintains the same ranking among the students as the original PEQA (Lemma 2).

The close connection between a grader's grading performance score and her marginal contribution to the student welfare, under PEQA, makes these possible. Alternative performance bonus schemes that do not use the idea of marginal social contributions, e.g., the one that compares and punishes graders whose peer-grading scores differ from the true scores, cannot satisfy these properties.

How does the mechanism PEQA work? The teaching staff evaluate a small subset of the total number of papers (called *probes*). Each grader is assigned $K > 2$ ($K$ even) papers (with $K/2$ probes) and they never grade their own paper. The probes are used to estimate the biases and reliabilities of the peer-graders so that the non-probe papers can be properly calibrated.[6] The peer-graders cannot tell apart the probes from the non-probes. PEQA compares the grader's and the teaching staff's evaluations of the probes to estimate each grader's bias and reliability. This requires two identifying assumptions: that the teaching staff can observe the true scores on the probe papers, and, that the graders grade identically on probes and non-probes. The estimated grader-bias is subtracted from the peer-reports to de-bias the reports. PEQA's score-assignment function assigns a weighted average of the de-biased grader-reports, with the weights being the inverse square-root of the estimated grader-variance. Thus, reports from high variance graders play a smaller role in the finally assigned score.

We allow students to raise regrading requests, after seeing their score. The teaching staff regrade such papers and assign them the true score. We assume that students raise these requests in a self-serving way: only when the student knows that her peer-graded score was lower than the true score. In the discussion following Theorem 1 in Section 5, we explain how PEQA utilizes the regrading requests and competitive preferences to incentivize reliable grading.

The schematic diagram of the stages of PEQA is shown in Figure 1.

Theoretical assumptions are often only an approximation of reality. To test some of our theoretical assumptions and to see how easily our mechanism could be implemented

[6] This is in spirit of the mechanism design with verification (e.g., (Li, 2020)) where the incentives of the agents depend on the agents' performance on the verification.

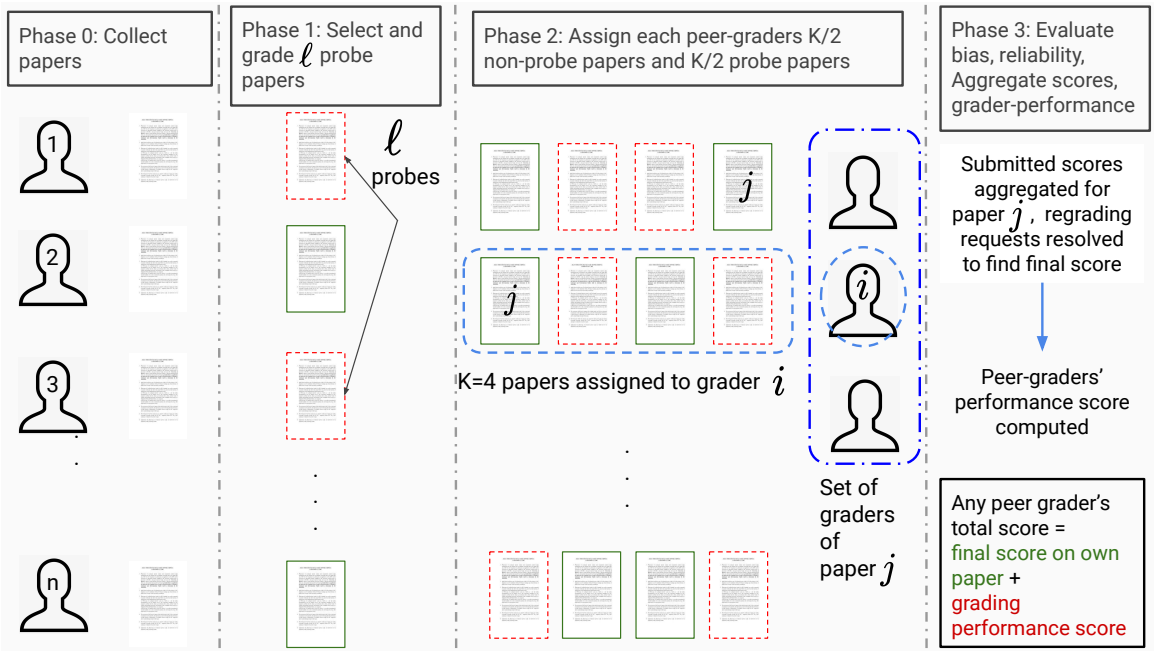



Figure 1: Schematic diagram of the PEQA mechanism decomposed into *four* phases. A typical non-probe paper is denoted by $j$ here.

in practice, we ran a classroom experiment (Section 7). Students enrolled in a computing course were asked to peer-grade a weekly class-quiz and were incentivized by PEQA. We independently graded all the exams to evaluate their true-scores. Compared to the true scores, PEQA assigned scores that were remarkably accurate: only 1 out of 41 sub-quizzes had a wrong score. Thus, despite the simplifying theoretical assumptions, the mechanism does very well consequentially.

Data from our PEQA sessions (Tables 1 to 3) is consistent with two of our assumptions.

1. The bias and variance were indeed not different across probes and non-probes under PEQA: students were not able to discern one from the other (Hypothesis 4).
2. Grade-manipulations, whenever present, reduced scores instead of inflating scores. This rejects the existence of collusive (i.e., the opposite of competitive) graders (Hypothesis 1).

We ran a second competitive session under a Median mechanism, which is currently the most popular mechanism used in MOOCs.[7] In our experiments, PEQA mechanism outperformed Median mechanism in terms of allocating accurate final scores (Hypothesis 3). These differences were statistically significant.

## Related Work

The existing research on peer-evaluation mechanisms can be broadly divided into three strands. The first strand of literature abstracts away any strategic motives of the peer-evaluators. Instead of providing a mechanism to incentivize strategic evaluators, they propose how the grader reports could be aggregated efficiently (Caragiannis, Krimpas, &

[7] Coursera and EdX use the median score for aggregation, as reported on Coursera and EdX websites.

Voudouris, 2015, 2020; Cho & Schunn, 2007; De Alfaro & Shavlovsky, 2014; Fiez, Shah, & Ratliff, 2020; Hamer, Ma, & Kwong, 2005; Kulkarni, Socher, Bernstein, & Klemmer, 2014; Noothigattu, Shah, & Procaccia, 2021; Paré & Joordens, 2008; Piech et al., 2013; Raman & Joachims, 2014; Shah, Bradley, Parekh, Wainwright, & Ramchandran, 2013; Wang, Stelmakh, Wei, & Shah, 2021; Wright, Thornton, & Leyton-Brown, 2015; Zarkoob, d'Eon, Podina, & Leyton-Brown, 2022).

The second strand of literature is based on *peer-prediction* approaches. These mechanisms incentivize *coordination* on similar evaluation reports by punishing evaluations that do not match each other. Thus, they do not necessarily incentivize *accuracy* (Dasgupta & Ghosh, 2013; Dhull, Jecmen, Kothari, & Shah, 2022; Faltings, Li, & Jurca, 2012; Jurca & Faltings, 2009; Lev, Mattei, Turrini, & Zhydkov, 2023; Miller, Resnick, & Zeckhauser, 2005; Prelec, 2004; Shnayder, Agarwal, Frongillo, & Parkes, 2016; Waggoner & Chen, 2014; Witkowski, Bachrach, Key, & Parkes, 2013; Witkowski & Parkes, 2013). Any such mechanism introduces uninformative equilibria alongside the truth-telling one (Jurca & Faltings, 2009; Waggoner & Chen, 2014).[8] More recent developments make the truthful equilibrium Pareto dominant, i.e., the truthful equilibrium is (weakly) more rewarding to every agent than any other equilibrium (Dasgupta & Ghosh, 2013; Kamble, Shah, Marn, Parekh, & Ramachandran, 2015; Radanovic & Faltings, 2015; Shnayder et al., 2016; Witkowski & Parkes, 2013). Shah (2022) provides a contemporary survey on the current solutions and challenges in peer-review.

The final strand consists of *hybrid* approaches where the true quality of some of the peer-assessed material can be found, for e.g, via evaluating a part of the materials by the mechanism designer (teaching staff in case of MOOCs) herself. Graders are then rewarded for agreement with the designer-agreed report (Dasgupta & Ghosh, 2013; Gao, Wright, & Leyton-Brown, 2016; Jurca & Faltings, 2005). Our mechanism also utilizes the feature that the true scores on a small subset of assignments can be revealed at a small cost. However, additionally, we address new and practical features of the peer-grading probem: we allow for competitive graders, we solve the efficient allocation problem under costly grading, and we allow regrading requests.

Alon, Fischer, Procaccia, and Tennenholtz (2011) and Holzman and Moulin (2013) study situations where peers have to choose a subset amongst themselves for a reward. The challenge here is to incentivize the peers to reveal their private information unselfishly. In particular, the goal is to guarantee that what peers report does not affect their chances of winning or getting selected. In these settings, there is no need to incentivize peers to gather information that is 'objective' (e.g., true score on an exam) and verifiable at a cost. There is also no need to ensure that peers enjoy higher utility when their gathered information is more precise. Finally, peers are purely selfish: they do not care about who wins in case they do not win themselves. Thus, by debriding the reports from personal winning chances, the mechanism makes the peers indifferent between all reports.

Cai, Daskalakis, and Papadimitriou (2015) consider a setting where data-sources (e.g., human labelers) can be paid monetarily to get their estimation of $f(x_i)$ at points $x_i$ allocated to them. The end goal is to estimate an exogenously provided $f$ using a given estimator $\hat{f}$. Data-sources can observe a noisy version of $f(x_i)$ with the noise decreasing in their effort,

[8] In particular, when the information is costly to obtain, it is generally easier for the agents to resort to coordinating on an uninformative low-effort equilibrium.

and they maximize the difference between the payment and the cost of the effort. They show that under their VCG-like payment mechanism and the assumption of a "well-behaved" $\hat{f}$, the dominant strategy for a data-source is to reveal its observation correctly and always participate in the data-providing exercise. Cai et al. (2015)'s data-sources naturally have no competitive preferences, like our graders do.

## 2. Peer-grading Mechanism

### 2.1 Definition

Each student $i \in N = \{1, \ldots, n\}$ has written an exam, and is also a participant in the peer-grading process. Thus $N = \{1, \ldots, n\}$ represents both the set of papers to be graded and the set of graders. We use $i$ as the index for a grader and $j$ as the index for a paper. For simplicity of exposition, we assume that each paper has only *one* question for evaluation. This is not a limitation. In Section 8, we discuss how the analysis of this section can be easily extended to multiple questions per exam.

Our mechanism would instruct the teaching staff to evaluate a fixed number $\ell (<< n)$ of these papers so that their true grades are known. These papers are called the *probe* papers. Let $G(j)$ denote the set of peer-graders of paper $j$ and $G^{-1}(i) := \{k \in N : i \in G(k)\}$ denote the set of papers assigned to evaluator $i$. The set $P_i \subset G^{-1}(i)$ and $NP_i = G^{-1}(i) \setminus P_i$ denotes respectively the probe and non-probe papers assigned to $i$. Both true and reported scores belong to $\mathbb{R}$. The co-graders of individual $i$ are $CG_i = \cup_{j \in NP_i} G(j) \setminus \{i\}$. We assume that the co-graders of $i$ grade at least one common non-probe paper with $i$.

Assuming that peer-reported scores are real numbers, a *peer-grading mechanism* $M$ is the tuple $\langle G, \mathbf{r}, \mathbf{t} \rangle$, where

- $G$ is the *assignment* function $G : N \to 2^N$ that maps papers to graders.
- $\mathbf{r} : \times_{j \in N} \mathbb{R}^{|G(j)|} \to \mathbb{R}^n$ is the *score-assignment* function, where the $j$th component $r_j(\cdot)$ is the function assigning the final score of paper $j$ based on the scores reported by $G(j)$.
- $\mathbf{t} : \times_{i \in N} \mathbb{R}^{|G^{-1}(i)|} \to \mathbb{R}^n$ is the *peer-grading performance score* function, where the $i$th component $t_i(\cdot)$ is the function that yields the peer-grading performance score to grader $i$.

Since every student $i$ has dual roles in peer-grading as explained in Section 1, $r_i$ and $t_i$ are the mechanism-assigned scores corresponding to her student and grader roles. For example, in a course that has 80 points on the exam and 20 points on peer grading performance, a student might score $r_i = 60$ and $t_i = 15$ on those two respectively. Her total course-score would be 75 out of 100.

### 2.2 Model of the True and Reported Scores

Let $\mathcal{F}(0, 1)$ be any general distribution with a support of $(-\infty, \infty)$, a *differentiable* density function $f(\cdot)$, mean zero, and variance one. We use $\mathcal{F}(a, b^2)$ to denote the distribution of the random variable $a + bX$ where $X \sim \mathcal{F}(0, 1)$.

We generalize the $\mathbf{PG}_1$ model of true score, bias, and reliability (Piech et al., 2013) to a strategic environment. We make two major changes. First, we replace their assumptions of normality with distribution $\mathcal{F}(a, b^2)$. Second, instead of assuming that bias and reliability are drawn randomly and independently from Normal and Gamma distributions respectively,

we make each a strategic choice by the peer-graders. Subject to these changes, the following features in our model resemble the $\mathbf{PG}_1$ model.

- ▷ The true score $y_j$ for paper $j$ is distributed as $\mathcal{F}(\mu, 1/\gamma)$, for all $j \in N$. This distribution is known from historical data of past examinations.
- ▷ Peer-graders do not see $y_j$ but after they choose their reliability $\tau_i \in \mathbb{R}_{>0}$ and bias $b_i$, they observe an independent draw from $\mathcal{F}(y_j, 1/\tau_i)$. Higher is $1/\tau_i$, noisier is the draw. We will use $1/\tau_i$ and $\sigma_i^2$ interchangeably.
- ▷ For the same grader $i$, the signals from two different papers $j_1$ and $j_2$ are independent draws from $\mathcal{F}(y_{j_1}, 1/\tau_i)$ and $\mathcal{F}(y_{j_2}, 1/\tau_i)$ respectively.
- ▷ Graders then add the bias $b_i \in \mathbb{R}$ to the signal before reporting. The reported score of paper $j$ by grader $i$ is $\tilde{y}_j^{(i)}$. Conditional on the true score $y_j$, it is distributed as $f(\tilde{y}_j^{(i)}|y_j) \sim \mathcal{F}(y_j + b_i, 1/\tau_i)$. Thus, $\tilde{y}_j^{(i)} = y_j + b_i + \sigma_i e_{ij}$ where $e_{ij} \sim \mathcal{F}(0, 1)$.
- ▷ We have used the same distribution $\mathcal{F}$ for both the true scores $y_j$ and the score observed by the grader $i$, i.e., $\tilde{y}_j^{(i)}$, to keep the model simpler and similar to $\mathbf{PG}_1$. However, this is not critical to our results. In particular, (a) we can have two different distributions for these two sets of random variables, and (b) the distribution of the observed score $\tilde{y}_j^{(i)}$ can be different from each grader $i$. None of these will affect the main conclusions of this paper.
- ▷ We have overloaded the notation $\tilde{y}$ to denote both individual grades and grade vectors. The grades of a paper $j$ given by its graders $G(j)$ is denoted by $\tilde{\mathbf{y}}_j^{G(j)} = (\tilde{y}_j^{(i)}, i \in G(j))$.

The dynamics of the grading process is shown in Figure 2. We define reliability as the

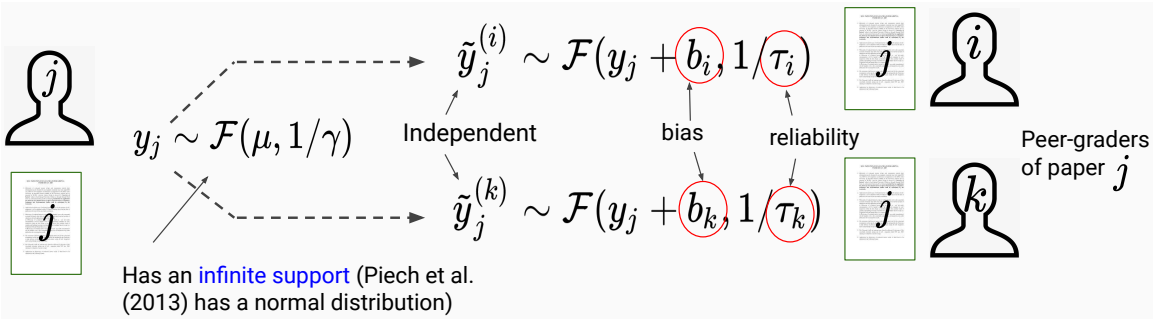



Figure 2: Peer-reports' generation process.

inverse of noise variance. Bias originates from a strategic manipulation or from non-strategic (generous or strict) grading-habits. In this paper, we would assume that the grader chooses her bias and reliability.

We assume that a grader grades all papers (probes and non-probes) with the same bias and reliability. This assumption is natural if the graders cannot identify the probes from the non-probes. Additionally, if a class uses multiple assignments (exams and problem sets) over the whole semester, then, the performance in any anonymized peer-graded assignment $j \in G^{-1}(i)$'s that $i$ grades, reveals very little to $i$ about $j$'s identity and overall class-rank over the semester. Thus, $i$ might feel equally competitive across all assignments she evaluates. We use the shorthand $\theta_i = (b_i, \tau_i) \in \mathbb{R} \times \mathbb{R}_{>0}$ to denote grader $i$'s strategic choices.

### 2.3 Other primitives of our mechanism

We have already defined a general peer-grading mechanism in Section 2.1. In this section, we fine-tune the $\langle G, \mathbf{r}, \mathbf{t} \rangle$ functions for our proposed mechanism.

Paper assignment rule $G^*(\cdot)$

Every paper is graded by at least one grader, and every grader grades at least two probe and one non-probe papers. Thus, (a) $G^*(j) \neq \emptyset$ and $j \notin G^*(j)$, $\forall j \in N$, (b) $|P_i| \geqslant 2$, $\forall i \in N$, and (c) $NP_i \neq \emptyset$, $\forall i \in N$. The graders know the proportion of probe and non-probe papers assigned to them, but cannot tell them apart.

Grade assignment and performance scores

The mechanism compares the peer-graded scores ($\tilde{y}_j^{(i)}$) with true scores ($y_j$) on the probe papers $P_i$, to statistically estimate the error parameters $\hat{\theta}_i = (\hat{b}_i, \hat{\tau}_i) \in \mathbb{R} \times \mathbb{R}_{>0}$ of each grader $i$. In the following, we use $\tilde{y}_j^{(i)} = y_j + b_i + \sigma_i e_{ij}$ where $e_{ij} \sim \mathcal{F}(0,1)$. First,

$$\hat{b}_i = \frac{\sum_{j \in P_i} (\tilde{y}_j^{(i)} - y_j)}{|P_i|} = \frac{\sum_{j \in P_i} (y_j + b_i + \sigma_i e_{ij} - y_j)}{|P_i|} = b_i + \frac{\sum_{j \in P_i} (\sigma_i e_{ij})}{|P_i|}. \tag{1}$$

Similarly,

$$\begin{aligned} \hat{\tau}_i &= \frac{|P_i| - 1}{\sum_{j \in P_i} (\tilde{y}_j^{(i)} - (y_j + \hat{b}_i))^2} \\ &= \frac{|P_i| - 1}{\sum_{j \in P_i} (y_j + b_i + \sigma_i e_{ij} - (y_j + b_i + \frac{\sum_{j \in P_i} (\sigma_i e_{ij})}{|P_i|}))^2} \\ &= \frac{|P_i| - 1}{\sum_{j \in P_i} (\sigma_i e_{ij} - (\frac{\sum_{j \in P_i} (\sigma_i e_{ij})}{|P_i|}))^2} = \frac{|P_i| - 1}{\sigma_i^2 \left( \sum_{j \in P_i} (e_{ij} - (\frac{\sum_{j \in P_i} (e_{ij})}{|P_i|}))^2 \right)}. \end{aligned} \tag{2}$$

Therefore, $\sqrt{\hat{\tau}_i} \propto \frac{1}{\sigma_i}$, where the proportionality constant is a function of the realized values of the random variables $y_j$s and $\tilde{y}_j^{(i)}$s. The estimated parameters are used in assigning performance-scores to papers and performance scores to peer-graders.

Definition **1 (Score and Accuracy)** *We define the score-assignment rule and the accuracy as follows.*

▷ *The score-assignment function* $\mathbf{r} = (r_j : j \in N)$ *is* inverse standard-deviation weighted de-biased mean (ISWDM) *if for every non-probe paper* $j$*, it assigns*

$$r_j^*(\tilde{\mathbf{y}}_j^{G(j)}, \hat{\boldsymbol{\theta}}_{G(j)}) = \frac{\sqrt{\gamma}\mu + \sum_{i \in G(j)} \sqrt{\hat{\tau}_i} (\tilde{y}_j^{(i)} - \hat{b}_i)}{\sqrt{\gamma} + \sum_{i \in G(j)} \sqrt{\hat{\tau}_i}}, \tag{3}$$

*where* $\tilde{y}_j^{(i)}$ *is the evaluation by the* $i$*th peer-grader and* $(\hat{b}_i, \hat{\tau}_i)$ *are her estimated parameters. Score* $\mathbf{r}^*$ *assigns the instructor-verified grade on every probe paper.*

▷ *The* accuracy *of paper $j$, at a score $r_j^*$ and true score $y_j$, is*

$$W_j^*(\tilde{\mathbf{y}}_j^{G(j)}, \hat{\boldsymbol{\theta}}_{G(j)}, y_j) = R(r_j^*(\tilde{\mathbf{y}}_j^{G(j)}, \hat{\boldsymbol{\theta}}_{G(j)}), y_j), \tag{4}$$

*where $\tilde{\mathbf{y}}_j^{G(j)}$ is the vector of peer-evaluated scores reported on paper $j$, $\hat{\boldsymbol{\theta}}_{G(j)}$ is the vector of evaluated error-parameters for the relevant graders $G(j)$, and $R : \mathbb{R}^2 \to \mathbb{R}$ is a continuous reward function that measures the closeness of the true score $y_j$ and the given score $r_j^*$. Formally, $R(x_1, y_1) < R(x_2, y_2)$ if $|x_1 - y_1| > |x_2 - y_2|$ for all $x_1, x_2, y_1, y_2 \in \mathbb{R}$. We assume that $R(x, x) = 0 \geqslant R(x, y) = R(y, x)$ for all $x, y \in \mathbb{R}$. One example of such a function would be $R(x, y) = -(x - y)^2$, which calculates the squared error in assigned scores.*

▷ *The* accuracy *of a score $r_j^*$ for paper $j$* without grader *$i$ when the true score is $y_j$ is denoted by $W_j^{(-i)*} = W_j^*(\tilde{\mathbf{y}}_j^{G(j)\setminus\{i\}}, \hat{\boldsymbol{\theta}}_{G(j)\setminus\{i\}}, y_j)$ where $W_j^*(\cdot)$ is defined as above.*

*The parameters $\gamma$ and $\mu$ are the parameters of the prior as defined by the $\mathbf{PG}_1$ model of Piech et al. (2013) (see Section 2.2), and $\hat{b}_i$ and $\hat{\tau}_i$ are the estimated bias and reliability of grader $i$.*

We will use the shorthands $W_j^*$ and $W_j^{(-i)*}$ for the accuracies with and without agent $i$ respectively when the arguments of such functions are clear from the context.

The ISWDM score-assignment function takes a *weighted average* of the prior mean $\mu$ and the de-biased (subtracting the estimated bias from the reported scores) reported scores. De-biasing ensures that the biases of the graders do not affect the finally assigned grade. The weight is chosen to be the square-root of reliability, which is the inverse of the variance for that grader. Higher the estimated reliability, higher is the weight on a grader.

Without incentive concerns, a statistician would have suggested a score-assignment function that would minimize the expected squared distance between the assigned score and true score on exam $j$, conditional on the true bias and variance parameters. Then, those true parameters could be approximated by the estimated bias and variance. In Equation (30) of Appendix C, we show that under the strong distribution-assumptions of Piech et al. (2013), such a score-assigment function on exam $j$ would come from the class of *weighted average (WA) score-assignment functions*:

$$r_j^{\mathrm{WA}}(\tilde{\mathbf{y}}_j^{G(j)}, \hat{\boldsymbol{\theta}}_{G(j)}) = \frac{\lambda_0 \mu + \sum_{i \in G(j)} \lambda_i (\tilde{y}_j^{(i)} - \hat{b}_i)}{\lambda_0 + \sum_{i \in G(j)} \lambda_i}, \tag{5}$$

where $\lambda_0, \lambda_i \geqslant 0, \forall i \in N$, not all zero. In particular, the parameters turn out to be $\lambda_0 = \gamma$ and $\lambda_i = \hat{\tau}_i, \forall i \in N$ (note the difference with $\lambda_i = \sqrt{\hat{\tau}_i}$ in Equation (3); see Appendix C for details). Here, $\mu$ is the prior mean of all papers, and the term $(\tilde{y}_j^{(i)} - \hat{b}_i)$ is the de-biased score on paper $j$ from grader $i$. This is indeed the *expected (social) reward maximizer (ERM)*, with the reward function $R(x, y) := -(x - y)^2$ as given below:

$$r_j^{\mathrm{ERM}}(\tilde{\mathbf{y}}_j^{G(j)}, \hat{\boldsymbol{\theta}}_{G(j)}) \in \operatorname*{argmax}_{x_j \in \mathbb{R}} \mathbb{E}_{y_j \mid \tilde{\mathbf{y}}_j^{G(j)}, \hat{\boldsymbol{\theta}}_{G(j)}} R(x_j, y_j). \tag{6}$$

However, in Theorem 2 we will show that in the class of WA score-assignment functions, ISWDM uniquely satisfies certain desirable properties. But, despite not being exactly the ERM, ISWDM does not compromise the expected accuracy ($W_j$) much (see Appendix D).

Regrading Requests

We consider peer-grading mechanisms that allow regrading requests. We assume that when a regrading request is raised, the instructor regrades the paper herself and assigns the true score on the paper. We also assume that the students know the true scores on their own papers and only raise a regrading request when they expect it to raise their score further.

Assumption **1** *Student $j$ knows $y_j$ and raises a regrading request only if $r_j^* < y_j$.*

In the next section, we lay down the peer-graders' incentive structure and the desirable properties of a mechanism.

## 3. Incentives and Design Desiderata

**Individual Preferences.** We assume that every individual $i$ cares about (a) her total score (sum of her exam score $r_i$ and peer-grading performance score $t_i$), and potentially, also about (b) the total scores of the other individuals. To model a potentially competitive grader who also cares about (b), we assume that her utility is increasing in (a), weakly decreasing in (b).

For agent $i$ in mechanism $M = \langle G, \mathbf{r}, \mathbf{t} \rangle$, the utility is given by

$$u_i^M = r_i + t_i - \left( \sum_{j \in N \setminus \{i\}} w_{ij} \cdot (r_j + t_j) \right), \tag{7}$$

where $w_{ij} \geqslant 0$. We nest the standard case of non-competitive graders under $w_{ij} = 0$. This linear formulation of competitive preferences can be interpreted as an approximation of more complicated formulations, and it allows theoretical tractability.[9]

In this section, we will assume that a more reliable grading does not come at any extra cost for the peer-grader, and hence we exclude such a cost component from the utility expression. The objective here is to understand whether a peer-grading mechanism can reward more reliable grading monotonically, despite the presence of competitive preferences, and when increasing costs are not at play: we define the desirable properties accordingly. One could have considered *costs of grading* to be increasing in reliability. We do this in Section 6, and the desiderata change accordingly.

Note that a few uncertainties are resolved after grader $i$ chooses her decision variables $(b_i, \tau_i)$ and before $r^*$ and $t^*$ are computed by the mechanism: (a) the true score $y_j$ on paper $j$ realizes, (b) the decision variables $(b_k, \tau_k)$ are chosen by co-grader $k$ (i.e., the strategic uncertainty), (c) the scores are reported by grader $i$, $\tilde{y}_j^{(i)}$ for paper $j$, which is realized from $(\tilde{y}_j^{(i)} | y_j) \sim \mathcal{F}(y_j + b_i, 1/\tau_i)$ and (d) the score on paper $j$ is reported by a co-grader $k$, which is realized from $(\tilde{y}_j^{(k)} | y_j) \sim \mathcal{F}(y_j + b_k, 1/\tau_k)$. We define two desirable properties of peer-grading mechanisms. The properties consider the grader $i$'s expected utility from the choice of strategies she makes. All expectations are taken *only with respect to* uncertainties

[9] This also takes care of the case where $i$ feels competitive against only a subset of her classmates who have had higher/comparable scores with her in the past experience. Student $i$ would assign $w_{ij} > 0$ to only those individuals to her utility function.

(a) and (c), i.e., the distribution of $i$'s grade-evaluation process $(\tilde{y}_j^{(i)}|y_j) \sim \mathcal{F}(y_j + b_i, 1/\tau_i)$ and the distribution of $y_j$. The properties hold for any ex-post realization of the other uncertainties (b) and (d), and there is no expectation taken on them. This is why both properties are defined as *ex-post*.

Definition **2 (Ex-Post Bias Insensitivity (EPBI))** *A peer-grading mechanism $M = \langle G, \mathbf{r}, \mathbf{t} \rangle$ is* ex-post bias insensitive (EPBI) *for grader $i$, if the expected utility of grader $i$ is independent of her bias $b_i$, irrespective of the biases and reliabilities of other graders $j \in N \setminus \{i\}$,* and *reported scores of the other graders. Define the following shorthand for the expectation, $\mathbb{E}_{i,b_i,\tau_i} \equiv \mathbb{E}_{y_j,\ j \in G^{-1}(i)} \mathbb{E}_{\tilde{y}_j^{(i)}|y_j \sim \mathcal{F}(y_j+b_i,1/\tau_i),\ j \in G^{-1}(i)}$.*[10] *Then we can mathematically define EPBI as*

$$\mathbb{E}_{i,b_i,\tau_i}\ u_i^M(\tilde{y}_j^{(i)}, \tilde{\mathbf{y}}_j^{(-i)}, y_j) = \mathbb{E}_{i,b_i',\tau_i}\ u_i^M(\tilde{y}_j^{(i)}, \tilde{\mathbf{y}}_j^{(-i)}, y_j), \\ \forall \{\tilde{y}_j^{(k)}, b_k, \tau_k\}_{k \neq i, j \in G^{-1}(i)}, \forall \tau_i, \forall b_i' \neq b_i. \tag{8}$$

*A peer-grading mechanism $M$ is EPBI, if it is EPBI for all participants $i \in N$.*

Definition **3 (Ex-Post Reliability Monotonicity (EPRM))** *A peer-grading mechanism $M = \langle G, \mathbf{r}, \mathbf{t} \rangle$ is* ex-post reliability monotone (EPRM) *for grader $i$, if her utility is* monotonically increasing *with her reliability, irrespective of the biases and reliabilities chosen by the other graders $j \in N \setminus \{i\}$,* and *the scores reported by the different graders. Mathematically (using the same shorthand as in Definition 2),*

$$\mathbb{E}_{i,b_i,\tau_i}\ u_i^M(\tilde{y}_j^{(i)}, \tilde{\mathbf{y}}_j^{(-i)}, y_j) > \mathbb{E}_{i,b_i,\tau_i'}\ u_i^M(\tilde{y}_j^{(i)}, \tilde{\mathbf{y}}_j^{(-i)}, y_j), \\ \forall \tau_i > \tau_i', \forall \{\tilde{y}_j^{(k)}, b_k, \tau_k\}_{k \neq i, j \in G^{-1}(i)}, \forall b_i. \tag{9}$$

*A peer-grading mechanism $M$ is EPRM, if it is EPRM for all participants $i \in N$.*

Both these properties are in some ways stronger than a *dominant strategy* version of the above definitions, as they hold for all realizations of uncertainties (b) and (d), as described on the last page.

We are now in a position to present the central mechanism of this paper.

## 4. The PEQA mechanism

Algorithm 1 shows the detailed steps of PEQA. For a simpler exposition of the algorithm, we provide a little non-rigorous description of PEQA in Algorithm 2 in the appendix.

[10] Note that the shorthand $\mathbb{E}_{i,b_i,\tau_i}$ explicitly means expectation with respect to $y_j$ and $\tilde{y}_j^{(i)}|y_j$ distributions.

**Algorithm 1** `PEQA` mechanism

**Require:** Parameters $\mu, \gamma$ of $\mathcal{F}$, $\alpha$ of the performance score, probe set $P$ with $|P| = \ell \in \left[\frac{K}{2}+1, \frac{n}{\frac{K}{2}+1}\right]$ ($K \geqslant 2$, even, is the number of papers assigned to each grader)

$(G(j), j \in N) \leftarrow$ `computeG`$(N)$, where $G(j)$: graders of $j$

**Require:** Reported scores $\{\tilde{y}_j^{(i)}, i \in G(j), j \in N\}$ given by the assigned graders $G$

Calculate $\hat{b}_i = \frac{\sum_{j \in P_i}(\tilde{y}_j^{(i)} - y_j)}{|P_i|}$, $\hat{\tau}_i = \frac{|P_i|-1}{\sum_{j \in P_i}(\tilde{y}_j^{(i)} - (y_j + \hat{b}_i))^2}$

Tentative score of the paper $j \leftarrow r_j^*(\tilde{\mathbf{y}}_j^{G(j)}, \hat{\boldsymbol{\theta}}_{G(j)})$ (via ISWDM, Equation (3))

Publish grades, invite regrading requests

After regrading period ends

**for** each paper $j \in N$ **do**

  **if** paper $j$ has regrading request **then**

    $y_j$ = true grade as checked by an instructor

  **else**

    $y_j = r_j^*$

$(t_i, i \in N) \leftarrow$ `computet`$(\tilde{\mathbf{y}}, \hat{\boldsymbol{\theta}})$, where $t_i$: performance score of $i$

**function** `computeG`$(N)$:

**for** each grader $i \in N$ **do**

  $G^{-1}(i) = \emptyset$

  **for** paper $k \in \{i+1, \ldots, i+1+K/2\}$ **do**

    $G^{-1}(i) \leftarrow G^{-1}(i) \cup (k \bmod \ell)$

    $G^{-1}(i) \leftarrow G^{-1}(i) \cup (\ell + k \bmod (n-\ell))$

**return** $G$

**end function**

**function** `computet`$(\tilde{\mathbf{y}}, \hat{\boldsymbol{\theta}})$:

$\mathbf{t} := (t_i, i \in N) \leftarrow \mathbf{0}$

**for** each paper $j \in N \setminus P$ **do**

  Calculate $W_j^*$ (Equation (4))

  **for** each grader $i \in G(j)$ **do**

    Calculate $W_j^{(-i)*}$ given by $W_j^*(\tilde{\mathbf{y}}_j^{G(j)\setminus\{i\}}, \hat{\boldsymbol{\theta}}_{G(j)\setminus\{i\}}, y_j)$ using Equation (4)

    $t_i \leftarrow t_i + \alpha(W_j^* - W_j^{(-i)*})$

**return** $\mathbf{t}$

**end function**

In short, the algorithm description specifies the three functions of a peer-grading mechanism $\langle G, \mathbf{r}, \mathbf{t}\rangle$ as defined in Section 2.1. The papers are assigned to the graders in a specific way. The assigned score on a paper is a weighted average (with appropriately chosen weights). Finally, the grading performance score is the *marginal contribution* of the grader towards the accuracy. The following lemma shows that the `computeG` function almost evenly distributes the non-probe papers.

Lemma **1** *In* `computeG`*, no agent gets her own paper for grading. Also, every non-probe paper is assigned to at least* $\frac{K}{2}$ *and at most* $\frac{K}{2}+1$ *graders.*

In the next section, we present our results on `PEQA`.

## 5. Properties of `PEQA`

Our first result shows that `PEQA` satisfies both the properties mentioned in Section 3, as long as the students care more about their own scores than others' scores.

Theorem **1** *If* $\sum_{k \in N\setminus\{i\}} w_{ik} \leqslant 1, \forall i \in N$*, then* `PEQA` *is EPBI and EPRM for all* $\alpha > 0$.

The expression $\sum_{k \in N\setminus\{i\}} w_{ik}$ denotes the sum of the relative weights assigned by $i$ to other peers' total scores. We believe that $\sum_{k \in N\setminus\{i\}} w_{ik} \leqslant 1$ is an intuitive restriction on

competitive preferences: Even competitive graders care more about their own score than they care about other's scores.[11]

A direct consequence of this result is that a grader will have no incentive in putting a deliberate upward or downward bias in this competitive environment and also will find it in her interest to maximize her reliability.

All the $r_k$ terms (for $k = i$ and $k \neq i$) in the utility expression (Equation (7)) would be replaced by $\max\{r_k^*, y_k\}$ where $r_k^*$ is the mechanism assigned score. This is due to how regrading requests are raised (Assumption 1) and because the instructor is assumed to give the correct score $y_k$ when a regrading request is received.

To make the proofs easily readable, we provide an intuition of the main ideas here. The complete details are available in Appendix B.

The EPBI result is driven by how the score-assignment function de-biases the grades through the estimated grader bias. Though the bias estimates from probes are noisy, in expectation, they are correct and are identical across probes and non-probes. Thus grader $i$'s bias cannot lower other's assigned final scores. We show that bias also does not effect the post-regrading expected score $\max\{r_j^*(\cdot), y_j\}$. Thus biasing reports does not provide any competitive incentives. Her grading performance score depends only on the assigned final scores on the papers she graded, and hence it is unaffected by bias too. EPBI is independent of the condition on $w_{ik}$s.

Intuitively, two forces drive the EPRM result.

- ▷ The link between $i$'s grading performance score and her marginal contribution to accurate grading plays a crucial role. A lower grading reliability of $i \in G(j)$ invariably lowers $i$'s marginal contribution to accurate score-assignment on paper $j$. This lowers $i$'s grading performance score and hence, her total utility.
- ▷ The score-assignment function and our regrading assumption (see Assumption 1) are crucial too. As mentioned previously, under our score-assignment function, grader $i$'s noisier grading leads to a noisier assigned grade on paper $j$. The noise moves the assigned grade above or below the true grade. Higher is the noise, larger is the potential movement in either direction. Grader $i$ determines the magnitude of the noise, but not the direction in which the noise moves the assigned grade. By selectively asking for regrades, student $j$ keeps any undeserved high grades and reverses any low grades that result from the noise. Thus, $i$'s noisier grading ends up increasing $j$'s grades post regrading-requests. Given $i$ dislikes when $j$ gets higher grades, this decreases $i$'s utility in expectation. Thus, $i$'s competitiveness also fuels $i$'s desire for an accurate grading.

Deriving the EPRM condition requires a bound on $w_{ik}$s. This is because the choice of reliability of grader $i$ affects the final grades of whoever she grades, and the marginal contributions (thus, grading performance scores) of her co-graders. We show that the condition on $w_{ik}$s is sufficient to ensure that the collective weight on other's grading performance scores never outweighs a competitive student's regard for her own performance score, irrespective of other's actions and noise. In the proof, we also show that the sufficient condition on the $w_{ik}$'s can be further weakened to a sum over only her co-graders. We kept the condition as mentioned in the theorem statement for simplicity and explainability.

[11] This expression is zero for the non-competitive grader who only cares about her own grades.

The relative weight $\alpha$ that an instructor assigns in `PEQA` (see Step 7 of the `computet` function in Algorithm 1) on the peer-grading performance score, determines what percentage of total grades come from own exam-score versus completing the peer-grading exercise, and can vary across different instructors and courses. It is, therefore, desirable to have a score-assignment function that is robust to any choice of the relative weight $\alpha$ while retaining the two properties above. It turns out that to meet this desirable criterion, the score-assignment function $r^*$ given by Equation (3) is crucial since any other score-assignment function in the weighted-average class would fail to keep the mechanism EPRM for some such percentage $\alpha$. What if one allowed some other performance score function $\mathbf{t}$, different from the one under `PEQA`? It turns out, "any other score-assignment function fails for some $\alpha$" stays true unless we satisfy the necessary condition of the following result, even if one starts with a different performance score function. This is the main idea of our next uniqueness result.

As a first step, we start with an arbitrary performance score function $\mathbf{t}$, which may be different from the one in `PEQA` and is chosen by the instructor. We will show that if that peer-grading mechanism needs to stay EPRM for all choices of $\mathbf{t}$, then ISWDM is necessary. This shows why even the ERM score-assigment function (Equation (6)) is not the optimal choice from the WA class in a world where reliability needs to be incentivized.

Theorem **2 (Uniqueness)** *Assume for every $i \in N$, $\exists j \in G^{-1}(i)$ s.t. $w_{ij} > 0$. Fix an arbitrary grader $i$ and any performance score function $\mathbf{t}$. The peer-grading mechanism $M = \langle G, \mathbf{r}^*, \mathbf{t} \rangle$, where $r_j^* \equiv r_j^{WA}, \forall j \in N$ (Equation (5)) is EPRM for grader $i$ for every peer-grading performance score function $\mathbf{t}$ only if $\lambda_i = \kappa_i/\sigma_i$, where $\kappa_i > 0$ is a factor independent of $\sigma_i$.*

As shown in Equation (2) and the discussion following it, we have $\sqrt{\hat{\tau}_i} \propto 1/\sigma_i$. Therefore, in the class of weighted average score computing function, the ISWDM score-assignment function, used by `PEQA` uniquely (upto constant multipliers) ensures EPRM for flexible performance score weight $\delta > 0$. This result shows why our score-assignment function is special, irrespective of the choice of grading performance scores.

At the risk of oversimplification, here is an intuition about how this result works. For the class of weighted average (WA) score-assignment functions, let us consider how the weights affect the *post-regrading* score.

$$\max\left\{r_j^{\text{WA}}(\tilde{\mathbf{y}}_j^{G(j)}, \hat{\boldsymbol{\theta}}_{G(j)}), y_j\right\} = y_j + \max\left\{\frac{\lambda_0(\mu - y_j) + \sum_{i\in G(j)} \lambda_i(\tilde{y}_j^{(i)} - \hat{b}_i - y_j)}{\lambda_0 + \sum_{i\in G(j)} \lambda_i}, 0\right\} \tag{10}$$

The first term on the RHS is the true score $y_j$ which is independent of grader $i$'s actions. We will focus on how grader $i$'s choices affect the numerator and the denominator of the second term, which is non-negative. The $(\tilde{y}_j^{(i)} - \hat{b}_i - y_j)$ terms are approximately a measure of the noise present in the signals that grader $i$ observed for paper $j$, which has a variance of $\sigma_i^2$. But, $\lambda_i = \kappa_i/\sigma_i$ uniquely makes the product $\lambda_i(\tilde{y}_j^{(i)} - \hat{b}_i - y_j)$ independent of grader $i$'s chosen $\sigma_i$, for all values of $\sigma_i$. This is true for all her co-graders too. Hence the numerator is independent of the variance of the graders, which is the first step of the proof.

The denominator is the sum of positive numbers. The term $\lambda_i = \kappa_i/\sigma_i$ guarantees that when $\sigma_i$ increases the whole fraction increases. Thus, noisier grading ends up increasing the

post-regrading score $\max\left\{r_j^{\text{WA}}(\tilde{\mathbf{y}}_j^{G(j)}, \hat{\boldsymbol{\theta}}_{G(j)}), y_j\right\}$. In the proof, we formalize this intuition, while accounting for what information is available to grader $i$ when she contemplates how her actions affect post-regrading scores.

Finally, we investigate the complexity of PEQA, which happens to be linear in the number of the agents and the papers each grader grades.

Theorem **3** *The worst-case complexity of computing* PEQA *is* $O(nK)$.

## 6. Welfare Under Costly Reliability

In this section, we extend our analysis to how PEQA deals with social welfare in a world where increasing reliability is costly to the grader. We calculate student welfare by subtracting the total reliability-cost from the sum of the grading-accuracy of all exams. We show that a modification of the grading performance score $\mathbf{t}^*$ of PEQA implements the student welfare-optimal level of costly reliability.

**Costly reliability.** As before, we assume that each paper has a single question.[12] All graders face the same reliability-cost function $c$ while grading that paper/question.

The estimated reliability for grader $i$, $\hat{\tau}_i$, is computed from her performance on the probe papers. Reliability is bounded above, i.e., $\tau_i \in [0, \bar{\tau}], \forall i \in N$. We summarize our assumptions below.

1. The cost $c : [0, \bar{\tau}] \to \mathbb{R}_{\geqslant 0}$ is convex, increasing, and equal for all graders $i \in N$.
2. The course instructor does not know $c$, only the graders do.

We simplify grader utility by assuming a uniform weight for the other-regarding component ($w_{ij} = w, \forall\ i, j \in N$) and is a common knowledge.

**Student welfare of grading.** For a non-probe paper, we assume that the social planner (e.g., the instructor) cares about two dimensions of welfare: (a) the accuracy of the final score (measured by the reward function $R(r_j^*, y_j)$) and (b) the total cost of grader-reliability.

We presume that *if the social planner was aware of the cost functions of grading*, she would have recommended a joint strategy profile $(\tau_i, \tau_{-i})$ that maximizes some linear combination of the reward and cost factors, which we call the *student welfare*. Define the set of all non-probe papers to be $NP := \cup_{i \in N} NP_i$. Then, the student welfare of grading all the papers is formally written as

$$\beta \sum_{j \in NP} E_{y_j} E_{(\tilde{y}_j^{(k)} | y_j) \sim \mathcal{F}(y_j + b_k, 1/\tau_k), k \in G(j)} R(r_j^*(\tilde{\mathbf{y}}_j^{G(j)}, \hat{\boldsymbol{\theta}}_{G(j)}), y_j) - K \sum_{i \in N} c(\tau_i), \qquad (11)$$

where $\beta > 0$ determines the relative weight between the two factors. The first term in the above expression excludes the probe papers which are accurately graded (by assumption) and are independent on $\tau_i$'s. However, the second term considers all papers since the graders incur costs for both probes and non-probes.

**Aligning social and individual incentives.** When the costs are private information, PEQA ensures that students exert welfare-optimal reliability in an equilibrium. There are three challenges on the way to aligning social and individual incentives. We discuss these below along with their solutions.

[12] In Section 8, we sketch how the analysis can be easily extended to multiple questions per exam.

- ▷ An instructor would care about the accuracy in grade-allocation, but how to make peer-graders care about the same? PEQA's grading performance score forces students to internalize accuracy in their decisions by paying each grader their marginal contribution to accuracy.
- ▷ Each competitive grader wants lower scores for others as part of her other-regarding utility. This is not aligned with the student welfare of grading and becomes a source of externality. We solve this by suggesting a modified grading performance score below, that additionally compensates graders for any potential losses from their other-regarding utility.
- ▷ The solution to the point above presents a new challenge. The other-regarding utility component would be different for each grader $i$, as their reference groups $N \setminus \{i\}$ would naturally be different. Thus they will be compensated different amounts. *Would this change the ordinal ranking of students in the class from that of* PEQA*?* We show that the answer is *no* (Lemma 2).

**Modified grading performance score.** Let the post-regrading request score be $g_i = \max\{r_i, y_i\}$. We propose the modified grading performance score

$$\pi_i := t_i + w \sum_{j \in N \setminus \{i\}} (g_j + \pi_j), \tag{12}$$

where $t_i$ is the original PEQA grading performance score. The additional terms on the RHS compensate for the other-regarding component in grader $i$'s utility. Though this simplifies the net utility of grader $i$, the simplicity comes at a price: if $i$ and $j$ are co-graders then $\pi_i$ has been described as a function of $\pi_j$ and vice-versa! How is the designer supposed to decide the values of $\pi_i$ and $\pi_j$ given the interdependency? We show that $\pi_i$ has an alternative expression that is independent of $\pi_j$s.

$$\pi_i = \frac{t_i + w \sum_{j \in N \setminus \{i\}} g_j + w\pi}{1 + w}, \tag{13}$$

$$\text{where, } \pi = \frac{t + w(n-1)g}{1 - w(n-1)}, \text{ and } t = \sum_{i \in N} t_i,\ g = \sum_{i \in N} g_i,\ w \neq \frac{1}{n-1}. \tag{14}$$

**The game of peer-grading.** The modified PEQA mechanism induces a game among the peer-graders after all the answerscripts of the exam have been submitted. The players (the graders) choose their reliabilities as their strategies to maximize their utility. Grader $i$'s utility is given by

$$u_i = g_i + \pi_i - w \sum_{j \in N \setminus \{i\}} (g_j + \pi_j) - \sum_{j \in G^{-1}(i)} c(\tau_i) = g_i + t_i - Kc(\tau_i), \tag{15}$$

where $K$ is the total number of papers graded by $i$ (including probes). *Thus, $\pi_i$ nullifies the other-regarding component of utility.*

The score assignment and performance score functions, that map players' strategies to players' utilities, are also common knowledge. Players simultaneously choose their reliabilities $\tau_i,\ i \in N$ to maximize their expected utility.

The following result shows that $\pi := (\pi_i, i \in N)$ retains the same order of course-scores as the original score functions $(t_i, i \in N)$.

Lemma **2 (Order Invariance)** *Fix a profile of player strategies and true scores in the peer-grading game. The modified* PEQA *performance score* $\pi$ *retains the same order among the students as the original* PEQA *performance score* $t$*.*

*Proof*: Consider two students $i$ and $k$ where $i$ received more total score in modified PEQA than $k$. We show that it is equivalent to $i$ getting more total score than $k$ in original PEQA. We show this in the following equivalent implications: $g_i + \pi_i > g_k + \pi_k \Leftrightarrow g_i + \frac{t_i + w\sum_{j\in N\setminus\{i\}} g_j + w\pi}{1+w} > g_k + \frac{t_k + w\sum_{j\in N\setminus\{k\}} g_k + w\pi}{1+w} \Leftrightarrow \frac{g_i + t_i + w(g+\pi)}{1+w} > \frac{g_k + t_k + w(g+\pi)}{1+w} \Leftrightarrow g_i + t_i > g_k + t_k$. ■

The next result shows that $\pi := (\pi_i, i \in N)$ implements the student welfare-optimal level of reliability in a pure Nash equilibrium. The proportion of non-probe papers (which is fixed once the mechanism is announced) is denoted by $p_{\text{NP}}$.

Theorem **4** *If the instructor uses the modified grading score* $\pi_i$ *and sets* $\alpha = \frac{\beta}{K}$*, every maxima of the expected student welfare is a Pure Strategy Nash Equilibrium (PSNE) of the induced game among the peer-graders.*

*Proof*: **Step 1:** $i$'s $\tau_i$-dependent utility component for grading paper $j$, is related to the accuracy of paper $j$ minus the cost of grading it.

As shown in Equation (15), $\pi_i$ already compensates for the other-regarding component and makes it inconsequential. The residual performance score $t_i$ in Equation (15) is the sum of performance scores $t_i^j$ from each paper $j \in G^{-1}(i)$. Now, $t_i^j = \alpha(W_j^* - W_j^{(-i)*})$, and $W_j^{(-i)*}$ do not depend on $i$'s reliability $\tau_i$. Hence the part of the $i$'s utility expression that depends on $\tau_i$ and is related to grading paper $j$ is:

$$\alpha W_j^* - c(\tau_i) = \alpha R(r_j^*(\tilde{\mathbf{y}}_j^{G(j)}, \hat{\boldsymbol{\theta}}_{G(j)}), y_j) - c(\tau_i), \tag{16}$$

where $\tilde{\mathbf{y}}_j^{G(j)}$ is the profile of all scores given by $G(j)$. Thus it depends potentially on the bias and reliability of co-graders, which are chosen strategically and simultaneously.

**Step 2:** Under $\pi_i$ and $\alpha = \frac{\beta}{K}$, $i$'s reliability-dependent utility is, in expectation, completely aligned with the student welfare.

Only non-probes have $W_j^* \neq 0$. Grader $i$, who is uncertain whether paper $j$ is a probe versus a non-probe paper, assigns a probability $p_{\text{NP}}$ to any paper being a non-probe. For any choice of bias and reliability by all the graders, the expected accuracy on paper $j$ is $E_{y_j} E_{(\tilde{y}_j^{(k)}|y_j)\sim\mathcal{F}(y_j+b_k,1/\tau_k),k\in G(j)} R(r_j^*(\tilde{\mathbf{y}}_j^{G(j)}, \hat{\boldsymbol{\theta}}_{G(j)}), y_j)$.[13] From the analysis of Theorem 1, we know that this expression is independent of bias under PEQA. Therefore, for simplicity, we assume that every grader strategizes only on her reliability. To emphasize the strategic and simultaneous choice of reliability, we denote the expected accuracy above using the shorthand $\bar{R}(\tau_i, \tau_{-i})$. Hence, for any reliability profile chosen by the set of graders on paper $j$, the part of the expected utility of grader $i$ that depends on $\tau_i$ is

$$U_i^j(\tau_i, \tau_{-i}) = \alpha \cdot p_{\text{NP}} \cdot \bar{R}(\tau_i, \tau_{-i}) - c(\tau_i). \tag{17}$$

[13] The distributions of these random variables are common knowledge of the graders.

When $\alpha = \frac{\beta}{K}$, grader $i$ maximizes $U_i^j(\tau_i, \tau_{-i}) = \beta \frac{p_{\text{NP}}}{K} \bar{R}(\tau_i, \tau_{-i}) - c(\tau_i)$.

Similarly, after taking the expectation w.r.t. the true and observed scores of the papers and graders respectively, the student welfare (Equation (11)) can be rewritten as $\beta \sum_{j \in NP} \bar{R}(\tau_i, \tau_{-i}) - K \sum_{i \in N} c(\tau_i)$. Note that $\bar{R}(\tau_i, \tau_{-i})$ is independent of $j$ and hence the expression can be simplified to $\beta \, |NP| \, \bar{R}(\tau_i, \tau_{-i}) - K \sum_{i \in N} c(\tau_i)$.

**Step 3:** Let $\tau^* = (\tau_k^*, k \in G(j))$ maximize student welfare. Then, it must be that $\beta \frac{p_{\text{NP}}}{K} \bar{R}(\tau_i^*, \tau_{-i}^*) - c(\tau_i^*) \geqslant \beta \frac{p_{\text{NP}}}{K} \bar{R}(\tau_i, \tau_{-i}^*) - c(\tau_i)$ for any alternative $\tau_i$ of player $i \in N$, where $p_{\text{NP}} = \frac{|NP|}{n}$. This is because any alternative $\tau_i$ that increases her expected utility would also increase the student welfare at $\tau^*$, creating a contradiction. Thus, if all except $i$ choose $\tau_{-i}^*$, player $i$ cannot do any better than choosing $\tau_i^*$. Thus, $\tau^* = (\tau_k^*, k \in G(j))$ is a PSNE of this game. ■

Instructors who believe that equilibrium is too restrictive a solution concept, could fall back upon our EPRM result (Theorem 1). Consider a grader operating under bonus $t_i$ with utility:

$$u_i = \underbrace{g_i + t_i - w \sum_{j \in N \setminus \{i\}} (g_j + \pi_j)}_{=: v_i} - K c(\tau_i) = v_i - K c(\tau_i), \tag{18}$$

Assuming an interior optimal, the grader chooses $\tau_i$ such that $\frac{dv_i}{d\tau_i} = K \frac{dc(\tau_i)}{d\tau_i}$. EPRM guarantees that the LHS is always positive, irrespective of beliefs about other graders. Thus, the LHS can be increased by scaling up $t_i$ to $\alpha t_i$ for $\alpha > 1$[14], to adjust reliability upwards, whenever necessary.

In the next section, we present our experimental study that tests some of our hypotheses made in the earlier results and verifies its practical usability.

## 7. Experimental study

The theoretical desirability of `PEQA` is established on restrictive assumptions about the domain of true and given scores, player utilities, and strategies. These assumptions approximate reality instead of describing it. How well does `PEQA` perform in a real-life exercise, where the scores and signals come from a bounded interval, or when player's utilities are competitive but not necessarily linear? This motivated our small `PEQA` experimental study.

Data from the `PEQA` study also help us investigate if two of our modeling assumptions are violated in reality: if bias and reliability are indeed identical in probes and non-probes, and, if competitive ($w_{ij} \geqslant 0$) preferences are a good model of the peer-grader behavior.

In a separate experimental study, we also run the *median mechanism*, that assigns the median peer-grader report without any performance score. It is the most popular mechanism used currently in MOOCs. We investigate the trade-off between the theoretical desirability of `PEQA` against the simplicity of the *median mechanism*. [15]

[14] In Step 3 of the proof of Theorem 1, we show $\frac{dt_i}{d\tau_i} > 0$.

[15] All data and code are available via: https://www.cse.iitb.ac.in/~swaprava/papers/Codes_Peer_Grading.zip

### 7.1 Experimental Design

We ran two experimental sessions: one with the median scoring mechanism (27 students), another with the PEQA mechanism (42 students). We recruited students through two open-calls to undergraduate students enrolled in a computing course (Prog101). The open calls did not contain any particulars of the two sessions. Every student who signed up for participation was assigned to one of the two sessions. An example view of the peer-grader while grading a paper is provided in Appendix F.

The experimental environment is not an exact replication of model assumptions, rather a replication of how a real-life peer-grading scenario would look like. In many classes, instructors grade on a curve: final numerical scores are converted to letter-grades (A to D) based on relative rankings. Grading on a curve creates a competitive classroom-environment that we wanted to replicate. We told participants that their total score is the sum of their peer-evaluated score and grading performance score. We paid students by the relative ranking of their total scores in the class, in both the sessions. The students who ranked in the first quartile of the total scores received M 650,[16] the next three quartiles received M 450, M 250, and M 50 respectively. They also received a show-up fee of M 50, irrespective of their total score. The monetary payments were placeholders for grades A to D in a class that grades on a curve: high relative performance resulted in high rewards.[17]

In the median mechanism session, the grading performance scores of all students were set to zero. The total-score ranking was identical to their peer-evaluated-score ranking. Thus, a student could decrease others' scores on the peer-evaluation task to increase her relative ranking and payment.

The PEQA session used the PEQA assignment and grading performance scores. Thus, manipulation on the peer-evaluation task risked getting a lower performance and total score, which would result in a lower payment.

The instructions and incentive-scheme, included in Appendix E, were explained in detail before each of the sessions began. In both sessions, we used numerical examples in our explanation. For PEQA, we showed the relation between performance score and grading reliability through a graph and verbally summarized the monotonic relationship.

We conducted both sessions during the weekly Prog101 labs, that happen in a large computer lab. Given this was a programming class, all questions being graded had objective criteria for being correct or incorrect. Further, the same questions were graded under both mechanisms. Our study lies at the intersection of *Lab* and *Field* experiments. We are interested in peer-grading behavior and the students are our population of interest. In this study, we observed our population of interest in their *naturally occurring environments*, like in Field experiments.

In both sessions, we asked students to peer-grade the same weekly class-quiz. We partitioned each quiz into three sub-quizzes (by treating one(two) question(s) of the quiz as a sub-quiz[18]), and divided each session into *three* rounds. In every round, the students were asked to peer-grade five sub-quizzes (each corresponding to one of five of her anonymous

[16] M = Indian Rupee (₹), a difference of M 200 is significant for a student.

[17] The ethics committee did not allow us to use university grades in the Prog101 class as incentives. They were worried that the students might feel coerced to provide us consent about accessing their data for our research, if we made the peer-grading part of the course-grading.

[18] The quiz had more than three questions.

peers). At the end of each round, students saw: (a) how peers had evaluated her performance on the sub-quiz, (b) her assigned score (median-scoring or `PEQA`), and (c) how her co-graders that round had evaluated the sub-quiz.

Within every sub-quiz, some (and not all) of the questions were 'regradable'. The students could raise a regrading request for only those questions at the end of the session. In the `PEQA` sessions, only the regradable questions were incentivized by the grading performance score. The non-regradable questions used the same assignment function but did not have any grading performance score.

We also graded all the papers ourselves (the instructor graded all of them), and we considered these scores to be the *true scores*. The difference between mechanism assigned scores and true scores is a measure of the quality of these mechanisms.

### 7.2 Hypotheses and results

Bias is calculated by subtracting the peer-assigned score from the true score. It measures the average direction and magnitude of manipulation. `PEQA` assumes that bias is zero or positive: students generally do not manipulate scores upwards (i.e., do not collude). The alternative hypothesis, that we would like to reject, would be that graders manipulate grading scores favorably for each other, implying a negative bias:

**Hypothesis 1** *Score-manipulation is collusive.*

Our second hypothesis suggests that bias should be maximum in the last round for two reasons. First, most repeated interactions have an end-game effect: selfish behavior unravels when no future interactions remain. Second, students who have experienced score-manipulation by others might retaliate as a punishment or reciprocal strategy in the later rounds of the treatment.

**Hypothesis 2** *Bias is maximum in the last round.*

In Tables 1 and 2, we summarize the bias in individual grading behavior in the three rounds of both treatments. To compare across questions and rounds, we normalize bias by the total score of the corresponding question.

| | Students | Round | Total grade | Avg bias (% of Total grade) | |
|---|---|---|---|---|---|
| | | | | regradable | non-regradable |
| Median | 27 | 1 | 1+1 | -0.4% | -0.6% |
| | | | | (-4%,3.2%) | (-3.1%,1.9%) |
| | | 2 | 2+2 | 1.7% | 1.2% |
| | | | | (-1%,4.4%)) | (-0.8%,3.1%) |
| | | 3 | 2+2 | 16.6% | 16.4% |
| | | | | (12%,21.2%) | (12%,20.7%) |

Table 1: Average bias from 3 rounds grading under the Median mechanism. We report the 95% confidence intervals below the averages.

| | Students | Round | Total grade | Avg bias (% of Total grade) | |
|---|---|---|---|---|---|
| | | | | regradable | non-regradable |
| PEQA | 42 | 1 | 1+1 | -0.6% | -0.5% |
| | | | | (-2%,1%) | (-1.9%,0.9%) |
| | | 2 | 2+2 | 0.6% | -0.3% |
| | | | | (0%,1.2%) | (-1.2%,0.6%) |
| | | 3 | 2+2 | 15.8% | 15% |
| | | | | (12.3%,19.2%) | (11.2%,18%) |

Table 2: Average bias from 3 rounds grading under the PEQA mechanism. We report the 95% confidence intervals under the averages.

In each round, every student graded a regradable and a non-regradable question.[19] The average bias is statistically identical to zero for the first two rounds, and significantly positive in the third round. This is true for both the regradable and non-regradable questions. Thus, the bias is either zero, or positive, and we can reject Hypothesis 1. The average bias (and the average absolute value of bias) is also significantly higher in the third round, based on t-tests with p-values smaller than .001. This holds for both regradable and non-regradable questions.

The lack of any bias, as evidenced by the tight confidence interval around 0, in the first two rounds parallels the results on honest reporting from the "die-roll in person and report" studies (Fischbacher & Föllmi-Heusi, 2013; Mazar, Amir, & Ariely, 2008). In these studies, subjects roll a die privately, self-report the outcome, and get paid based on the report. Fischbacher and Föllmi-Heusi (2013) report that only 20% of people lie to the fullest extent, 39% choose to be honest, and a sizable proportion cheats only marginally. Lying aversion (Dufwenberg & Dufwenberg, 2018), caring about lie-credibility, and a notion of self-concept maintenance (Mazar et al., 2008) are potential reasons for why people do not lie completely even under full anonymity.

How do the two mechanisms perform? The median assignment rule, due to its robustness to outliers, is immune to insincere grading as long as only a minority of graders are insincere. PEQA is bias invariant (EPBI), incentivizes effort, and should outperform the Median mechanism. We use the accuracy of the mechanism-assigned scores as a metric of relative performance. Given students graded most insincerely in the third round, we use this round to test Hypothesis 3.

**Hypothesis 3** PEQA *assigns accurate final scores. In the presence of strategic manipulation, the final score assigned under* PEQA *is closer to the true scores, than that assigned under median-scoring.*

In Table 3, we present the means of fractional-difference and squared fractional-difference between the mechanism-assigned score and true score. Thus, for the former we calculate $d_j = (\text{true score}_j - \text{mechanism assigned score}_j)/\text{total score}_j$ on student $j$'s third-round subquiz, and then take the average over all $j$. Similarly, the latter is the average of $d_j^2$. We find the true score on an exam by grading it ourselves.

[19] We wanted to check if students manipulate grades more when questions are non-regradable. We do not find any statistically significant effect.

The average difference between true and mechanism assigned scores is 14.8% under Median and only -1.2% under PEQA. The negative sign indicates that PEQA assigned slightly higher scores than the true score. Both difference and squared difference are significantly smaller under PEQA. Under the median mechanism, the difference and squared difference were equal because $d_j$ almost always took values of 0 or 1.

| | Median | PEQA | Median-d | Median-p |
|---|---|---|---|---|
| Mean Difference | 14.8% | -1.2% | 0% | 0% |
| Mean Squared Difference | 14.8% | 0.6% | 0% | 0.5% |
| $N$ | 27 | 42 | 27 | 27 |

Table 3: Difference and squared-difference. The first two columns are from Median mechanism and PEQA. The last two columns are from an ablation study where we used the Median data, but additionally debiased the reports (Median-d), or additionally ran the PEQA assignment function on the reports along with debiasing (Median-p).

The median mechanism assigned lower than true grade (assigned a 1 instead of a 2) for 15% (4 out of 27) of the sub-quizzes. In comparison, the PEQA mechanism was (almost) always point-precise: only one sub-quiz (out of 42) assigned a grade of 0.5 points higher. Thus, the proportion of cases with incorrect grades is significantly smaller under the PEQA mechanism at p-value of $p = .03$.[20] The number of regrading requests in the median and PEQA sessions were 4/27 and 3/42 respectively, a difference that is not statistically significant.

We also ran an *ablation study* where we removed parts of the Median mechanism and replaced it with features of the PEQA to understand why the latter works better. We randomly chose two of the five sub-quizzes that Median subjects graded in each round, and treated them as *probes*, and the rest as *non-probes*.[21] We estimated each grader's bias and reliability from those probes. For our first ablation rule, we apply the Median assignment rule on the debiased scores to create a synthetic set of scores (Median-d). For our second ablation rule, we apply the ISWDM score assignment function (Equation 3) to the scores reported. This means, we debiased the scores, as well as weighted the debiased scores by the square root of the reliability of the respective graders, to create another synthetic set of scores (Median-p). The Median-d scores should outperform the Median scores if Median graders were introducing significant bias. Further, the Median-p scores should outperform the Median-d scores if graders varied on their reliability, and hence weighing graders differently was valuable.

We report the performance of both ablation rules in Table 3, in terms of the mean difference and mean squared difference between the true scores and the assigned scores. The ablation study reveals that both Median-d and Median-p scores perform as well as PEQA[22], which implies that introducing bias was the main reason that the original Median mechanism performed poorly as compared to PEQA. This is perhaps explained by the fact that the peer-grading sessions were run during a regular lab session, where students were under supervision and hence could not neglect grading duties to indulge in more entertaining

[20] We used a one-sided Equality of proportions hypothesis test

[21] The results are not sensitive to which sub-quizzes are chosen are probes.

[22] There was no statistical difference between any pair of them.

activities (for example, browsing social media). Thus, students were equally attentive across both incentive schemes, and their only form of manipulation was introducing a bias.

One of the crucial assumptions of PEQA was that bias and noise are invariant across probes and non-probes used in that mechanism. An alternative scenario would be one where graders somehow figured out which are the probes, then gamed the system by manipulating scores on the non-probes, while not manipulating scores on the probes. The relevant hypothesis, where PEQA assumption would fail, would be:

**Hypothesis 4** *Bias is higher in non-probes than in probes.*

To test this, we pooled across all three rounds, and all questions where PEQA incentives were used, to maximize power. In a t-test, we could reject this hypothesis with the p-value of 0.09.

## 8. Discussions and limitations

The assumption of one question per exam, used in Sections 2 and 6 can easily be generalized to multiple questions per exam.

For Section 2, the definitions and analysis, done for one question per paper, can be extended for multiple questions per paper assuming that agent $i$ has a bias $b_{iq}$ and reliability $\tau_{iq}$ for question $q$ (which can be different from that of a different question $q'$, but same for question $q$ on all the papers she grades $j \in G^{-1}(i)$). The definitions of EPBI and EPRM, and PEQA will be updated accordingly by *indexing the question $q$ of paper $j$* as $(j, q)$. For EPBI, the equalities will be for all questions $q$ of paper $j$ and for all papers $j$ that agent $i$ grades. For EPRM, the inequalities will be for two reliabilities $\tau_{iq} > \tau'_{iq}$ for each question $q$ for each paper $j$. PEQA will similarly get updated by calculating this pair $(j, q)$'s score-assignment function and performance score function.

The setup of Section 6 can be generalized to multiple questions per exam by calculating both the reward[23] from accurate grading and the student welfare of grading (Equation (11)) *question-wise*. To see this, assume that question $q$ on any exam has a reliability-cost function $c_q$ (which can be different for a different question $q'$).[24] If agent $i$ chooses a reliability $\tau_{iq}$ question $q$, she will face a corresponding cost $c_q(\tau_{iq})$. The total student welfare calculated for that question will be the reward minus the total cost (across all graders) of grading for that question. Next, one needs to extend the analysis presented in Section 6 to maximize the sum of student welfare for each question over all the exams.

Like all mechanisms, PEQA has its limitations. For example, it relies on the assumption that a grader grades all papers with the same bias and reliability. Among other things, this requires that the graders cannot distinguish probe versus non-probe papers. While the assumption has been utilized in other papers (Gao et al., 2016; Piech et al., 2013), it is definitely a strong assumption which might no longer hold if the the same set of probe papers are reused repeatedly. However, this issue can be handled using a set of *synthetic* papers used as probes. These are certain papers which are manufactured by the course staff with known marks (i.e., known mistakes for example) and deliberately mixed in the set of

[23] This is the distance between the true score and given score.

[24] For instance, in a physics exam, a question on the general theory of relativity is more difficult to grade than that on Newton's laws of motion.

probe papers. The advantages of synthetic papers are (a) the marks are already known, hence they do not need grading, (b) they can be created in large numbers with little cost, (c) these papers can help create the required ratio of the probes and non-probe papers so that the students cannot tell apart the probes from the non-probes.[25]

Another way for the assumption to fail would be if grader-assigned grades depend on the true quality of the papers. It is possible that graders only negatively grade good papers, but when a paper is clearly bad, they do not care. We suggest a partial solution which can perhaps be implemented and tested in future work: `PEQA` can be extended so that subjects are assigned probe papers of both high and low quality, so that for each individual we have two measures of bias, one when they grade high quality papers and another when they grade low quality papers. Then, those measurements can be used while debiasing the non-probe papers that have been assigned high and low scores respectively.

## 9. Conclusion

We introduce a new mechanism, `PEQA`, that uses a score-assignment rule and grading performance scores to incentivize graders. Our mechanism is robust to grader's competitive preferences. The rule and the performance score guarantee unbiased grades. They also guarantee that any grader's utility increases monotonically with her grading reliability, irrespective of her competitiveness and how her co-graders act. Our assignment rule is unique in its class to satisfy this utility-reliability monotonicity while allowing flexibility in how large performance scores need to be. When grading is costly, a special version of `PEQA` implements socially optimal reliability-choices in an equilibrium of the peer-evaluation game among co-graders. Finally, in our classroom experiments, `PEQA` assigns accurate final scores and outperforms the popular median mechanism.

## 10. Ethics statement

This paper provides a method for efficient *peer-grading* with the option of raising regrading requests if students are dissatisfied with their peer-graded scores. Peer-grading is a practice widely used in MOOCs. In our method, since the students get a chance to ask for regrades, it does not deny any student from getting a just and fair grade.

All exam papers in the experiments of this paper and also in the proposed method for future use are given to the peer-graders after removing every identifiable information. Hence, students grade answerscripts without knowing whose answerscript it is. We planned it that way to preserve the privacy of the individuals.

The related literature (Sadler & Good, 2006, e.g.) shows that peer-grading improves learning of the students in addition to their regular learning through a course. In our proposed method, the instructor holds the decision on how much weight of the total score of an exam/quiz should be given for peer-grading, e.g., about 10% or less. This is not unusual since such small course weights are often given to several related activities of a course, e.g., class participation or scribing lectures, etc.

[25] If a probe is sent to the same number of peer-graders as that of a non-probe, this will create the ideal situation where the probes will be impossible to distinguish from non-probes.

## Acknowledgments

This work has been supported by the Indian Institute of Technology Kanpur under the grant number 2017198. We would like to thank the Institutional Ethics Committee (IEC) of the Indian Institute of Technology Kanpur for providing us with the opportunity to run the human subject study with the students of the institute via the IEC Communication Number: IITK/IEC/2020-21/II/30. We also thank Bikramaditya Datta, Debasis Mishra and the seminar/ workshop participants at UC Davis, Academia Sinica, and WED 2019 (ISI Delhi) for their many valuable comments .

## References

Alon, N., Fischer, F., Procaccia, A., & Tennenholtz, M. (2011). Sum of us: Strategyproof selection from the selectors. In *Proceedings of the 13th conference on theoretical aspects of rationality and knowledge* (pp. 101–110).

Cai, Y., Daskalakis, C., & Papadimitriou, C. (2015). Optimum statistical estimation with strategic data sources. In *Conference on learning theory* (pp. 280–296).

Campanario, J. M. (1998). Peer review for journals as it stands today - Part 1. *Science communication*, *19*(3), 181–211.

Caragiannis, I., Krimpas, G. A., & Voudouris, A. A. (2015). Aggregating partial rankings with applications to peer grading in massive online open courses. In *Proceedings of the 2015 international conference on autonomous agents and multiagent systems* (pp. 675–683).

Caragiannis, I., Krimpas, G. A., & Voudouris, A. A. (2020). How effective can simple ordinal peer grading be? *ACM Transactions on Economics and Computation (TEAC)*, *8*(3), 1–37.

Cho, K., & Schunn, C. D. (2007). Scaffolded writing and rewriting in the discipline: A web-based reciprocal peer review system. *Computers & Education*, *48*(3), 409–426.

Dasgupta, A., & Ghosh, A. (2013). Crowdsourced judgement elicitation with endogenous proficiency. In *Proceedings of the 22nd international conference on world wide web* (pp. 319–330).

De Alfaro, L., & Shavlovsky, M. (2014). Crowdgrader: A tool for crowdsourcing the evaluation of homework assignments. In *Proceedings of the 45th acm technical symposium on computer science education* (pp. 415–420).

Dhull, K., Jecmen, S., Kothari, P., & Shah, N. B. (2022). The price of strategyproofing peer assessment. *arXiv preprint arXiv:2201.10631*.

Dufwenberg, M., & Dufwenberg, M. A. (2018). Lies in disguise–a theoretical analysis of cheating. *Journal of Economic Theory*, *175*, 248–264.

Faltings, B., Li, J. J., & Jurca, R. (2012). Eliciting truthful measurements from a community of sensors. In *Internet of things (iot), 2012 3rd international conference on the* (pp. 47–54).

Fiez, T., Shah, N., & Ratliff, L. (2020). A super* algorithm to optimize paper bidding in peer review. In *Conference on uncertainty in artificial intelligence* (pp. 580–589).

Fischbacher, U., & Föllmi-Heusi, F. (2013). Lies in disguise–an experimental study on cheating. *Journal of the European Economic Association*, *11*(3), 525–547.

Gao, A., Wright, J. R., & Leyton-Brown, K. (2016). Incentivizing evaluation via limited access to ground truth: Peer-prediction makes things worse. *arXiv preprint arXiv:1606.07042*.

Hamer, J., Ma, K. T., & Kwong, H. H. (2005). A method of automatic grade calibration in peer assessment. In *Proceedings of the 7th australasian conference on computing education-volume 42* (pp. 67–72).

Holzman, R., & Moulin, H. (2013). Impartial nominations for a prize. *Econometrica*, *81*(1), 173–196.

Jurca, R., & Faltings, B. (2005). Enforcing truthful strategies in incentive compatible reputation mechanisms. In *International workshop on internet and network economics* (pp. 268–277).

Jurca, R., & Faltings, B. (2009). Mechanisms for making crowds truthful. *Journal of Artificial Intelligence Research*, *34*, 209–253.

Kamble, V., Shah, N., Marn, D., Parekh, A., & Ramachandran, K. (2015). Truth serums for massively crowdsourced evaluation tasks. *arXiv preprint arXiv:1507.07045*.

Kulkarni, C. E., Socher, R., Bernstein, M. S., & Klemmer, S. R. (2014). Scaling short-answer grading by combining peer assessment with algorithmic scoring. In *Proceedings of the first acm conference on learning@ scale conference* (pp. 99–108).

Lev, O., Mattei, N., Turrini, P., & Zhydkov, S. (2023). Peernomination: A novel peer selection algorithm to handle strategic and noisy assessments. *Artificial Intelligence*, *316*, 103843.

Li, Y. (2020). Mechanism design with costly verification and limited punishments. *Journal of Economic Theory*, *186*, 105000.

Mazar, N., Amir, O., & Ariely, D. (2008). The dishonesty of honest people: A theory of self-concept maintenance. *Journal of marketing research*, *45*(6), 633–644.

Miller, N., Resnick, P., & Zeckhauser, R. (2005). Eliciting informative feedback: The peer-prediction method. *Management Science*, *51*(9), 1359–1373.

Noothigattu, R., Shah, N., & Procaccia, A. (2021). Loss functions, axioms, and peer review. *Journal of Artificial Intelligence Research*, *70*, 1481–1515.

Paré, D. E., & Joordens, S. (2008). Peering into large lectures: examining peer and expert mark agreement using peerscholar, an online peer assessment tool. *Journal of Computer Assisted Learning*, *24*(6), 526–540.

Piech, C., Huang, J., Chen, Z., Do, C., Ng, A., & Koller, D. (2013). Tuned models of peer assessment in MOOCs. *arXiv preprint arXiv:1307.2579*.

Prelec, D. (2004). A bayesian truth serum for subjective data. *science*, *306*(5695), 462–466.

Radanovic, G., & Faltings, B. (2015). Incentives for subjective evaluations with private beliefs. In *Proceedings of the 29th aaai conference on artificial intelligence (aaai 15)* (pp. 1014–1020).

Raman, K., & Joachims, T. (2014). Methods for ordinal peer grading. In *Proceedings of the 20th acm sigkdd international conference on knowledge discovery and data mining* (pp. 1037–1046).

Sadler, P. M., & Good, E. (2006). The impact of self-and peer-grading on student learning. *Educational assessment*, *11*(1), 1–31.

Shah, N. B. (2022). Challenges, experiments, and computational solutions in peer review. *Communications of the ACM*, *65*(6), 76–87.

Shah, N. B., Bradley, J. K., Parekh, A., Wainwright, M., & Ramchandran, K. (2013). A case for ordinal peer-evaluation in moocs. In *Nips workshop on data driven education* (pp. 1–8).

Shnayder, V., Agarwal, A., Frongillo, R., & Parkes, D. C. (2016). Informed truthfulness in multi-task peer prediction. In *Proceedings of the 2016 acm conference on economics and computation* (pp. 179–196).

Waggoner, B., & Chen, Y. (2014). Output agreement mechanisms and common knowledge. In *Second aaai conference on human computation and crowdsourcing.*

Wang, J., Stelmakh, I., Wei, Y., & Shah, N. B. (2021). Debiasing evaluations that are biased by evaluations. In *Proceedings of the aaai conference on artificial intelligence* (Vol. 35, pp. 10120–10128).

Witkowski, J., Bachrach, Y., Key, P., & Parkes, D. C. (2013). Dwelling on the negative: Incentivizing effort in peer prediction. In *First aaai conference on human computation and crowdsourcing.*

Witkowski, J., & Parkes, D. C. (2013). Learning the prior in minimal peer prediction. In *Proceedings of the 3rd workshop on social computing and user generated content at the acm conference on electronic commerce* (Vol. 14).

Wright, J. R., Thornton, C., & Leyton-Brown, K. (2015). Mechanical ta: Partially automated high-stakes peer grading. In *Proceedings of the 46th acm technical symposium on computer science education* (pp. 96–101).

Zarkoob, H., d'Eon, G., Podina, L., & Leyton-Brown, K. (2022). Better peer grading through bayesian inference. *arXiv preprint arXiv:2209.01242*.

# Appendices

## Appendix A. Simpler description of PEQA

Algorithm 2 gives the description.

**Algorithm 2** PEQA

1: Inputs: (1) the parameters $\mu$ and $\gamma$ of the priors on $y_j$, $\forall j \in N$, which is distributed as $\mathcal{F}(\mu, 1/\gamma)$, (2) the reported scores $\tilde{\mathbf{y}}_P^N$ of the graders on the probe papers, and (3) reported scores $\tilde{\mathbf{y}}_{N\setminus P}^N$ on the non-probe papers.

2: Set the probe set $P$ with $|P| = \ell$, a pre-determined constant $\leqslant \frac{n}{\frac{K}{2}+1}$, where $K$ (even) is the number of papers assigned to each grader.

3: $G = G^*$: every grader $i \in N$ is assigned $\frac{K}{2}$ probe and $\frac{K}{2}$ non-probe papers, in such a way that every non-probe paper is assigned to at least $\frac{K}{2}$ and at most $\frac{K}{2}+1$ graders. This is always possible by assigning the $(n-\ell)$ non-probe papers to $(n-\ell)$ graders with each paper assigned to exactly $\frac{K}{2}$ graders. The rest $\ell$ graders can be assigned to the same $(n-\ell)$ papers arbitrarily such that these papers get at most one additional grader (since $\ell\frac{K}{2} \leqslant n-\ell$). Note that this is the reason $\ell$ cannot be larger than $n/(K/2+1)$. Ensure that a grader does not get her own paper for evaluation.

4: Estimate $\hat{b}_i, \hat{\tau}_i, \forall i \in N$ as given in Section 2.3.

5: $\mathbf{r}$: the score of the paper $j$ is given by the ISWDM $\mathbf{r}^*$ (Equation (3)).

6: At this stage, students may request for regrading. Instructor learns the correct grade $y_j$ for the papers which came for regrading. For other papers, $y_j = r_j^*$ is assumed.

7: $\mathbf{t}$: the performance score to grader $i$ for grading paper $j \in NP_i$ is given by $t_i^j = \alpha(W_j^* - W_j^{(-i)*})$, where $\alpha > 0$ is a constant chosen at the designer's discretion. The total performance score to grader $i$ is therefore $t_i = \sum_{j \in NP_i} t_i^j$.

## Appendix B. Omitted Proofs

### B.1 Proof of Lemma 1

Since $K \geqslant 2, \ell \leqslant \frac{n}{\frac{K}{2}+1} \leqslant \frac{n}{2}$. Hence, there are more non-probe papers than probes. Also, since $\ell \geqslant \frac{K}{2}+1$ and computeG gives grader $i$ gives $K/2$ probe and $K/2$ non-probe papers starting from $i+1$, the grader can never get her own paper. To see this, note that the papers are given in a round-robin manner individually within the pools of probe and non-probe papers. Since the probes are fewer in number, it is sufficient to argue that probe paper $i$ does not go to agent $i$. This is obvious since $\ell \geqslant \frac{K}{2}+1$. For the non-probes, the round-robin cycle is larger and therefore the non-assignment of a paper to the same agent is maintained.

Now, consider the *coverage* issue of each non-probe paper by the graders. We will call a grader a *probe(non-probe) grader* if her paper is a probe(non-probe). Since the $(n-\ell)$ non-probe agents get a round-robin assignment of size $K/2$ from the non-probe papers, each of these papers is *covered* by exactly $K/2$ non-probe graders. Additionally, there are $\ell$ probe graders and each of them gets $K/2$ non-probe papers. Therefore, $\ell\frac{K}{2}$ more graders are uniformly assigned on the non-probe papers. But since $\ell \leqslant \frac{n}{\frac{K}{2}+1}$, i.e., $\ell\frac{K}{2} \leqslant (n-\ell)$,

each non-probe paper may get covered by one extra probe grader. This proves that each non-probe paper has a grader coverage of at least $K/2$ and at most $K/2+1$.

### B.2 Proof of Theorem 1

By Assumption 1, the student knows her $y_j$ perfectly and if $r_j^* \geqslant y_j$, she does not raise a regrading request. PEQA will assume $r_j^*$ to be the true score and design the peer-grading performance score accordingly when there is no regrading request. The student asks for regrading *only if* $r_j^* < y_j$. The utility of grader $i$ after the regrading requests have been addressed is (we have omitted the arguments of the functions in Equation (7) where it is understood) therefore

$$u_i(\cdot) = \max\{r_i^*(\cdot), y_i\} + t_i - \left( \sum_{j \in G^{-1}(i)} w_{ij} \max\{r_j^*(\cdot), y_j\} + \sum_{k \in CG_i \setminus \{i\}} w_{ik} t_k \right) - \phi(\cdot) \quad (19)$$

We decomposed the utility expression to gather together the terms that are affected by the grading (and hence, the choices of $b_i, \tau_i$) of student $i$. They are (a) the exam scores of the papers graded by $i$ (first term in the parentheses), and (b) the peer-grading performance score of the co-graders of $i$ (second term in the parentheses). The function $\phi$ is the remaining part of $u_i$ that is independent of $i$'s grading. Hence, taking expectation for agent $i$ as defined in Definition 2, we get

$$\begin{aligned} \mathbb{E}_{i,b_i,\tau_i}\ u_i(\cdot) &= \max\{r_i^*(\cdot), y_i\} + \mathbb{E}_{i,b_i,\tau_i}\ t_i \\ &\quad - \left( \sum_{j \in G^{-1}(i)} w_{ij} \mathbb{E}_{i,b_i,\tau_i}\ \max\{r_j^*(\cdot), y_j\} + \sum_{k \in CG_i \setminus \{i\}} w_{ik} \mathbb{E}_{i,b_i,\tau_i}\ t_k \right) - \phi(\cdot) \end{aligned} \quad (20)$$

We prove that PEQA is EPBI and EPRM in *four* steps. First, we observe that the first term on the RHS is independent of the values of $b_i$ and $\tau_i$. In the second step, we show that each summand $\max\{r_j^*(\cdot), y_j\}$ in the first summation term is independent of $b_i$ and decreasing in $\tau_i$. The third step shows that $t_i$ is independent of $b_i$ and increasing in $\tau_i$, and the fourth step shows that this conclusion is true even for $t_i - \sum_{k \in CG_i \setminus \{i\}} w_{ik} t_k$ for the sufficient condition of the theorem.

**Step 1: $\max\{r_i^*(\cdot), y_i\}$ is independent of the values of $b_i$ and $\tau_i$.** This is obvious since student $i$ does not grade her own paper and hence she has no control on the grade given by PEQA on her paper.

**Step 2: Each individual term $\mathbb{E}_{i,b_i,\tau_i} \max\{r_j^*(\cdot), y_j\}$ is independent of $b_i$ and increasing in $\sigma_i$.** Note that $\mathbb{E}_{i,b_i,\tau_i} \equiv \mathbb{E}_{y_k \sim \mathcal{F}(\mu, 1/\gamma),\ k \in G^{-1}(i)} \mathbb{E}_{\tilde{y}_k^{(i)} | y_k \sim \mathcal{F}(y_k + b_i, 1/\tau_i),\ k \in G^{-1}(i)}$. Now, as the processes that determine the true scores and thence the reported scores conditional on the true on exams $j \in G^{-1}(i)$ are all mutually independent, for any fixed exam $j$ graded by $i$, $\mathbb{E}_{i,b_i,\tau_i} \max\{r_j^*(\cdot), y_j\}$ simplifies to $\mathbb{E}_{y_j \sim \mathcal{F}(\mu,1/\gamma)} \mathbb{E}_{\tilde{y}_j^{(i)} | y_j \sim \mathcal{F}(y_j + b_i, 1/\tau_i)} \max\{r_j^*(\cdot), y_j\}$.

Next, recall that the score-assignment function for PEQA is ISWDM (Definition 1)

$$r_j^*(\tilde{\mathbf{y}}_j^{G(j)}, \hat{\boldsymbol{\theta}}_{G(j)}) = \frac{\sqrt{\gamma}\mu + \sum_{i \in G(j)} \sqrt{\hat{\tau}_i}(\tilde{y}_j^{(i)} - \hat{b}_i)}{\sqrt{\gamma} + \sum_{i \in G(j)} \sqrt{\hat{\tau}_i}}.$$

The final grade after regrading is

$$\max\{r_j^*(\cdot), y_j\} = \max\{r_j^*(\cdot) - y_j, 0\} + y_j. \tag{21}$$

Grader $i$'s estimated bias is given by $\hat{b}_i = \frac{\sum_{k\in P_i}(\tilde{y}_j^{(i)} - y_j)}{x}$, where $x = |P_i|$. In PEQA, we use the same number $K/2$ as $|P_i|$, for all $i$. Hence, $x = K/2$, is a constant in our analysis.

Given our model of peer-reports, $\tilde{y}_j^{(i)} = y_j + b_i + n_{ij}$, where $n_{ij} \sim \mathcal{F}(0, 1/\tau_i)$ is a noise term. Hence, it is easy to show that $\hat{b}_i = b_i + \frac{\sum_{k\in P_i} n_{ik}}{x}$ and $\frac{1}{\hat{\tau}_i} = \hat{\sigma}_i^2 = \frac{\sum_{k\in P_i}(n_{ik} - \frac{1}{x}\sum n_{ik})^2}{x}$, where $n_{ik} \sim \mathcal{F}(0, \sigma_i^2)$.

Substituting these values we get the expression for

$$r_j^*(\cdot) - y_j = \frac{\sqrt{\gamma}(\mu - y_j) + \sum_{l\in G(j)} \sqrt{\hat{\tau}_l}(n_{lj} - \frac{\sum_{k\in P_l} n_{lk}}{x})}{\sqrt{\gamma} + \sum_{l\in G(j)} \sqrt{\hat{\tau}_l}}. \tag{22}$$

Note that $z_j = \sqrt{\gamma}(\mu - y_j)$ is a $\mathcal{F}(0,1)$ variable, that is independent of all the other variables in the expression. In the following, we take the expectation of the term $\max\{r_j^*(\cdot) - y_j, 0\}$ w.r.t. $z_j$ and show that it is independent of $b_i$ and increasing in $\sigma_i = 1/\sqrt{\tau_i}$, which implies that irrespective of the values of the other graders' biases and reliabilities, it is best for grader $i$ to reduce her $\sigma_i$ to increase this component of her utility (since the term comes with a negative sign in the utility expression).

Note that among other things, this also means that we are changing the order of expectations in $\mathbb{E}_{y_j\sim\mathcal{F}(\mu,1/\gamma)}\mathbb{E}_{\tilde{y}_j^{(i)}|y_j\sim\mathcal{F}(y_j+b_i,1/\tau_i)} \max\{r_j^*(\cdot), y_j\}$. We are first taking the expectation $\mathbb{E}_{y_j\sim\mathcal{F}(\mu,1/\gamma)}$ after a change of variable $z_j = \sqrt{\gamma}(\mu - y_j)$, and we call this term $I_j$. Then, we take expectation of $I_j$ w.r.t. $\mathbb{E}_{\tilde{y}_j^{(i)}|y_j}$, which is the same as integrating w.r.t $n_{ij}$.[26]

$$\begin{aligned}
I_j &= E_{z_j} \max\{r_j^*(\cdot) - y_j, 0\} \\
&= \int_{-\sum_{l\in G(j)} \sqrt{\hat{\tau}_l}(n_{lj} - \frac{\sum_{k\in P_l} n_{lk}}{x})}^{\infty} \frac{z_j + \sum_{l\in G(j)} \sqrt{\hat{\tau}_l}(n_{lj} - \frac{\sum_{k\in P_l} n_{lk}}{x})}{\sqrt{\gamma} + \sum_{l\in G(j)} \sqrt{\hat{\tau}_l}} f(z_j)dz_j + 0 \\
&= \frac{1}{\sqrt{\gamma} + \sum_{l\in G(j)} \sqrt{\hat{\tau}_l}} \int_{-\sum_{l\in G(j)} \sqrt{\hat{\tau}_l}(n_{lj} - \frac{\sum_{k\in P_l} n_{lk}}{x})}^{\infty} \left( z_j + \sum_{l\in G(j)} \sqrt{\hat{\tau}_l}(n_{lj} - \frac{\sum_{k\in P_l} n_{lk}}{x}) \right) \times \\
&\qquad\qquad f(z_j)dz_j \\
&= \frac{1}{\sqrt{\gamma} + \sum_{l\in G(j)} \sqrt{\hat{\tau}_l}} \int_0^{\infty} v_j f(v_j - \sum_{l\in G(j)} \sqrt{\hat{\tau}_l}(n_{lj} - \frac{\sum_{k\in P_l} n_{lk}}{x}))dv_j \\
&= \frac{1}{\sqrt{\gamma} + \sum_{l\in G(j)} \sqrt{\hat{\tau}_l}} \int_0^{\infty} v_j f \left( v_j - \sum_{l\in G(j)\setminus\{i\}} \sqrt{\hat{\tau}_l}(n_{lj} - \frac{\sum_{k\in P_l} n_{lk}}{x}) \right. \\
&\qquad\qquad \left. - \sqrt{\hat{\tau}_i}(n_{ij} - \frac{\sum_{k\in P_i} n_{ik}}{x}) \right) dv_j
\end{aligned}$$

[26] If the original integrand is integrable, its absolute value must also be integrable, and thus one can use Fubini's theorem to change the order of expectation.

$$= \frac{1}{\sqrt{\gamma}+\sum_{l\in G(j)}\sqrt{\hat{\tau}_l}}\int_0^\infty v_j f\left(v_j - \sum_{l\in G(j)\setminus\{i\}}\sqrt{\hat{\tau}_l}(n_{lj}-\frac{\sum_{k\in P_l}n_{lk}}{x}) - \frac{(m_{ij}-\frac{\sum_{k\in P_i}m_{ik}}{x})\sigma_i}{\sqrt{\frac{\sum_{k\in P_i}(m_{ik}-\frac{1}{x}\sum m_{ik})^2}{x}\sigma_i^2}}\right)dv_j$$

$$= \frac{1}{\sqrt{\gamma}+\sum_{l\in G(j)\setminus\{i\}}\sqrt{\hat{\tau}_l}+1/\sqrt{\frac{\sum_{k\in P_i}(m_{ik}-\frac{1}{x}\sum m_{ik})^2}{x}\sigma_i}}\times$$

$$\int_0^\infty v_j f(v_j - \sum_{l\in G(j)\setminus\{i\}}\sqrt{\hat{\tau}_l}(n_{lj}-\frac{\sum_{k\in P_l}n_{lk}}{x}) - \frac{(m_{ij}-\frac{\sum_{k\in P_i}m_{ik}}{x})}{\sqrt{\frac{\sum_{k\in P_i}(m_{ik}-\frac{1}{x}\sum m_{ik})^2}{x}}})dv_j \tag{23}$$

In the third equality, we have substituted $v_j = z_j + \sum_{l\in G(j)}\sqrt{\hat{\tau}_l}(n_{lj}-\frac{\sum_{k\in P_l}n_{lk}}{x})$, and in the fifth equality, we substituted $n_{ik} = m_{ik}\cdot\sigma_i$. Since $n_{ik}\sim\mathcal{F}(0,\sigma_i^2)$, we get $m_{ik}\sim\mathcal{F}(0,1)$. Note that $f$ is the density of a $\mathcal{F}(0,1)$ random variable. Hence the whole expression within the integral is independent of $\sigma_i$. It is easy to see that the pre-multiplied term is increasing in $\sigma_i$. Hence, we conclude that the integral $I_j$ is independent of $b_i$ and increasing in $\sigma_i = 1/\sqrt{\tau_i}$.

For any integral outside $I_j$ over any $n_{ik} =_{m_{ik}} \times\sigma_i$, we perform a change of variable to make it an integral over $m_{ik}$, and there are no extra $\sigma_i$ terms originating there as $f_{n_{ik}}(n_{ik})dn_{ik}{=}f_{m_{ik}}(m_{ik})dm_{ik}$.[27]

**Step 3: The expected value of $t_i^j$ is independent of $b_i$ and decreasing in $\sigma_i$.** We assumed in Section 2 that the reward function is decreasing in the difference $|r_j^* - y_j|$ and the mechanism assigns reward to be zero when $r_j^* > y_j$. Hence, we calculate the condition on $y_j$ when the reward is non-zero.

$$r_j^*(\tilde{\mathbf{y}}_j^{G(j)}, \hat{\boldsymbol{\theta}}_{G(j)}) \leqslant y_j \iff \frac{\sqrt{\gamma}\mu + \sum_{l\in G(j)}\sqrt{\hat{\tau}_l}(\tilde{y}_j^{(l)}-\hat{b}_l)}{\sqrt{\gamma}+\sum_{l\in G(j)}\sqrt{\hat{\tau}_l}} \leqslant y_j$$

$$\iff \sqrt{\gamma}\mu + \sum_{l\in G(j)}\sqrt{\hat{\tau}_l}(\tilde{y}_j^{(l)}-\hat{b}_l) \leqslant y_j(\sqrt{\gamma}+\sum_{l\in G(j)}\sqrt{\hat{\tau}_l})$$

$$\iff y_j\sqrt{\gamma} \geqslant \sqrt{\gamma}\mu + \sum_{l\in G(j)}\sqrt{\hat{\tau}_l}(\tilde{y}_j^{(l)}-\hat{b}_l-y_j) = \sqrt{\gamma}\mu + \sum_{l\in G(j)}\sqrt{\hat{\tau}_l}(n_{lj}-\frac{\sum_{k\in P_l}n_{lk}}{x})$$

$$\iff y_j \geqslant \frac{\sqrt{\gamma}\mu + Z + \sqrt{\hat{\tau}_i}(n_{ij}-\frac{\sum_{k\in P_i}n_{ik}}{x})}{\sqrt{\gamma}}, \quad \text{where } Z = \sum_{l\in G(j)\setminus\{i\}}\sqrt{\hat{\tau}_l}(n_{lj}-\frac{\sum_{k\in P_l}n_{lk}}{x})$$

$$= \frac{\sqrt{\gamma}\mu + Z}{\sqrt{\gamma}} + \frac{(m_i-\frac{\sum_{k\in P_i}m_{ik}}{x})\sigma_i}{\sqrt{\gamma}\sqrt{\frac{\sum_{k\in P_i}(m_{ik}-\frac{1}{x}\sum m_{ik})^2}{x}\sigma_i^2}}$$

Note that the RHS is independent of $\sigma_i$. Hence the limits of the integral where the reward $R$ is non-zero is also independent of $\sigma_i$.

[27] Note that as $F_{n_{ik}}(\sigma_i x) = F_{m_{ik}}(x)$, differentiation on both sides gives $\sigma_i f_{n_{ik}}(\sigma_i x) = f_{m_{ik}}(x)$, where $F$ and $f$ are the CDF and PDF respectively.

By definition, the $W_j^{(-i)*}$ component of the performance score is independent of bias and reliability of grader $i$. Hence, we only consider the first component which is dependent on the bias and reliability of grader $i$. We will consider the integral only w.r.t. $y_j$ to compute $t_i^j$ and we just showed that the limits of this integral is independent of $\sigma_i$. Hence, if we show that the reward function $R(r_j^*, y_j)$ is independent of $b_i$ and decreasing in $\sigma_i$, then we are done. Consider the argument of the reward function

$$
\begin{aligned}
r_j^* - y_j &= \frac{\sqrt{\gamma}(\mu - y_j) + \sum_{l \in G(j)} \sqrt{\hat{\tau}_l}(n_{lj} - \frac{\sum_{k \in P_l} n_{lk}}{x})}{\sqrt{\gamma} + \sum_{l \in G(j)} \sqrt{\hat{\tau}_l}} \\
&= \frac{[\sqrt{\gamma}(\mu - y_j) + \sum_{l \in G(j) \setminus \{i\}} \sqrt{\hat{\tau}_l}(\tilde{y}_j^{(l)} - \hat{b}_l - y_j)] \sqrt{\frac{\sum_{k \in P_i} (n_{ik} - \frac{1}{x} \sum_{l \in P_i} n_{il})^2}{x}} + (n_j - \frac{1}{x} \sum_{l \in P_i} n_{il})}{(\sqrt{\gamma} + \sum_{l \in G(j) \setminus \{i\}} \sqrt{\hat{\tau}_l}) \sqrt{\frac{\sum_{k \in P_i} (n_{ik} - \frac{1}{x} \sum_{l \in P_i} n_{il})^2}{x}} + 1} \\
&= \frac{Z_{-i} \sqrt{\frac{\sum_{k \in P_i} (m_{ik} - \frac{1}{x} \sum_{l \in P_i} m_{il})^2}{x}} \cdot \sigma_i + \sigma_i \cdot (m_j - \frac{1}{x} \sum_{l \in P_i} m_{il})}{X_{-i} \sqrt{\frac{\sum_{k \in P_i} (m_{ik} - \frac{1}{x} \sum_{l \in P_i} m_{il})^2}{x}} \cdot \sigma_i + 1}
\end{aligned} \tag{24}
$$

In the last equality, we substituted $Z_{-i} = [\sqrt{\gamma}(\mu - y_j) + \sum_{l \in G(j) \setminus \{i\}} \sqrt{\hat{\tau}_l}(\tilde{y}_j^{(l)} - \hat{b}_l - y_j)]$ and $X_{-i} = (\sqrt{\gamma} + \sum_{l \in G(j) \setminus \{i\}} \sqrt{\hat{\tau}_l})$. As before, we substituted $n_{ik} = m_{ik} \cdot \sigma_i$. Since $n_{ik} \sim \mathcal{F}(0, \sigma_i^2)$, we get $m_{ik} \sim \mathcal{F}(0, 1)$. We see that the absolute value of the above expression is independent of $b_i$ and increasing in $\sigma_i$. Hence $R(r_j^*, y_j)$ is independent of $b_i$ and decreasing in $\sigma_i$.

**Step 4: $t_i^j - \sum_{k \in CG_i^j \setminus \{i\}} w_{ik} t_k^j$ is independent of $b_i$ and decreasing in $\sigma_i$ for $\sum_{k \in N \setminus \{i\}} w_{ik} \leqslant 1$.**

First, we show that $t_i^j - t_k^j$ is independent of $b_i$ and decreasing in $\sigma_i$. This is because $W_j^*$ cancels and this difference reduces to $W_j^{(-k)*} - W_j^{(-i)*}$. The second term is independent of $b_i$ and $\sigma_i$. The first term is independent of $b_i$ and decreasing in $\sigma_i$ by the same argument as step 3, with the set of graders reduced to $N \setminus \{k\}$.

Observe that, in the utility of grader $i$, the difference in these two performance score terms appear as follows.

$$
t_i - w_{ik} \cdot \sum_{k \in CG_i \setminus \{i\}} t_k = \sum_{j \in NP_i} \left( t_i^j - w_{ik} \cdot \sum_{k \in CG_i^j \setminus \{i\}} t_k^j \right).
$$

Consider the terms in the parentheses on the RHS.

$$
t_i^j - w_{ik} \cdot \sum_{k \in CG_i^j \setminus \{i\}} t_k^j = \cdot \sum_{k \in CG_i^j \setminus \{i\}} w_{ik} \cdot (t_i^j - t_k^j) + (1 - \sum_{k \in CG_i^j \setminus \{i\}} w_{ik}) t_i^j.
$$

Both terms in the RHS is independent of $b_i$ and decreasing in $\sigma_i$ as we have already shown and since $\sum_{k \in CG_i^j \setminus \{i\}} w_{ik} \leqslant \sum_{k \in N \setminus \{i\}} w_{ik} \leqslant 1$. (Note that the first inequality can also be written as $\leqslant \max_{i \in N} \max_{A \subset \mathcal{F}^i{}_{K/2}} \sum_{k \in A} w_{ik}$, since $|CG_i^j| \leqslant K/2 + 1$, which proves the other version of the theorem with a weaker sufficient condition.)

Combining all steps, we have shown that the expected utility of grader $i$, where the expectation is taken as $\mathbb{E}_{i,b_i,\tau_i}$ is independent of $b_i$ and decreasing in $\sigma_i$. Hence these two properties hold for any choice of actions by the other graders. Hence we have proved that `PEQA` is EPBI and EPRM.

### B.3 Proof of Theorem 2

Since the given condition of the theorem is an arbitrary performance score function $\mathbf{t}$, we choose some performance score function $\mathbf{t}'$ and take $\mathbf{t} \equiv \delta\mathbf{t}'$, where $\delta > 0$ is arbitrary. We use the $\delta$ as a scaling factor, which can be increased or decreased arbitrarily to leverage the arbitrary choice of a performance score function. We will show that for every $\delta > 0$, the claim of the theorem holds as discussed below.

Consider the utility expression for agent $i$ with peer-grading performance score weight $\delta$ (we consider only the component of $u_i(\cdot)$ that depends on agent $i$'s bias and reliability, from Equation (20))

$$\begin{aligned} u_i(\cdot) &= \max\{r_i^{\text{WA}}(\cdot), y_i\} + \delta t'_i - \sum_{j \in G^{-1}(i)} w_{ij} \cdot \max\{r_j^{\text{WA}}(\cdot), y_j\} - \delta \sum_{k \in CG_i \setminus \{i\}} w_{ik} \cdot t'_k \\ &= \max\{r_i^{\text{WA}}(\cdot), y_i\} - \sum_{j \in G^{-1}(i)} w_{ij} \cdot \max\{r_j^{\text{WA}}(\cdot), y_j\} + \delta \left( t'_i - \sum_{k \in CG_i \setminus \{i\}} w_{ik} \cdot t'_k \right) \end{aligned} \tag{25}$$

Taking expectation, we get

$$\begin{aligned} \mathbb{E}_{i,b_i,\tau_i}\ u_i(\cdot) = \max\{r_i^{\text{WA}}(\cdot), y_i\} &- \sum_{j \in G^{-1}(i)} w_{ij} \cdot \underbrace{\mathbb{E}_{i,b_i,\tau_i}\ \max\{r_j^{\text{WA}}(\cdot), y_j\}}} \\ &+ \delta \mathbb{E}_{i,b_i,\tau_i} \left( t'_i - \sum_{k \in CG_i \setminus \{i\}} w_{ik} \cdot t'_k \right) \end{aligned} \tag{26}$$

The first term on the RHS, $i$'s own grade, is independent of $\sigma_i$. Hence, we will focus on the other two terms.

We will show that for all realizations of the random variables as defined in EPRM (Definition 3), and for all values of $\delta$, the utility is monotone decreasing in $\sigma_i$ only if $\lambda_i(\sigma_i) = \kappa_i/\sigma_i$. For brevity of notation, we will use $\lambda_i$ to denote the function where the argument is clear from the context.

We claim that for EPRM to hold for all $\delta$, it is necessary that the second term of Equation (26) (including the negative sign) is monotonically non-increasing in $\sigma_i$ for all realizations of the random variables. Suppose not, i.e., there exists some $\sigma'_i$, where the second term is increasing in $\sigma_i$. Then at that $\sigma'_i$, there exists a $\delta_{\sigma'_i} > 0$, sufficiently small, such that the sign of the derivative (w.r.t. $\sigma_i$) of the expected utility (LHS of Equation (26)) is determined by the sign of the derivative of the second term. This implies that the overall utility will be increasing in $\sigma_i$ for that choice of $\delta_{\sigma'_i} > 0$ at $\sigma'_i$. This violates EPRM. Hence the claim.

With that background, our objective is now to show that each underbraced individual term inside the second term in the RHS of Equation (26), $\mathbb{E}_{i,b_i,\tau_i}\ \max\{r_j^*(\cdot), y_j\}$, is mono-

tonically non-decreasing in $\sigma_i$ only if $\lambda_i(\sigma_i) = \kappa_i/\sigma_i$.[28] Define the following terms to shorten the forthcoming expressions.

$$K_1 = \lambda_0 + \sum_{l \in G(j)\setminus\{i\}} \lambda_l,\ K_j = \sum_{l \in G(j)\setminus\{i\}} \lambda_l \left( n_{lj} - \frac{\sum_{k \in P_l} n_{lk}}{x} \right),\ K_{3j} = m_{ij} - \frac{\sum_{k \in P_i} m_{ik}}{x}.$$

We follow the same set of arguments, particularly on changing the order of expectation and subsequent simplification, as we did after Equation (21). Consider the second term in Equation (26). Using Equation (21), we reduce the expression for paper $j$ in the sum into the difference term and consider its expectation w.r.t. $z_j = \lambda_0(\mu - y_j)$, which is a $\frac{\lambda_0}{\sqrt{\gamma}}\mathcal{F}(0,1)$ random variable, to get a similar expression like Equation (23) as follows. We ignore the positive constant $\frac{\lambda_0}{\sqrt{\gamma}}$ as it does not play a role in determining the sign of the variation.

$$I_j = E_{z_j \sim \mathcal{F}(0,1)} \max\{r_j^*(\cdot) - y_j, 0\} = \frac{1}{K_1 + \lambda_i} \times \int_0^\infty v_j f(v_j - K_j - K_{3j}\lambda_i) dv_j$$

To find the change w.r.t. $\sigma_i$, we take its partial derivative and find $\frac{\partial I_j}{\partial \sigma_i}$ to be

$$\frac{-K_{3j}\left(\frac{\partial(\lambda_i \cdot \sigma_i)}{\partial \sigma_i}\right)(K_1 + \lambda_i) \cdot \int_0^\infty v_j f'(v_j - K_j - K_{3j}\lambda_i) dv_j - \left(\frac{\partial(\lambda_i)}{\partial \sigma_i}\right) \cdot \int_0^\infty v_j f(v_j - K_j - K_{3j}\lambda_i) dv_j}{(K_1 + \lambda_i)^2}$$

Note that $K_1$ is positive, while $K_j$ and $K_{3j}$ can take any sign. To ensure that the expression above is non-negative for *all* values of the realized random variables, it is necessary and sufficient that

$$\frac{\partial(\lambda_i \cdot \sigma_i)}{\partial \sigma_i} = 0, \quad \text{and} \quad \frac{\partial(\lambda_i)}{\partial \sigma_i} \leqslant 0. \tag{27}$$

This is because the second integral in the numerator is always positive. Therefore, the above condition is equivalent to $\lambda_i = \kappa_i/\sigma_i$, where $\kappa_i > 0$ is a factor independent of $\sigma_i$. This concludes the proof.

### B.4 Proof of Theorem 3

We compute the complexity of `PEQA` with reference to the steps in Algorithm 1.

- ▷ Line 1 requires running `computeG` function which will compute the set of graders for each paper. This step takes $O(nK)$ time.
- ▷ In Line 2, computation of $\hat{b}_i$ and $\hat{\tau}_i$ for a grader $i$ is $O(K)$ since a grader is given $K/2$ probe papers. So the total time complexity for $n$ graders is $O(nK)$.

[28] Note that, in case one of the underbraced terms ($\mathbb{E}_{i,b_i,\tau_i} \max\{r_j^*(\cdot), y_j\}$) of Equation (26) is decreasing for some exam $j^*$ and for some realization of the random variables of the graders except $i$ (i.e., $\{\tilde{y}_j^{(k)}, b_k, \tau_k\}_{k \neq i}$) on that exam $j^*$, one could replicate the same realizations of the same random variables for all the other exams $j \in G^{-1}(i) \setminus \{j^*\}$. Then, the $\mathbb{E}_{i,b_i,\tau_i} \max\{r_j^*(\cdot), y_j\}$ terms would be decreasing for all $j \in G^{-1}(i)$.

▷ In Line 3, the computation of score for paper $j$, i.e., $r_j$, is $O(K)$ since every non-probe paper is given to at most $K/2+1$ graders as from the `computeG` function. So, the total time to compute scores of all the papers is $O(nK)$.

▷ Lines 4, 5 do not contribute to the computational complexity.

▷ In the `for` loop of Line 6, the computation of $y_j, \forall j \in N$ would require $O(n)$ complexity.

▷ In Line 11, we claim that calculating `computet` takes $O(nK)$ time. After that getting performance for each grader is just $O(n)$.

Lemma **3** `computet` *can be calculated in* $O(nK)$ *time.*

*Proof*: the computation of $W_j^*$ can be retrieved in $O(1)$ since we can store the $r_j^*$ result computed in Line 3 (and the corresponding accuracy $W_j^*$) in a hash map. Note that we can calculate $r_j^{(-i)*}\ \forall i \in G(j)$ in $O(K)$ as well if we precompute the numerator and denominator of $r_j^*$ in $O(K)$ time and then subtract the grader $i$'s contribution $(\sqrt{\hat{\tau}_i}(\tilde{y}_j^{(i)} - \hat{b}_i))$ from the precomputed numerator and denominator (subtract $\sqrt{\hat{\tau}_i}$) in $O(1)$ time to calculate $r_j^{(-i)*}$. By doing so, the computation of $W_j^{(-i)*}, \forall i \in G(j)$ takes $O(K)$ time. Hence, the computation of $t_i^j, \forall i \in G(j)$ takes $O(K)$ time. Computing this quantity for all non-probe papers, i.e., $t_i^j, \forall i \in G(j), \forall j \in N \setminus P$ will take $O(nK)$ time. Therefore, the computation of $t_i \forall i \in N$ would just take $O(nK)$ time. ■

Hence, the overall time complexity of `PEQA` is $O(nK)$.

## Appendix C. Calculation of $r_j^{\mathbf{ERM}}(\tilde{\mathbf{y}}_j^{G(j)}; \hat{\boldsymbol{\theta}}_{G(j)})$ under $\mathbf{PG}_1$ (Piech et al., 2013) model:

Below, we find a score-assignment function $r_j^{\mathrm{ERM}}(\tilde{\mathbf{y}}_j^{G(j)}, \hat{\boldsymbol{\theta}}_{G(j)})$ that would maximize a quadratic reward function, i.e., that would minimize the expected squared distance between the assigned score and true score on exam $j$. We will calculate the expression w.r.t. the true error parameters $\boldsymbol{\theta}$. Then, given $\boldsymbol{\theta}$ is not observed, we will approximate $\boldsymbol{\theta}$ with the estimated value of the same parameters, i.e., $\hat{\boldsymbol{\theta}}$ to find a new expression.

To calculate $r_j^{\mathrm{ERM}}(\tilde{\mathbf{y}}_j^{G(j)}; \boldsymbol{\theta}_{G(j)})$, we first need to calculate the conditional distribution of the true score $y_j$, $\psi(y_j | \tilde{y}_j^{G(j)}; b_{G(j)}, \tau_{G(j)}, \mu, \gamma)$ under the $\mathbf{PG}_1$ (Piech et al., 2013) model, where $\psi(\cdot)$ is the density of the normal distribution with the mean and variance given by that model. For the convenience of the reader, we restate the $\mathbf{PG}_1$ model below.

**Model $\mathbf{PG}_1$ (grader bias and reliability)** This model puts prior distributions over the latent variables and assumes for example that while an individual grader's bias may be nonzero, the average bias of many graders is zero. Specifically,

$$\text{(Reliability) } \tau_v \sim \mathcal{G}(\alpha_0, \beta_0) \text{ for every grader } v,$$

$$\begin{aligned}
\text{(Bias) } b_v &\sim \mathcal{N}(0, 1/\eta_0) \text{ for every grader } v,\\
\text{(True score) } s_u &\sim \mathcal{N}(\mu_0, 1/\gamma_0) \text{ for every user } u,\\
\text{(Observed score) } z_u &\sim \mathcal{N}(s_u + b_v, 1/\tau_v) \text{ for every observed peer grade } s_u,
\end{aligned}$$

where $\mathcal{G}$ and $\mathcal{N}$ refers to the gamma and normal distributions respectively with appropriate hyperparameters. The hyperparameters $\alpha_0, \beta_0, \eta_0, \mu_0, \gamma_0$ are the hyperparameters for the priors over reliabilities, biases, and true scores, respectively.

Hence, the conditional distribution of the true score $y_j$, $\psi(y_j|\tilde{y}_j^{G(j)}; b_{G(j)}, \tau_{G(j)}, \mu, \gamma)$ is calculated as follows.

$$\begin{aligned}
\psi(y_j|\tilde{y}_j^{G(j)}; b_{G(j)}, \tau_{G(j)}, \mu, \gamma) &= \frac{\psi(y_j; \mu, \gamma)\psi(\tilde{y}_j^{G(j)}|y_j; b_{G(j)}, \tau_{G(j)})}{\int_{y_j} \psi(y_j; \mu, \gamma)\psi(\tilde{y}_j^{G(j)}|y_j; b_{G(j)}, \tau_{G(j)})dy_j}\\
&\propto \psi(y_j; \mu, \gamma)\psi(\tilde{y}_j^{G(j)}|y_j; b_{G(j)}, \tau_{G(j)})\\
&\propto \psi(y_j; \mu, \gamma) \prod_{i \in G(j)} \psi(\tilde{y}_j^{(i)}|y_j; b_i, \tau_i)\\
&\propto \exp\Big( -\frac{1}{2}\gamma(y_j - \mu)^2 + \sum_{i \in G(j)} \Big( -\frac{1}{2}\tau_i(\tilde{y}_j^{(i)} - (y_j + b_i)\Big)^2 \Big)\\
&\propto \exp\Big( -\frac{1}{2}\Big[\gamma(y_j - \mu)^2 + \sum_{i \in G(j)} \tau_i\big(\tilde{y}_j^{(i)} - (y_j + b_i)\big)^2\Big] \Big)
\end{aligned}$$

The expression inside the exponent is quadratic. We consider the exponent as follows.

$$\begin{aligned}
&\gamma(y_j - \mu)^2 + \sum_{i \in G(j)} \tau_i\big(\tilde{y}_j^{(i)} - (y_j + b_i)\big)^2\\
&\quad = const. + \gamma(y_j^2 - 2y_j\mu) + \sum_{i \in G(j)} \tau_i\Big((y_j + b_i)^2 - 2\tilde{y}_j^{(i)}(y_j + b_i)\Big)\\
&\quad = const. + \Big(\gamma + \sum_{i \in G(j)} \tau_i\Big)y_j^2 - 2\Big(\gamma\mu + \sum_{i \in G(j)} \tau_i(\tilde{y}_j^{(i)} - b_i)\Big)y_j,\\
&\quad = const. + R\Bigg(y_j - \frac{1}{R}\Bigg(\gamma\mu + \sum_{i \in G(j)} \tau_i(\tilde{y}_j^{(i)} - b_i)\Bigg)\Bigg)^2\\
&\quad (\text{where, } R = \gamma + \sum_{i \in G(j)} \tau_i)
\end{aligned}$$

Therefore the resultant distribution is Gaussian:

$$\psi(y_j|\tilde{y}_j^{G(j)}; b_{G(j)}, \tau_{G(j)}, \mu, \gamma) \sim \mathcal{N}\left(\frac{\gamma\mu + \sum_{i \in G(j)} \tau_i(\tilde{y}_j^{(i)} - b_i)}{\gamma + \sum_{i \in G(j)} \tau_i}, \frac{1}{\gamma + \sum_{i \in G(j)} \tau_i}\right)$$

$$\mathbb{E}_{y_j|\tilde{y}_j^{G(j)}; b_{G(j)}, \tau_{G(j)}}[y_j] = \frac{\gamma\mu + \sum_{i \in G(j)} \tau_i(\tilde{y}_j^{(i)} - b_i)}{\gamma + \sum_{i \in G(j)} \tau_i} \tag{28}$$

Now we are in a position to calculate $r_j^{\text{ERM}}(\tilde{\mathbf{y}}_j^{G(j)}; \boldsymbol{\theta}_{G(j)})$. The reward function is $R(x_j, y_j) = -(x_j - y_j)^2$ where $x_j$ is the estimated score and $y_j$ is the true score for paper $j$. The score-assignment rule *expected reward maximizer* (ERM) is given below.

$$\begin{aligned} r_j^{\text{ERM}}(\tilde{\mathbf{y}}_j^{G(j)}; \boldsymbol{\theta}_{G(j)}) &= \operatorname*{argmax}_{x_j \in S} \int_{y_j} \psi(y_j | \tilde{y}_j^{G(j)}; b_{G(j)}, \tau_{G(j)}) R(x_j, y_j) dy_j \\ &\text{where } b_i, \tau_i \text{ are the estimated bias and reliabilities } \forall i \in G(j) \\ &= \operatorname*{argmax}_{x_j \in S} \Big[ - \int_{y_j} \psi(y_j | \tilde{y}_j^{G(j)}; b_{G(j)}, \tau_{G(j)}) (x_j - y_j)^2 dy_j \Big] \\ &= \operatorname*{argmin}_{x_j \in S} \int_{y_j} \psi(y_j | \tilde{y}_j^{G(j)}; b_{G(j)}, \tau_{G(j)}) (x_j - y_j)^2 dy_j \end{aligned}$$

Let $g_j(x_j) = \int_{y_j} \psi(y_j | \tilde{y}_j^{G(j)}; b_{G(j)}, \tau_{G(j)}) (x_j - y_j)^2 dy_j$. Hence we need to find $x_j$ that minimizes $g_j(x_j)$. The first and second order conditions are given as follows.

$$\begin{aligned} &\frac{\partial g_j(x_j)}{\partial x_j} = \frac{\partial}{\partial x_j} \left[ \int_{y_j} \psi(y_j | \tilde{y}_j^{G(j)}; b_{G(j)}, \tau_{G(j)}) (x_j - y_j)^2 dy_j \right] \\ &= \int_{y_j} \frac{\partial}{\partial x_j} \left[ \psi(y_j | \tilde{y}_j^{G(j)}; b_{G(j)}, \tau_{G(j)}) (x_j - y_j)^2 dy_j \right] \\ &= 2 \int_{y_j} \psi(y_j | \tilde{y}_j^{G(j)}; b_{G(j)}, \tau_{G(j)}) (x_j - y_j) dy_j \\ &= 2x_j \int_{y_j} \psi(y_j | \tilde{y}_j^{G(j)}; b_{G(j)}, \tau_{G(j)}) dy_j - 2 \int_{y_j} y_j \psi(y_j | \tilde{y}_j^{G(j)}; b_{G(j)}, \tau_{G(j)}) dy_j \\ &= 2x_j - 2\mathbb{E}_{y_j | \tilde{y}_j^{G(j)}; b_{G(j)}, \tau_{G(j)}} y_j \\ &\frac{\partial g_j(x_j)}{\partial x_j} = 0 \Leftrightarrow x_j = \mathbb{E}_{y_j | \tilde{y}_j^{G(j)}; b_{G(j)}, \tau_{G(j)}} y_j \\ &\frac{\partial^2 g_j(x_j)}{\partial x_j^2} = 2 > 0 \end{aligned}$$

The first and second order conditions show that $x_j = \mathbb{E}_{y_j | \tilde{y}_j^{G(j)}; b_{G(j)}, \tau_{G(j)}} y_j$ is a global minima. Hence

$$r_j^{\text{ERM}}(\tilde{\mathbf{y}}_j^{G(j)}; \boldsymbol{\theta}_{G(j)}) = \mathbb{E}_{y_j | \tilde{y}_j^{G(j)}; b_{G(j)}, \tau_{G(j)}} y_j = \frac{\gamma\mu + \sum_{i \in G(j)} \tau_i (\tilde{y}_j^{(i)} - b_i)}{\gamma + \sum_{i \in G(j)} \tau_i}. \tag{29}$$

The last equality follows from Equation (28).

Replacing $\boldsymbol{\theta}$ with the estimated parameters, i.e., $\hat{\boldsymbol{\theta}}$, we get,

$$r_j^{\text{ERM}}(\tilde{\mathbf{y}}_j^{G(j)}; \hat{\boldsymbol{\theta}}_{G(j)}) = \mathbb{E}_{y_j | \tilde{y}_j^{G(j)}; \hat{b}_{G(j)}, \hat{\tau}_{G(j)}} y_j = \frac{\gamma\mu + \sum_{i \in G(j)} \hat{\tau}_i (\tilde{y}_j^{(i)} - \hat{b}_i)}{\gamma + \sum_{i \in G(j)} \hat{\tau}_i}. \tag{30}$$

## Appendix D. Comparison of Mean Squared Error under ERM (squared error minimizer) and ISWDM (PEQA)

Expected reward maximizer (ERM) minimizes the squared error by definition (Equation (6)). Thus, $W_j^{\text{ERM}} - W_j^{\text{ISWDM}} \geqslant 0$, where the reward function $R$ in the definition of $W_j$ (Equation (4)) is the negative of the squared error. How worse does ISWDM (PEQA) perform w.r.t. accuracy? To understand that, we run a simulation.

We consider a paper graded by 5 peer-graders– we call this *a set of grader-reports*. The graders are *symmetric*, i.e., have the same bias and reliabilities. To abstract away from the estimation process that is identical across ERM and ISWDM, we assume that the estimated bias and reliability are equal to the true values.

The simulations are run w.r.t. Piech et al. (2013). We generate 100 i.i.d. true scores with the parameters $\mu = 1, \gamma = 16$. For each generated true score, for a fixed bias and reliabilty $(\bar{b}, \bar{\tau})$, we generate 100 i.i.d sets of grader-reports, each set having the report of 5 graders with $(\bar{b}, \bar{\tau})$. We calculate the fractional difference $d = (W_j^{\text{ERM}} - W_j^{\text{ISWDM}})/|W_j^{\text{ERM}}|$ for each of the $100^2$ observations with $(\bar{b}, \bar{\tau})$.

To study the effect of bias, we vary the bias between 0 and 1 in steps of 0.1, keeping reliability fixed at 10.5. In Figure 3, for each bias $b \in \{0, 0.1, .., 1\}$ on the x-axis, we plot the mean and standard error of the observed $d$s under $(\bar{b} = b, \bar{\tau}{=}10.5)$. Similarly, to study the effect of reliability, for each $\tau \in \{6, 7, .., 15\}$ on the x-axis, we plot the mean and standard error of the observed $d$s under $(\bar{b} = 0.5, \bar{\tau} = \tau)$.

It shows that the average sub-optimality is small. It is insensitive to bias (roughly 25%) and monotonically decreasing in reliability.

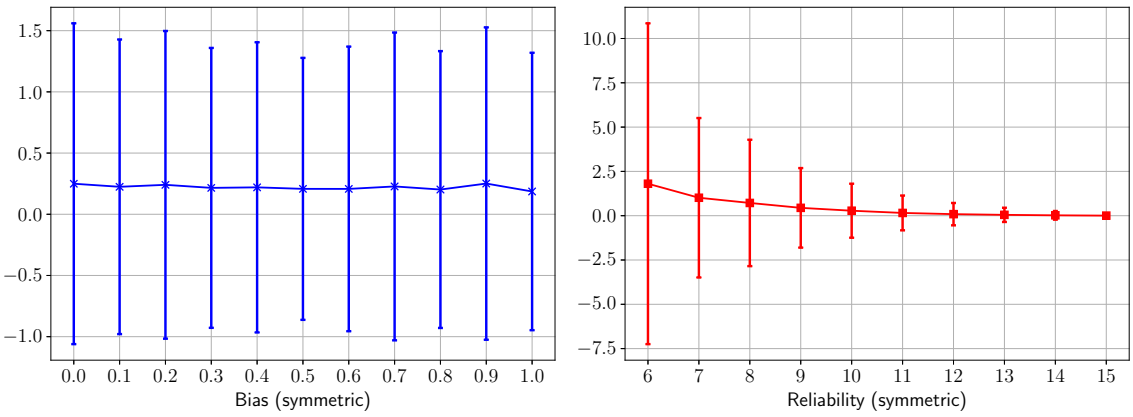



Figure 3: Normalized sub-optimality $((W_j^{\text{ERM}} - W_j^{\text{ISWDM}})/|W_j^{\text{ERM}}|)$ with increasing bias and reliability.

## Appendix E. Instructions provided to the human subjects (Section 7)

The instructions for both the mechanisms were as follows.

### E.1 Median Mechanism Instructions

First, please register yourself on: [registration link] and solve the problem[29] therein. The example there would help you understand how your decisions map into your final payments, through the median-mechanism and payment system used in this study. This is a study on peer-grading. You should read the following instructions carefully, as they would help you perform successfully in the study. In this study, each of you will be asked to grade the assignments of five anonymous students in this room. Similarly, your own assignment would be graded by five anonymous students from this room. Your peer-graded marks and the relative ranks in this peer-grading exercise only determine your payment from this session. It will not be used to determine your actual score for your final grade in the course. The assignment score used towards your university grades will be provided to you by the instructor (i.e., tutors or myself) later.

**Would I know whose exam papers I might be grading / correcting?** You would not have this information. We will take maximal precautions to make sure that the grader or the assignment-owner's identities are anonymous to each other during and after this session. Further you would also not know which other four participants are grading the same papers as you. Thus, this procedure is double-blind. We will provide you a solution manual to help you in the grading process. Follow the explanation of the questions and correct answers presented before the study. Please be respectful and encouraging in the grading process. Scores should reflect the learner's understanding of the assignment and points should not be deducted for difficulties with language or differences in opinion or for using a different but correct methodology.

How are the final grades on my own assignment decided? All five peer-graders independently assign you grades on all of the questions (there are 5 in total, all worth 2 points). Then for each question-part your final grade is the median of those five grades. For example, if on the second round of peer-grading, the five graders assign you 0, 1, 1.5, 2, and 2 respectively, then your final assigned grade on that question would be 1.5. We would calculate your grades on all the questions separately by the above median-method, and then aggregate those median grades from all the questions. For example, if there are five questions and the median grades on the questions are 0, 1, 1.5, 2 and 2 respectively, then the total grade on the assignment is 6.5.

**How does one calculate the median of five numbers?** Sort the numbers in increasing order and the third highest number would be the median.

**Can I dispute my peer-assigned grades?** Yes, for certain questions you can, and for others you cannot. In case you think your true grade is different than the grade that has been assigned to you on these questions, you can privately indicate that on a form, that would be sent at the end of the peer-grading and that will immediately notify us. We would then reassign you the grade the Teaching staff had assigned to your assignment previously. This whole process would be completed in a click of a button and you would be shown your updated grade in a matter of seconds. Please note that once a dispute is lodged, your grade would become the Teaching Staff assigned grade irrespective of whether that results in an increase or decrease over your original grade.

[29] The problem tests the participant's understanding of median and similar simple techniques.

**How are my payments decided?** Every participant would get a show up fee of M 50 for participating in and completing this session. You would also get an additional amount depending on your ranking in the pool of 'n' participants today. The ranking would be done in decreasing order of the final grades assigned to you all on the whole assignment. A ranking of x means that there are (x-1) other people who have a strictly higher grade than you. The additional amount would be equal to M 650 for the top 25% (first quartile) ranked students, M 450 for the next 25% (second quartile) ranked students, M 250 for the third quartile ranked students, and M 50 for the bottom quartile students. If the number of students that scored the same overlaps to two or more different quartiles, then all of them get the average payment of those quartiles. E.g., suppose 7 students got the same marks, and 3 students are in first quartile while 4 are in second quartile – then all 7 get M 600 (average of M 700 and M 500). Hence, in this study, the higher you are in the ranking based on your peers' judgment (and a potential review), higher is your total payment.

**How do the grades you submit affect your own payment?** The grades you submit obviously do not affect your own grade, because you are never grading your own paper, but they can still affect your own payment. Your grading would potentially affect the grades of others, and that can change the relative rank between you and the person(s) you are grading. For example, when you assign someone a higher grade, that might change the median grade they are assigned, and thus move them to a higher rank than you. Similarly, when you give them a lower grade, it might move them to a relatively lower rank than you. Obviously, both of these scenarios would affect the final payments of both you and the other person, as everyone is paid according to the final rankings.

Time-line for the study in chronological order:

- ▷ Stage 0: The whole assignment to be graded is broken up into 3 small parts, that would be peer-graded in three stages. The total grade from the whole assignment determines your final ranking and payment. At this stage, you are expected to complete the questionnaire successfully.

- ▷ Stage 1: Every one of you peer-grades the first part of the assignment of 5 of your peers. Therefore, for any question you are grading in this stage, you know that 4 other anonymous participants are also grading that question. Also, the first part of your own assignment is also being peer-graded by 5 other participants. One part of these questions will have options for regrading, while the other part will not (it will be mentioned in the response sheet, but all regrading requests will be collected at the end of stage 3).

- ▷ Feedback Stage 1: For each paper you graded in Stage 1, we will show you the grades assigned by you and the 4 other anonymous graders. We will also show you how part 1 of your own assignment got graded by the assigned graders.

- ▷ Stage 2: Similar to Stage 1, now part 2 of the assignment gets peer-graded. But the papers are now sent to a new random set of peer-graders. One part of these questions will have options for regrading, while the other part will not (it will be mentioned in the response sheet, but all regrading requests will be collected at the end of stage 3).

- ▷ Feedback Stage 2: Feedback of Stage 2 (similar to Stage 1) observed.

▷ Stage 3: Similar to Stage 2 (one part has regrading requests, the other does not), now part 3 of the assignment gets peer-graded.

▷ Feedback Stage 3: Feedback of Stage 3 (similar to Stages 1 and 2) is sent to all students, along with their tentative total score. You may raise regrading requests for the part that is regradable (as mentioned above). Any regrading requests that are lodged will be acted on. Performance on the whole assignment is aggregated, and the final ranking and payments are sent via email. To finish the study, complete the survey that comes in the last email. Study ends.

**Is my data confidential?** Yes, your data is completely confidential. Before observing and analyzing the collected data, we would be removing every personal identifier from the data, so that none of the decisions can be traced back to the individual who made the decision.

The first practice example tests you on your understanding of the mechanism how the peer-grading leads to your final grade, rank, and payment. You must complete this practice example with a score of 80% or more (i.e., correctly answer at least 4 questions out of 5). You will get one chance only, so please do this carefully. Failing this, you would be asked to leave this session with a M 20 reward.

**Important: Please do not communicate with any other participants during this session. For the grading, open one file at a time, finish grading, submit the grade in the google form and then move on. Please keep seated even if you are done with grading before time. If you have any questions, please raise your hand and one of us will come by to answer your query. Please use your university domain email id throughout this session. Please come remembering your google id/password, since that may be needed for some form filling.**

### E.2 PEQA Instructions

Before you begin, please register yourself on: [registration link]. Submit the form only once.

This is a study on peer-grading. In this study, each of you will be asked to grade **five** anonymous assignments. Similarly, your own assignment would be graded by a certain number of anonymous students from this room. Your peer-graded marks and your performance in the peer-grading exercise will only determine your payment from this session. It will not be used to determine your actual score for your final grade in the course. The assignment score used towards your university grades will be provided to you by the instructor (i.e., tutors or myself) later.

**Would I know whose exam papers I might be grading / correcting?** You would not have this information. We will take maximal precautions to make sure that the grader or the assignment-owner's identities are anonymous to each other during and after this session. Further you would also not know which other four participants are grading the same papers as you. Thus, this procedure is double-blind. We will provide you a solution manual to help you in the grading process. Follow the explanation of the questions and correct answers presented before the study. Please be respectful and encouraging in the grading process. Scores should reflect the learner's understanding of the assignment and

points should not be deducted for difficulties with language or differences in opinion or for using a different but correct methodology.

**How are the final grades on my own assignment decided?** Your peer-graders independently assign you grades on all of the questions. Then for each question your final grade is decided by running it through a **new mechanism called PEQA (Peer Evaluation with Quality Assurance)**. This is a mechanism which is designed to remove the individual biases in grading, and selectively weight and reward graders by how precise they are (details to follow). We would calculate your grades on all the questions separately by the above method, and then aggregate those grades from all the questions. In each round, you will have some **regradable** and **non-regradable** questions. For the regradable part, you will earn the peer-given score computed through PEQA **and** an additional PEQA reward for grading. For the non-regradable part, you will only receive the peer-given score computed through PEQA, **but no additional reward for grading**.

**Can I dispute my peer-assigned grades?** Yes, for certain questions you can, and for others you cannot. In case you think your true grade is different than the grade that has been assigned to you on these questions, you can privately indicate that on a form, that would be sent at the end of the peer-grading and that will immediately notify us. We would then reassign you the grade the Teaching staff had assigned to your assignment previously. This whole process would be completed in a click of a button and you would be shown your updated grade in a matter of seconds. Please note that once a dispute is lodged, your grade would become the Teaching Staff assigned grade irrespective of whether that results in an increase or decrease over your original grade.

**What is the PEQA mechanism?** Let us describe PEQA in short in the following two steps:
**Step 1 Probes:** Out of the five questions you (a grader) grade, two are randomly assigned to be *probes* (rest three are *non-probes*). On the probe papers, we would directly assign the teaching staff assigned grades and also use the teaching staff assigned grades to get an estimate of your individual *average deviation (or bias)* and *variance* in the assignments you graded. We will do this for all the graders. For a grader who on average, assigns a grade higher than the true-grade, the estimated deviation would be negative, and otherwise would be positive.
**Step 2 Non-Probes:** The non-probes would be graded using the information from, (i) the assigned grades of all the graders, and (ii) the estimated average deviation (or bias) and variance of grading by peer-graders in Step 1. The assigned scores would be "de-biased" using the information in 2.
Here is a numerical example that goes through these two steps. Suppose on the five questions you graded, the first two questions are randomly assigned as probes (this is for illustration only, the actual probes will be interspersed and not the first two, and you won't know which are the probes).

| Paper | Status | Score you assigned (A) | True Score (B) | Deviation (A-B) | Bias=Avg of Deviation |
|---|---|---|---|---|---|
| 1 | *Probe* | 3 | 3 | 3-3=0 | $\frac{0+(-.5)}{2} = -.25$ |
| 2 | *Probe* | 2.5 | 2 | 2-2.5=-.5 | |
| 3 | | 3.5 | | | |
| 4 | | 4 | | | |
| 5 | | 2 | | | |

On the probe questions, your evaluation would be compared with the evaluation done by the course instructors (True score), to calculate an average deviation in your grading. We would then use this to calculate the variance of your deviation.

| Paper | | Score you assigned | True Score | Deviation | Bias=Avg Deviation | Variance of Deviation |
|---|---|---|---|---|---|---|
| 1 | *Probe* | 3 | 3 | 3-3=0 | $-.25$ | $\frac{(0+.25)^2+(-.5+.25)^2}{2} = .0625$ |
| 2 | *Probe* | 2.5 | 2 | 2-2.5=-.5 | | |
| 3 | | 3.5 | | | | |
| 4 | | 4 | | | | |
| 5 | | 2 | | | | |

Suppose the (bias,variance) pairs of the other two graders, who are also grading question 4, are (.25, .05) and (-.5,.2) respectively. Suppose the scores they had assigned to the same Q4 was 3 and 2 respectively, while you have given 4 to that question.

Then, the final grade on Q4 (a typical non-probe question) would be calculated as ($k_1$ and $k_2$ are some appropriately chosen constants)

$$\text{assigned\_score} = \frac{k_1 + \frac{1}{\sqrt{.0625}}(4 + (-.25)) + \frac{1}{\sqrt{.05}}(3 + .25) + \frac{1}{\sqrt{.2}}(2 + .5)}{k_2 + \frac{1}{\sqrt{.0625}} + \frac{1}{\sqrt{.05}} + \frac{1}{\sqrt{.2}}}$$

When we assign the final grade on any non-probe question, we will "**de-bias**" the reports from all the graders by subtracting out the bias, and also selectively over-weight the information from the low-variance graders. We consider the inverse of the square-root of your variance as your **precision of grading**, and use this precision to weight your assigned score on this paper. The accuracy of the mechanism_assigned_score is given by $-(\text{assigned\_score-true\_score})^2$.

If you were not one of the graders, and the mechanism only assigned scores using the reports of the other graders,

$$\text{assigned\_score\_without\_you} = \frac{k_1 + \frac{1}{\sqrt{.05}}(3 + .25) + \frac{1}{\sqrt{.2}}(2 + .5)}{k_2 + \frac{1}{\sqrt{.05}} + \frac{1}{\sqrt{.2}}}$$

The new accuracy is $-(\text{assigned\_score\_without\_you-true\_score})^2$. Now, your `PEQA` performance score from peer-grading question 4 would be calculated as the difference between the accuracy with you, and the accuracy without you. This is intuitively equivalent to you

getting paid for your relative contribution in your group towards making the final assigned grade accurate. **The more accurate the assigned score is, when you are included in the group of graders, the higher would be your performance score!**

The PEQA performance score on each question you have graded that is worth $x$ points, is assigned on the scale of $[0, \frac{x}{2}]$. So, in round 1, where each regradable question is worth one point, and you grade a total of 3 non-probe questions, the maximum PEQA performance score you could get is $3 \times 0.5 = 1.5$ and the minimum is 0.

This PEQA grade and performance scores have the following properties:

**Bias Invariance:** Suppose you had reported grades of 3+x, 2.5+x, 3.5+x, 4+x, and 2+x, on all the questions instead, and thus had an individual deviations x points higher than before. This would have no effect on the PEQA performance scores, as it would be de-biased as described above. This is a mathematical property of the mechanism described.

With the new reported grades, your average deviation is changed to $-.25-x$ from $-.25$. The $+x$ and $-x$ cancel out in the expression of the assigned score, leaving it unchanged.

$$\text{assigned\_score} = \frac{k_1 + \frac{1}{\sqrt{.0625}}(4 + \cancel{x} + (-.25 - \cancel{x})) + \frac{1}{\sqrt{.05}}(3 + .25) + \frac{1}{\sqrt{.2}}(2 + .5)}{k_2 + \frac{1}{\sqrt{.0625}} + \frac{1}{\sqrt{.05}} + \frac{1}{\sqrt{.2}}}$$

**Clearly the assigned_score_without_you also cannot change if your bias changes, so your expected PEQA performance score cannot change here!**

**Precision Monotonicity:** For every set of (bias, variance) your co-graders might have, your expected PEQA performance score from the peer-grading task is monotonically increasing in your grading-precision (*precision is the inverse square-root of your variance*). This is a mathematical property that can be easily showed by using calculus and statistics. Thus the more precisely you evaluate a paper in the peer-grading task, (or alternatively the lower your grading variance) the higher your peer-grading score.

Here is a graph that shows how the PEQA performance score changes with the Precision for a grader, who is grading alongside with two graders, one of highest precision and one of lowest precision.

**How do you calculate $-$(assigned_score-true_score)$^2$?** If there is no regrading request, then we would assume that true_score=assigned_score, and this value is zero. If there are regrading requests, then we would evaluate the paper ourselves and assign the course-instructor assigned score as true_score to calculate the value.

**What is my consolidated score?** Your consolidated score is the sum of (i) the score on your own assignment (consolidated score from the regradable and non-regradable parts), and (ii) your PEQA performance score (peer-grading score). For example, if the peer-assigned score (computed via PEQA) on your own assignment is x, and your peer-grading score is y, your consolidated score is x+y.

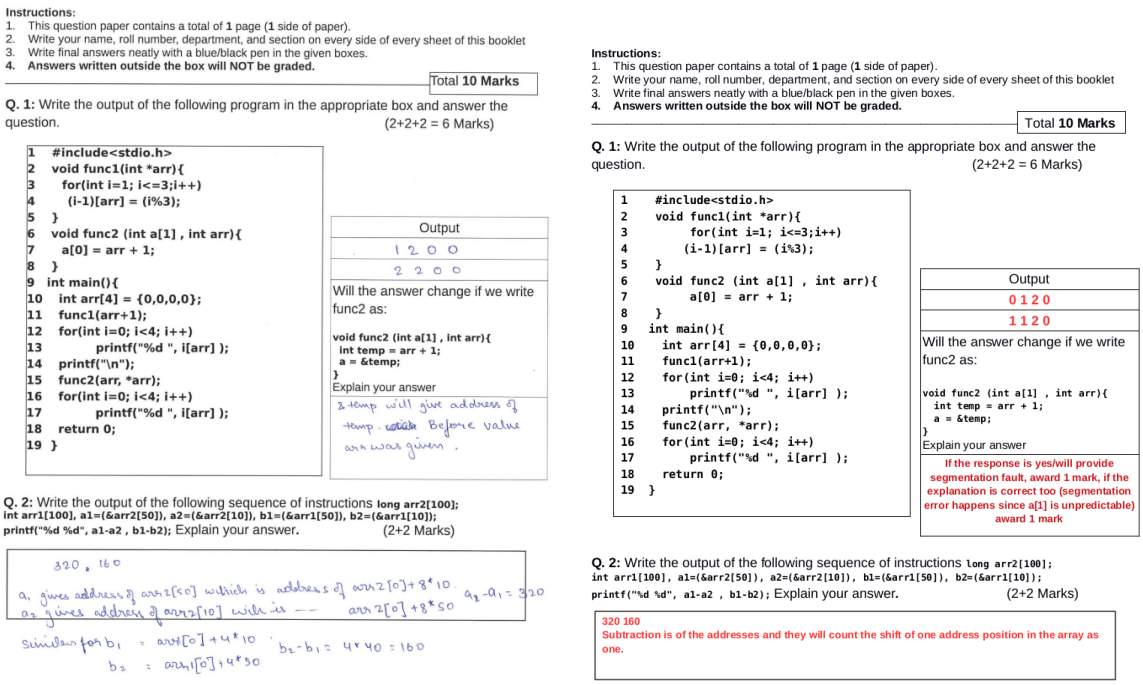
**How are my payments decided?** Every participant would get a show-up fee of M 50 for participating in and completing this session. You would also get an additional amount depending on your ranking in the pool of 'n' participants today, based on the consolidated score. **The ranking would be done in decreasing order of the final grades (i.e.,**

**the consolidated score) assigned to you all on the whole assignment**. A ranking of x means that there are (x-1) other people who have a strictly higher consolidated score than you. The additional amount would be equal to M 650 for the top 25% (first quartile) ranked students, M 450 for the next 25% (second quartile) ranked students, M 250 for the third quartile ranked students, and M 50 for the bottom quartile students. If the number of students that scored the same overlaps to two or more different quartiles, then all of them get the average payment of those quartiles. For example, suppose 13 students out of a population of 40 got 10/10, then all 13 get M $(700 \times 10 + 500 \times 3)/13 = 654$ – the next rank starts from 14. Hence, in this study, the higher is your consolidated score, higher is your total payment.

**How do the grades you submit affect your own payment?** The grades you submit obviously do not affect your own grade, because you are never grading your own paper, but they can still affect your own payment, in two ways.
1) **By affecting the grade of others**: Your grading could potentially affect the grades of others, *only if the question is chosen as non-probe question*, and consequently that can change the relative rank between you and the person(s) you are grading. For example, when you assign someone a higher/ lower grade on a question that is chosen as a non-probe question, that might change the `PEQA`assigned quiz score (and thus the consolidated score) they are assigned, and thus affect the relative rankings. But, note that Bias Invariance result described above already tells you that a **different bias would not change the expected quiz scores** of any peers.
2) **By affecting your peer-grading score**: Assigning a higher/ lower score on any question, could change your payments in two ways. If this happened on a question that was chosen as probe, we would be calculating your precision and bias to a different number, and **a lower (respectively higher) precision would result in a lower (respectively higher) marginal impact of your peer-grading reports, and hence, a lower (respectively higher) peer-grading score (and hence lower consolidated score)** for you. If this was a non-probe question instead, then you might be able to change the peer-graded score on that paper, depending on how much weight we assign to your evaluation.

**Is my data confidential?** Yes, your data is completely confidential. Before observing and analyzing the collected data, we would be removing every personal identifier from the data, so that none of the decisions can be traced back to the individual who made the decision.

You would be given a questionnaire of three questions that tests you on your knowledge of calculation of median. Failure in answering at least two correctly out of those three questions would disqualify you from participation in this study. In this case you would be asked to leave this session with a M 20 reward. Important: Please do not communicate with any other participants during this session. For the grading, open one file at a time, finish grading, submit the grade in the google form and then move on. Please keep seated even if you are done with grading before time. If you have any questions, please raise your hand and one of us will come by to answer your query. Please use your university domain email id throughout this session. Please come remembering your google id/password, since that may be needed for some form filling.




(a) Sample paper (b) Answer key

Figure 4: A typical view of the peer-grader.

## Appendix F. View of a typical paper by the peer-graders

Figure 4 shows a typical view of the peer-graders for a specific paper. The paper was chosen such that it has both subjective and objective components. During the experiments, a few emails are sent to the individual peer-graders. The first email provides the set of papers to be graded by them along with a sketch of solutions (as shown in the figure). The second email is sent after the peer-grades are available and asks if the student wants to place a regrading request. The final email gives the final scores after regrading the paper (if requested) and their final bonus scores (monetary payments in the experiment).

All data and code of this paper are available at: https://www.cse.iitb.ac.in/~swaprava/papers/Codes_Peer_Grading.zip