Education and Information Technologies
https://doi.org/10.1007/s10639-022-11175-5

A dynamic review allocation approach for peer assessment in technology enhanced learning

Gabriel Badea[1] · **Elvira Popescu**[1]

Received: 8 November 2021 / Accepted: 16 June 2022

Abstract

Peer assessment plays an important part in education, fostering involvement and critical thinking skills for the students, while reducing the grading workload for the teacher. However, research on review allocation mechanisms in the context of peer assessment is relatively scarce in the literature. Although the first electronic peer assessment systems emerged over two decades ago, the matching of the solutions to reviewers has been done predominantly statically and randomly. The current paper proposes an innovative dynamic review allocation mechanism with extra bidding that attempts to solve some of the issues exhibited by the static and random approaches. The new method splits the review period in two stages: a first review phase, where students have to submit required assessments, and an extra review phase, where students can offer bidding points to perform additional optional reviews. The mechanism was integrated as part of our LearnEval peer assessment platform. We employed the approach in the context of a computer science course. A comparison of the proposed mechanism with a static review allocation approach applied in the previous year shows increased fairness. Furthermore, the results from a dedicated simulation module for emulating the novel review allocation mechanism attest the scalability of the approach.

Keywords Dynamic review allocation · Extra review phase · Peer assessment · Review allocation mechanism · Simulation module

✉ Gabriel Badea
gabriel.badea@edu.ucv.ro

Elvira Popescu
elvira.popescu@edu.ucv.ro

[1] University of Craiova, Craiova, Romania

Published online: 27 June 2022

1 Introduction

Peer review, also known as peer assessment, represents the process through which learners assess the quality of the work submitted by peers with the same level of knowledge (Topping, 1998). The field is a well-known research topic that has gained the attention of many academics in the recent decades (Estévez-Ayres et al., 2013; Giannoukos et al., 2010; Verleger et al., 2010). Initially, the activity was applied mainly in physical classrooms, where a relatively low number of students were enrolled. Nowadays, it is applied in classes of any size. Furthermore, it has become increasingly common to be employed in MOOCs (Massive Open Online Courses) with many learners registered (Gamage et al., 2021). On the other hand, the method constitutes a viable solution in grading complex and subjective assignments where it would be difficult to apply automatic machine grading.

The peer assessment activity offers many advantages for both learners and teachers, such as: encourages reflection and metacognition, fosters critical-thinking skills (Politz et al., 2014), allows peers to spend more time on the task and receive more and faster feedback (Mader & Bry, 2019), and improves the quality of learning for both the reviewer and the reviewee (Topping, 2009; Murakami et al., 2012). In addition, peer assessment can be used by the instructor in several ways to improve the educational process: to peer-moderate the grades assigned by the students in group projects where it is difficult to know the individual contribution of each member (Loddington et al., 2009; Sukstrienwong, 2017), to reduce the grading workload (Rubin & Turner, 2012), and to allow the instructor to focus more on the students that need their attention (Mader & Bry, 2019). A reliable peer assessment process can be achieved only by considering the various facets of the activity, such as: quality of the feedback provided by learners, accuracy of the assigned grades, fairness of the approach used for allocating the submitted solutions to reviewers, relevance of the review criteria, or learners' confidence in the process. In the literature many papers have focused on and have examined the various actors, mechanisms and facets of the process, however, relatively little attention has been dedicated to the mechanisms employed for allocating the submitted solutions to assessors[1] (Abrache et al., 2021; Crespo García et al., 2005; Gehringer & Cui, 2002; Staubitz et al., 2016; Verleger et al., 2010).

More than two decades ago, before the advent of electronic peer reviewing, the allotment of deliverables was done manually by the instructor in the classroom. However, the task was very time-consuming and cumbersome, even when few students were enrolled. Furthermore, it was even more intricate to personalize the activity, such as, make it anonymous or allocate the submissions based on particular traits of the peers (e.g., high quality submissions to novice reviewers). The activity was not scalable, thus, in classes with high enrollment levels it was not practical to apply it (Crespo et al., 2005). The advent of web platforms dedicated for peer assessment (Davies, 2000; Lin et al., 2001; Freeman & McKenzie, 2002) has

[1] Students' work submitted for peer assessment can be referred to as deliverable, submission or solution and these terms are used interchangeably in this paper.

resulted in the automation of a part of the common redundant tasks. Representative automatized tasks include: the delivery of solutions, provision of reviews, assignment of grades, or generation of reports and statistics. Furthermore, the allocation of the submissions to evaluators was no longer done manually, but instead it started to be performed by the system. The time allotted by the instructor for the activity was considerably reduced and the scalability of the procedure increased (Davies, 2000).

The reviewer assignment strategy plays a crucial role on the attitudes of reviewers, authors, review efficiency and quality (Wang & Sun, 2018). The strategy affects not only learners' satisfaction, but also the educational equality (Wang et al., 2018). However, for a long period, most of the systems have employed a simple, static assignment technique where each deliverable is allocated randomly to assessors (Abrache et al., 2021). Static refers to the approach in which the solutions are allocated only once, at the beginning of the reviewing phase. This implies several negative consequences on the fairness and equity of the assessment process. In case a student drops out of a course or does not provide the assigned reviews on time, a part of the solutions may be left with very few or even no evaluations at all. In practice, this approach results in some deliverables receiving many evaluations, whereas others very few or none. Furthermore, in case no constraints exist on the allocation, unfair and undesirable assignments can emerge, such as a novice reviewer to assess a low quality submission, and thus, no gain would be obtained by either of the actors.

To date a few papers have focused on finding and proposing new alternatives to the random, static allocation of the submissions (Abrache et al., 2021; Estévez-Ayres et al., 2013; Staubitz et al., 2016). The most practical, fair and unbiased alternative to the static assignment represents the *dynamic* allocation (Gehringer & Cui, 2002; Staubitz et al., 2016). In this approach a submission is not allocated to an assessor beforehand, but only in the moment when the student requests to review it, thus reducing the risk of not submitting the review on time.

In this paper we propose a novel review allocation mechanism that assigns dynamically, at student's request, solutions to assess, and attempts to counteract the flaws existing in the static approaches. Furthermore, the mechanism entails an extra review phase where peers can bid points to evaluate additional solutions. The paper addresses the following objectives: 1) introduce the new dynamic review allocation mechanism with extra bidding; 2) compare the proposed mechanism's allocation results with the ones from a static review allocation approach; 3) assess the outcomes of the proposed mechanism; 4) examine students' opinion regarding the review allocation; 5) showcase and run a highly configurable simulation module that emulates the behavior of the proposed mechanism. The new mechanism was developed and integrated in our LearnEval peer assessment platform (Badea & Popescu, 2019a, b, 2022).

The paper is structured as follows: Section 2 surveys related work regarding different classifications of review allocation mechanisms, as well as various complex approaches; Section 3 describes several static review allocation approaches employed by our peer assessment platform and introduces the new dynamic review allocation mechanism with extra bidding; Section 4 depicts a practical application of the proposed mechanism, compares the outcomes with the ones obtained from a static allocation, and examines students' perception regarding the new approach; Section 5 introduces a simulation module developed to assess the applicability and scalability of the proposed mechanism, and displays the results of an experiment ran with various configuration parameters values; Section 6 discusses the results obtained, draws some conclusions, and presents limitations and future work directions for improving the reliability of the mechanism and the peer assessment process.

2 Related work

In the following, we present the current state of the literature by surveying different classifications of review allocation mechanisms, followed by several complex approaches and the various features, advantages and challenges they exhibit.

2.1 Classification of review allocation mechanisms

Gehringer (2001) discusses, in an early review study which classifies the strategies employed for matching assessors to submissions, methods for increasing feedback quality and for preventing the clustering of grades offered by the students around the mean. Various review allocation approaches are identified, such as: randomly, pseudo-randomly, self-selection by the assessors, strategies employed when working in groups, strategies employed when developing projects, and strategies for team peer assessment (i.e., students assess the work of their team members so that the instructor can consider the relative contribution of each member and assign a grade according to it). The author suggests that a peer assessment software should support the majority of the discussed techniques as each review allocation mechanism is appropriate for a given setting.

Wang and Sun (2018) proposed a taxonomy of the reviewer assignment strategies describing stability, anonymity, review author ratio, and assigner aspects. In terms of stability, the strategy can be fixed or random. In the fixed approach, the same reviewers are assigned to the same authors the entire semester and the actors can collaborate and work together to finalize the task. On the other hand, in the random scenario, different reviewers are assigned in distinct peer assessment sessions allowing actors to learn more by viewing more diverse writing styles and review perspectives. In terms of anonymity, the matching can be double-blind, or single-blind, when the identity of one of the actors is disclosed to the other. In terms of reviewer author ratio, four possible ways exist: one evaluator to one author, multiple evaluators to one author, one evaluator to multiple authors, and multiple evaluators to multiple authors. The paper recommends one should consider the student's workload and process reliability when choosing the strategy. In terms of assigner, it can be the student, the teacher, or the system. The work concludes there is no optimal assignment strategy that fits all the conditions, and the strategy should be selected based on the context.

Four reviewer assignment approaches and the consequences and inequalities they exhibit are presented in Wang et al. (2018). The strategies are based on the review competence difference between the reviewer and the reviewee: {high, low} when highly skilled peers are assigned to low competence students; {low, high} when low competence peers are assigned to highly skilled students allowing evaluators to learn from the reviewed solutions; {high, high} when high competence students are assigned to highly skilled students, fostering knowledge sharing and learners' satisfaction; {low, low} which is a consequence of the {high, high} assignment as there are not enough high competence students remaining. In the last approach the low skilled students might feel discriminated. The outcomes show that inequality happens when there is little competence difference between the two peers assigned as a review pair, such as the patterns {high, high} and {low, low}.

Different ways have been used to classify the review allocation approaches considering aspects, such as: feedback quality (Gehringer, 2001), reviewers' competence, actors' anonymity, assigner, stability, reviewer author ratio (Wang and Sun, 2018), or review competence difference between the reviewer and the reviewee (Wang et al., 2018). However, as can be seen, most review allocation approaches presented in these taxonomies are simple, static, or random; there are also a few exceptions of more complex approaches, which we survey in the following subsection.

2.2 Complex review allocation mechanisms

Staubitz et al. (2016) and Estévez-Ayres et al. (2013) present two interesting dynamic review allocation mechanisms that allocate the submissions to peers on demand. In both works, in case the student does not submit the review on time, the system reallocates the submission to another peer. In Staubitz et al. (2016) each submission is assigned a priority to provide a fair distribution of the reviews and to avoid the case when multiple students review simultaneously the same deliverable. The priority is computed based on the remaining number of required reviews to be submitted for the solution. The deliverables with the highest priority, that still need evaluations, are allotted first. Furthermore, when a student provides an assessment, their submission priority is increased, fostering more active reviewers to get more evaluations than the students who evaluate less. However, the literature reports that skilled students usually write more reviews than the rest, thus this approach might result in competent students receiving more evaluations, at the harm of the weaker peers. Similarly, Estévez-Ayres et al. (2013) present two review allocation mechanisms whose main roles are to reduce students' frustration by decreasing the number of committed learners that do not receive assessments from their peers. The algorithms set a specific deadline for each review relative to the moment it was assigned to the student, instead of setting a global review deadline. The main novelty of the approaches is that peers can start the assessment process earlier, once they have submitted a solution. Once the student decides to move to the next phase, N submissions are assigned to them and a sliding deadline (similar for all students) is set to submit the assessments. The student can request additional reviews (only once) after providing the mandatory ones, but now the deadline will be the global review deadline. A comparison of the two algorithms with a baseline allocation approach, via simulation, showed that both algorithms substantially minimized the number of committed students that do not receive any reviews. However, the baseline method does not penalize the peers who do not provide the assessments.

Another approach is to model the learner based on their assessment skills and construct their profile before allocating the submissions, as in the works of Abrache et al. (2021), Crespo et al. (2005), or Giannoukos et al. (2010). Abrache et al. (2021) for example model the process of allocating the submissions to reviewers as a Many to Many Assignment Problem (MMAP). Two steps are required for modelling the allocation process: firstly, peer modeling as a reviewer based on descriptive features and calculation of their review competence, and secondly, creation of clusters of peers representing various levels of review competence. The review categories of the assessors are weighted when computing the grade to be assigned for a submission. The number of evaluations per submission was set to four to not overload the reviewers, thus, each student had to perform between one and six assessments depending on their review category. Crespo et al. (2005) also present an effective and generic mechanism for matching solutions to reviewers for tailoring the peer assessment process based on learners' needs and traits. The mechanism can be applied to team-based projects and it is not limited to educational settings. Each peer is depicted as a profile based on their characteristics and pairs are generated according to different pedagogical criteria specified by the instructor. The review allocation task is depicted as an optimization problem to find a solution that meets a set of criteria. The approach was trialed in a computer engineering course and the students' scores were used to construct the profiles. Each submission was assigned to three assessors: a reviewer with a profile complementary with the author, a highly skilled reviewer, and a reviewer with a profile similar to the author. The three assessments are weighted when computing the grade for a submission. Learners' opinions about the process, gathered by means of surveys and informal talks, were overall positive. By contrast, Giannoukos et al. (2010) employ feed forward neural networks to determine the best assessor for a given author. Machine learning techniques are used to construct reviewer profiles based on past review data and the *helpfulness* assigned by the authors to the reviewer comments. The algorithm makes the assessment process more efficient by determining the reviewer that would offer the most useful evaluation for an author. Two types of profiles are created for each student: a profile as reviewer, and a profile as author. Initially, the mapping between evaluators and authors is performed randomly. At the end of the first reviewing phase, the authors rate the usefulness of the comments received and the algorithm constructs a list of optimal reviewers for each author based on the perceived helpfulness the authors would probably receive by assigning them. The preliminary results are promising as the neural network achieved over 72% accuracy in predicting the three to five most optimal assessors for an author.

Additional approaches for review allocation are applied in Verleger et al. (2010) and Gamage et al. (2017). In Verleger et al. (2010) an algorithmic assignment method matched the highly skilled reviewers with teams requiring more support, and less skilled assessors with teams needing less support. The aim of the algorithm was to make the allocation more helpful for the reviewed students compared to a random assignment mechanism. At the first step, the approach assigns to each team a highly skilled reviewer. Next, the remaining assessors are evenly allocated between the teams, however, to the teams in need are assigned highly skilled reviewers and to the rest of the teams are assigned low skilled evaluators. The approach attempts to offer the best support for the teams in need without penalizing too much the rest. The outcomes show that the simple assignment mechanism still represents the easiest and fairest method for matching as although it does not guarantee every team will receive a highly skilled reviewer, it does not have a clear bias for some of the teams to not receive one. By contrast, Gamage et al. (2017) present a peer assessment framework where reviewers and solution authors know the identities of each other and can communicate directly. The authors state that disclosing identities could make peers to be more accountable and offer more consideration when reviewing, increasing the overall feedback quality. Similar with Giannoukos et al. (2010), the proposed framework allows authors to rate the feedback received based on *helpfulness* and at the beginning of the assessment process every peer was treated equally and the allocation was done randomly. The algorithm assigned students who provided high quality feedback to reviewers who also provided helpful feedback, stimulating peers to offer good reviews. However, in some cases a reviewer had to wait until an assessor with a corresponding feedback usefulness was found. The grades received were available only after the student submitted all the reviews to mitigate grade inflation. The student had a time frame to submit the review and rate the feedback received. The outcomes show that students communicate more and the feedback is more relevant when the peer assessment process is not double-blind.

The complex mechanisms reported above represent a powerful solution and alternative to the static and random allocation commonly applied in the peer assessment activity. However, the proposed mechanisms also have some limitations, such as: the assessment skills of the reviewers are not considered, thus only unskilled assessors might be assigned to some solutions (Staubitz et al., 2016; Estévez-Ayres et al., 2013); if some evaluations are not submitted, then some solutions might be left without any reviews (Abrache et al., 2021; Crespo et al., 2005); the number of reviews different students performed varied to a large extent (Abrache et al., 2021); a student offering low quality feedback may be assessed by a similar student (Gamage et al., 2017); no improvements are found compared to a random allocation approach (Verleger et al., 2010); important delays exist in some cases in the assignment process (Gamage et al., 2017). Furthermore, some papers present only preliminary experimental results (Giannoukos et al., 2010) or are not tested in real settings at all (Estévez-Ayres et al., 2013).

Moreover, very few papers focused on the student's motivation to perform reviews and even fewer focused on the student's incentive to submit extra reviews. The current paper proposes a novel dynamic review allocation mechanism with extra bidding that attempts to solve important issues pertaining to the current approaches, such as unfair allocation of the reviewers, low involvement, or unequal number of evaluations submitted for different solutions. The presented approach entails an extra review phase where students can provide additional evaluations. A new concept, bidding points, is introduced to stimulate peers to get more involved in the reviewing process. Firstly, the students gain bidding points by submitting reviews and back-reviews. Subsequently, these points can be used to request extra reviews. The additional assessments allow faster gathering of points, motivating students to review more.

3 LearnEval review allocation mechanisms

In this section we showcase the various review allocation mechanisms supported by our peer assessment platform, called LearnEval (Badea & Popescu, 2019a, b, 2022) and we report on the applicability and suitability of each allocation mechanism.

3.1 LearnEval system overview

LearnEval is a highly configurable peer assessment platform that allows instructors to easily tailor the assessment workflow based on the specific requirements and context of the course (Badea & Popescu, 2019b, 2022). The core modules that enhance the teacher experience include:

- *Assignments* - where the instructor can create the assignments for the course and set various properties, such as: specifications, submission or review deadline, review criteria
- *Calibration* - where the instructor can define calibration assignments for the students to practice their assessment skills
- *Settings* - where the instructor can configure various aspects of the assessment process, such as: anonymity of the reviewers, anonymity of the solution authors, define the weights of the metrics used for computing personal scores, or the mechanism employed for allocating the submissions to reviewers
- *Statistics* - where the instructor can visualize, inspect and analyze various aspects related to the peer assessment activity at different granularity levels
- *Scores* - where the instructor can view the personal scores of each student
- *Notifications* - where the instructor can view the notifications automatically sent by the system when various actions of interest occur.

In addition, the student area (Badea & Popescu, 2019a, 2022) also provides a wide range of modules, such as:

- *Assignments* - where the student can view information about the assignments and submit solutions
- *Calibration* - where the student can practice, test and improve their assessment skills before assessing actual peers' work
- *Review solutions* - where the student can review peers' submissions
- *My solutions* - where the student can access their submitted solutions
- *My reviews* - where the student can access their submitted reviews, but also the reviews submitted by other peers for the solutions they reviewed

- *Statistics* - where the student can visualize, inspect and analyze various aspects related to their peer assessment activity
- *Scores* - where the student can view their personal scores; furthermore, an open learner model visualization is available (Badea & Popescu, 2020a)
- *Notifications* - where the student can view the notifications automatically sent by the system when various actions of interest occur.

LearnEval has already been applied for several years at the University of Craiova, Romania, in the context of computer science courses and project-based learning (PBL) scenarios (Badea & Popescu, 2019c, 2020b, c, 2022). For the first two years, we employed a static review allocation mechanism that automatically assigns the student deliverables to assessors once the submission deadline is reached. However, given the multiple issues found during the static allocation, the platform was extended and a new dynamic review allocation mechanism was integrated.

Indeed, a powerful and robust peer assessment platform should support multiple review allocation approaches and let the instructor select the mechanism according to the specific requirements and the context of the course. LearnEval is very comprehensive and supports a wide range of review allocation mechanisms, as described next: *Manual by the student*, *Manual by the teacher*, *Random by the system*, *Automatic based on three categories of reviewers* (low, medium and high review skills), *Automatic based on N categories of reviewers* (according to the number of reviews per solution), and *Dynamic review allocation with extra bidding*.

3.2 Initial review allocation approaches

The static review allocation approaches assign the submissions to the evaluators only a single time, once the submission deadline is reached. In a typical static review allocation approach, a peer assessment session comprises the following phases:

1. The teacher creates the assignment and sets the submission and review deadline
2. The students submit solutions until the submission deadline is reached
3. The submissions are allocated to assessors based on the review allocation mechanism selected by the instructor
4. The students provide the assigned reviews until the review deadline is reached
5. Each submission is assigned a grade based on the evaluations received.

In the following we describe the static review allocation mechanisms supported by the initial version of the LearnEval platform.

3.2.1 *Manual by the student* approach

The instructor specifies the number of assessments each submission must receive and according to the value, the system computes the minimum and the maximum limit of reviews each student is required to perform for the current assignment:

$$minimumRequiredReviews = floor\left(\frac{solutionsSubmitted * reviewsPerSubmission}{studentsEnrolled}\right)$$

, where: *minimumRequiredReviews* represents the minimum number of evaluations the student should perform, *solutionsSubmitted* represents the number of solutions submitted for the current assignment, *reviewsPerSubmission* represents the number of assessments each submission must receive, and *studentsEnrolled* represents the number of students enrolled in the course.

$$maximumRequiredReviews = minimumRequiredReviews + 1$$

, where: *maximumRequiredReviews* represents the maximum number of evaluations the student could perform.

Note that only a part or even none of the students can perform *maximumRequiredReviews* evaluations to not exceed the required number of assessments per submission. The mechanism allows students to self-select the submissions they wish to review. By default, the available information about a submission includes the link where it can be downloaded and the review deadline. The instructor can further configure the solution's author name to be visible. However, this approach is not recommended as it can open ways for bias such as students selecting only submissions from a few preferred peers. In the pool of submissions to select are displayed only the submissions that have not been assigned to the required number of assessors yet.

3.2.2 *Manual by the teacher* approach

The instructor specifies the number of assessments each submission must receive. This mechanism allows the instructor to map the submissions to reviewers in an easy and intuitive manner (see Fig. 1). The page draws a row for each submission containing a link where it can be downloaded and a list of cells with possible reviewers. The peers that are currently assigned as reviewers for a given solution have a grey background cell color. The instructor clicks the student's name in the corresponding row to assign a particular student to a submission. In case the required number of reviewers assigned to a submission was met, the background color turns to green (see the first solution in Fig. 1), otherwise it is depicted in red. Several metrics are displayed for each reviewer to ease the instructor's work in mapping assessors: number of assigned solutions to review for the current assignment, total number of assigned solutions to review, and total number of submitted reviews for the current course. Furthermore, the metrics' values are displayed in various colors to indicate if some conditions are met (e.g., total number of submitted reviews is displayed in red if it is above the class average). The approach is not very scalable and it is recommended in classes with a relatively low number of students enrolled. Therefore, in courses with high enrollment rate and many assignments, the workload can be very difficult to handle by the instructor. Furthermore, the instructor must know the competence and assessment skills of the students for a fair assignment and to maximize the reliability of the process.

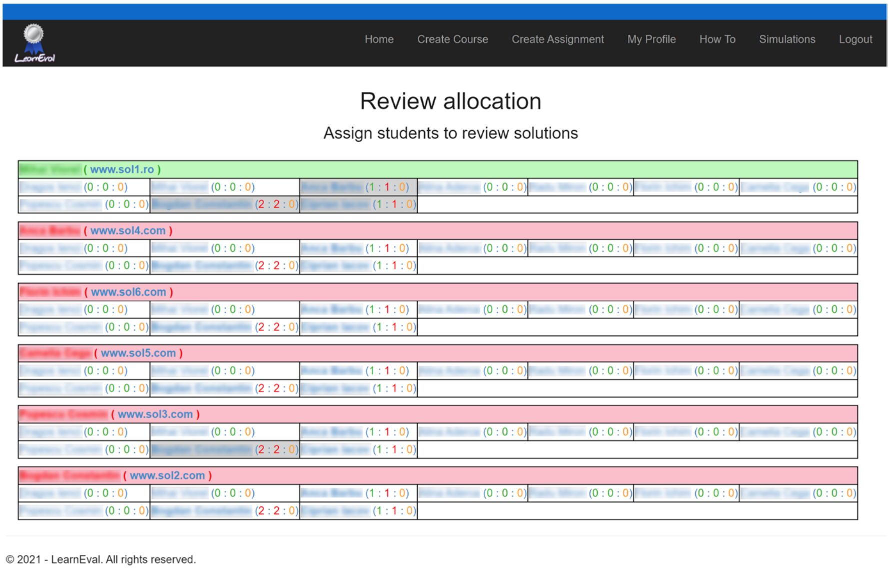

Fig. 1 LearnEval "Manual by the teacher" allocation mechanism - Assign reviewers to submissions page

3.2.3 *Random by the system* approach

In this case, the submissions are automatically allocated to the number of reviewers specified by the instructor once the submission deadline is reached. In the allocation process, priority is given to the peers with the lowest number of assessments assigned in the past and, in case several students have the same amount, priority is given to the ones that have submitted a solution for the current assignment. The allocation can be unfair as a student with low review competence could be assigned to a low quality submission, resulting in little gain for both the reviewer and the author of the solution. As reported in the literature, this type of allocation has been extensively applied in many peer assessment systems.

3.2.4 *Automatic based on three categories of reviewers* approach

Again, the submissions are automatically allocated to the number of reviewers specified by the instructor once the submission deadline is reached. However, the reviewers are split in three distinct categories according to their review competence: students with high reviewing skills (HRS), students with medium reviewing skills (MRS) and students with low reviewing skills (LRS).

In general, the number of evaluators to assign from a category is computed as follows:

$$reviewersToAssignFromCategory_{reviewCategory} = reviewersToAssign/3$$

, where: $reviewersToAssignFromCategory_{reviewCategory}$ represents the number of reviewers to assign from category *reviewCategory* and *reviewersToAssign* represents the total number of assessors necessary to assign to the solution.

Nonetheless, two distinctive cases need to be considered:

- case I: *reviewersToAssign* mod 3 = 1 ➔ an extra reviewer is assigned, i.e., the student with the lowest number of submissions assigned to review until the present from the review category with the lowest total number of submissions allocated until the present
- case II: *reviewersToAssign* mod 3 = 2 ➔ two extra reviewers are assigned, i.e., the two students with the lowest number of submissions assigned to review until the present from the review categories with the lowest total number of submissions allocated until the present.

In the following we detail the assignment process. The first step in the allocation process represents the categorization of the students based on their reviewing skills. LearnEval computes for each student a *Reviewing Score* according to their assessment skills. *Reviewing Score* depends on several metrics, such as: average of the back-reviews received from the peers, average of the back-reviews received from the teacher, agreement between the grades assigned by the student and the final grades assigned to the submissions reviewed, and a calibration score. A back-review represents an evaluation that a solution author can perform on a received review. The calibration score depicts student's assessment skills before the actual reviewing process starts. Next, the students are ordered descending based on their *Reviewing Score* and are allocated to the review categories: the first third are allotted to HRS, the second third are allotted to MRS, and the remaining ones are allotted to LRS.

Once the submission deadline is reached, the next step represents the allocation of the submitted solutions to reviewers. LearnEval allows the instructor to configure the number of reviewers per submission. In the context of our studies, the solutions were allotted to three evaluators, one from each review category, to achieve a fair process. The algorithm selects from each category the student with the lowest number of assigned reviews until the present, thus, in the end, each peer has relatively the same number of assigned solutions to review. Hence, the maximum difference between the highest and lowest number of assigned reviews a peer can have is one. However, in practice this allocation approach poses several problems that make it unfair for some of the participants. In an ideal unfolding, where all or most of the peers submit the assigned reviews, each solution receives the required evaluations, and implicitly, the authors receive sufficient feedback regarding their work. However, in real settings, the actual involvement of the reviewers varies, and thus, some of the submissions receive all the three required reviews, while others receive only one or even no reviews.

We employed this initial allocation mechanism in 2018-2019 and 2019-2020 academic years in the context of several computer science courses (Badea & Popescu, 2019c, 2020b, c, 2022). The issues related to the static allocation were especially visible in the Web Applications Design course (Badea & Popescu, 2020b) where the peer assessment activity was optional and in the last assignments very few students

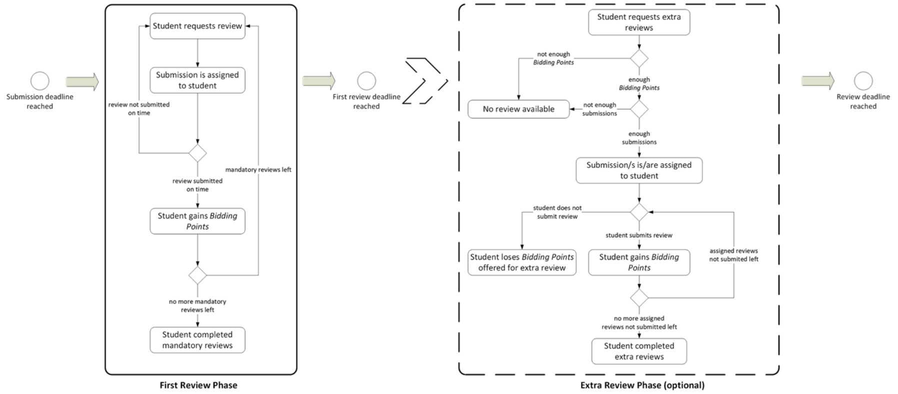

Fig. 2 Workflow of the new dynamic review allocation mechanism with extra bidding

decided to perform the assigned reviews resulting in many submissions with no evaluation.

3.2.5 *Automatic based on N categories of reviewers* approach

This mechanism is similar with the one presented previously. The submissions are automatically allocated to the number of reviewers specified by the instructor once the submission deadline is reached. In this approach the reviewers are split in N categories (instead of three: HRS, MRS, LRS) based on their assessment skills. The mechanism assigns from each review category the student with the lowest number of assigned reviews until the present.

3.3 Dynamic review allocation with extra bidding

The latest version of LearnEval includes the new dynamic review allocation mechanism with extra bidding. As opposed to the static review mechanisms presented above, where the assignment is done only once, at the beginning of the reviewing phase, a dynamic allocation mechanism implies that the assignment is updated throughout the session, based on students' requests. This feature allows to rebalance the number of reviews per solution on the fly as the current allocation state is taken into consideration.

The review period is split into two phases: a mandatory first review phase, and an optional extra review phase (see Fig. 2).

In the first review phase the students are required to submit reviews to increase their involvement and reviewing score, whereas the extra review phase builds on the students' motivation, desire and availability to request and perform additional evaluations, apart from the required ones. At the core of the review allocation mechanism are the bidding points. They introduce an aspect of gamification (Deterding et al., 2011; Tenório et al., 2016) representing an incentive for the students to get involved and contribute more to the assessment process, and even perform extra evaluations.

In our case, the instructor used these bidding points to provide a bonus score at the end of the course for the highly engaged students. The points can be earned in three ways: by providing the mandatory reviews, by submitting extra reviews, or by offering back-reviews.

The *bidding* mechanism is highly configurable and can be easily personalized by the instructor. The following settings related to the new review allocation mechanism, and implicitly, the bidding mechanism, can be configured:

- *Bonus Review Bidding Points First Review Phase (BRBPFRP)*, specifies the amount of bidding points earned by the student by submitting a review in the first review phase. By default, the value is 100.
- *Bonus Review Bidding Points Extra Review Phase (BRBPERP)*, specifies the amount of bidding points earned by the student by submitting a review in the extra review phase. By default, the value is 300. It is recommended that this value is higher than *BRBPFRP.*
- *Points to Bid per Additional Review (PBAR)*, specifies the amount of bidding points necessary to be offered by the student in order to be assigned an extra review. By default, the value is 100. This value must be lower than *BRBPERP.*
- *Maximum Allowed Additional Reviews to Bid for (MAARB)*, specifies the number of additional reviews the student can request in the extra review phase. By default, the value is 3. When setting this value, the instructor should seek a balance between the students' workload and the reliability of the grade assigned to the submission.
- *Hours Allowed to Review Solution Allocated (HARSA)*, specifies the number of hours the student has available to submit a review in the first review phase, before the submission returns to the allocation pool. By default, the value is 5.
- *Bonus Bidding Points Back Review (BBPBR)*, specifies the amount of bidding points earned by the student by submitting a back-review. By default, the value is 50. It is recommended that this value is lower than *BRBPFRP.*

3.3.1 First review phase

An essential feature of the first review phase is its dynamic character that counteracts some of the important issues found with the static approach. This method allows for all the submissions to receive relatively the same number of submitted reviews at the end of the assessment activity (the maximum difference between the most and the least reviewed solution can be one). The submissions are not allotted statically (only once, when the submission deadline is reached), but at the reviewer's request. The learner must submit the number of required reviews specified by the instructor to increase their *involvement score*. Therefore, the number of assessments a student can perform is not dependent anymore on the number of submitted solutions, as was the case in the previously presented static approaches. Every student can perform three reviews (or the number of required reviews defined by the instructor) and each submission will receive at least three evaluations. By contrast, in the static approach, if the submission rate is low, each student is assigned a relatively low number of reviews to perform.

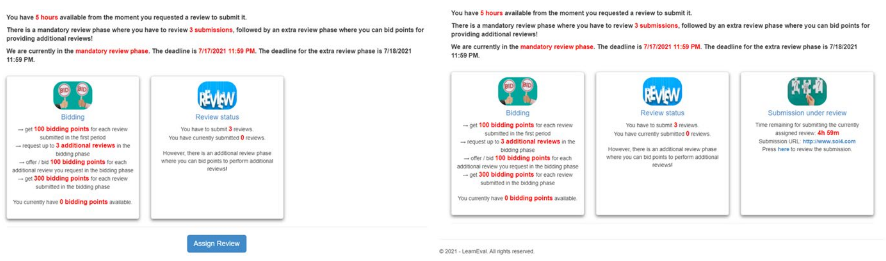

Fig. 3 LearnEval mandatory review phase - Assign review (left) & Review assigned (right)

Once the submission deadline is reached, the student enters the *Review Solutions* module (Fig. 3) and various information regarding the review activity is available, such as: number of evaluations required to submit, number of evaluations already submitted, time available to submit an assigned review, deadlines for both review phases, and detailed information about the bidding process. The student can request a deliverable to review by pressing the *Assign Review* button.

The selection of the submission to assess, in order to allow a fair distribution of the reviewers, involves the following steps:

1. *Compute X - the lowest number of submitted reviews a deliverable currently has*
2. *Retrieve all the deliverables with X submitted reviews*
3. *Select from the previous list the deliverable with the lowest number of submitted reviews from the category of the current reviewer*

The second major feature of the allocation mechanism is represented by the review expiration time. A similar approach can be found in the solution proposed by Staubitz et al. (2016). The reviewer has a limited time window to submit the review assigned to them; this period can be configured by the instructor, with the default value being 5 hours. The deliverable is returned to the allocation pool in case the student does not submit the review before the expiration deadline. This allows the system to subsequently allocate the submission to another student, avoiding the case when the deliverable could be left without any assessments. Finally, when a review is submitted, the reviewer is rewarded with *BRBPFRP* bidding points and the number of submitted evaluations for the assessed deliverable is increased by one.

3.3.2 Extra review phase

The extra review phase is an optional stage that allows students to perform additional evaluations. It tries to solve various cases that happen in practice with students: they did not have time to perform the mandatory reviews and wish to recoup, they want to gain more reviewing experience, they wish to view more of their colleagues' solutions, or they want to provide more help to their peers.

This phase starts once the deadline for the first review phase is reached. Again, the student must enter the *Review Solutions* module to take part in the review

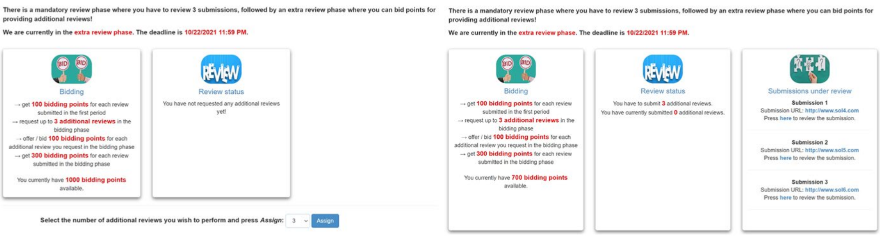

Fig. 4 LearnEval extra review phase - Assign additional reviews (left) & Reviews assigned (right)

activity (Fig. 4). The page is to a large extent similar with the previous phase, with a few changes. In the lower part of the page the student can select the number of extra submissions they wish to review. The upper limit of extra reviews a student can request is 3 by default, but it can be configured by the instructor. The student must offer *PBAR* bidding points for each additional extra assigned review. Therefore, the student must have the necessary bidding points when requesting extra reviews, otherwise a notification is displayed informing the student about the maximum limit of extra evaluations that can be requested. The allocation of the submissions to reviewers is performed in a similar way to the first review phase, i.e., taking into consideration the number of submitted reviews for the deliverable and the review category of the requester. The student receives as a reward *BRBPERP* bidding points for each extra assessment performed. On the other hand, to encourage the submission of the review, in case it is not performed before the deadline, the offered bidding points (*PBAR*) are lost. In this stage, the only time limit to submit the assessment is the extra review phase deadline.

Several core features of this phase increase the reliability of the peer assessment process. The confidence in the final grades assigned to the solutions raises as the number of assessments submitted for the solutions increases. The bidding points are helpful in two ways: to stimulate the students to perform additional reviews, and to allow the instructor to employ them in different ways (e.g., as bonus score for the final grade). However, as participation to this phase is optional, the amount of bidding points gathered and the number of extra assessments performed by the student do not affect their *involvement score*.

Figure 5 illustrates the *Settings* module where the instructor can select the review allocation mechanism and configure the various parameters related to it (default values are provided for each parameter).

A comparison between the initial (static) review allocation approaches presented in section 3.2 and the new (dynamic) review allocation approach introduced in this section is included in Table 1.

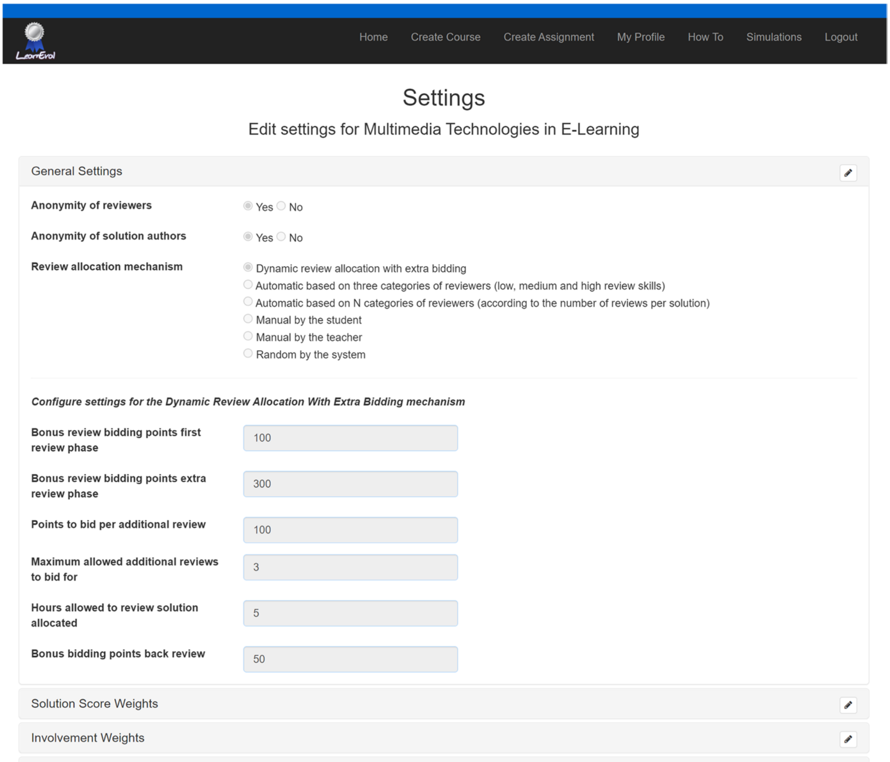

Fig. 5 LearnEval settings module (teacher area)

4 Practical application of the review allocation mechanism

4.1 Course context

The new review allocation mechanism was employed in the context of a Multimedia Technologies in E-Learning (MTEL) course, involving 4th year students from University of Craiova, Romania, in the first semester of the 2020-2021 academic year. 40 students were registered in the LearnEval peer assessment platform, together with one instructor. The platform was also applied in the previous delivery of the course, in 2019-2020 academic year, but employing the static review allocation mechanism, *Automatic based on three categories of reviewers* (presented in section 3.2.4). Students' informed consent regarding data collection, usage and processing was obtained when they registered in the LearnEval platform, according to ethical guidelines in research (Muravyeva et al., 2020). The course followed a PBL approach where each student had to individually develop a project. The requirements of the project were the same in both editions of the course: the learners had to build a website presenting an Informatics lesson, as well as include multimedia content, such as images, video, audio, animations or

Table 1 Comparison of review allocation approaches implemented in LearnEval

Mechanism name	Manual vs automatic	Allocator	Based on reviewing skills	Static vs dynamic	Main advantages	Main disadvantages
Manual by the student	Manual	Student	No	Static	The students can select the submissions they are interested to evaluate	Biased, the students might select only submissions from a few preferred peers
Manual by the teacher	Manual	Teacher	No	Static	The teacher can allocate submissions according to specific needs and inside knowledge	Not scalable, in courses with high enrollment rate and many assignments the workload can be very difficult to handle
Random by the system	Automatic	System	No	Static	Very simple, eliminates the time required for manual allocation	Unfair, a student with low reviewing skills could be assigned to a low-quality submission, resulting in little gain for both the reviewer and the reviewee
Automatic based on three categories of reviewers	Automatic	System	Yes	Static	Relatively fair for both reviewers and reviewees	If the involvement of reviewers varies, some submissions can be left with few evaluations
Automatic based on N categories of reviewers	Automatic	System	Yes	Static	Relatively fair for both reviewers and reviewees	If the involvement of reviewers varies, some submissions can be left with few evaluations
Dynamic review allocation with extra bidding	Automatic	System	Yes	Dynamic	Fair for both reviewers and reviewees; the students can perform additional reviews, increasing the number of evaluations per submission	In rare cases, after the extra review phase, the difference among submissions regarding the number of reviews can be higher than one

Table 2 Number of solutions and reviews submitted by the students in the first review phase

	Assignment I	Assignment II	Assignment III	Overall
Solutions submitted	36 (90%)	35 (88%)	37 (93%)	108 (90%)
Reviews submitted (first review phase)	88 (81%)	81 (77%)	90 (81%)	259 (80%)

educational games. The first assignment demanded learners to provide the layout of the website and the navigational structure, the second assignment demanded learners to code the algorithm behind the lesson and include multimedia material, whereas the last assignment demanded learners to incorporate educational games, animations and knowledge assessment surveys.

The peer assessment activity was mandatory and represented 30% of the final grade assigned to the project. Furthermore, the peer assessment settings were similar in both editions of the course. The students had to upload the project deliverables in LearnEval before each presentation and a dedicated peer assessment session was devised for each of the three assignments. The students could attend a calibration phase where they assessed two test solutions at the beginning of the semester, prior to the unfolding of the peer assessment activity. The review criteria were identical in all the sessions, being related to: implementation, functionality, and aesthetic and pedagogical quality. The first review phase lasted one week and required learners to submit three evaluations. The student had five hours to submit the review after requesting it, otherwise the deliverable would return to the allocation pool. The students gained 100 bidding points for each assessment submitted in this phase. As the latest edition of the course employed the new review allocation mechanism, an extra review phase was held where the students could perform up to three additional assessments, over a period of one week. The students had to bid 100 points for each extra review assigned and gained in return 300 bidding points for each submitted assessment; furthermore, 50 bidding points were gained for each submitted back-review.

4.1.1 Overview of the students' involvement

Table 2 summarizes the number of solutions and reviews provided by the students in 2020-2021 academic year[2]. The number of solutions submitted by the students was high and relatively similar for all the three sessions. Furthermore, it can be noticed that the number of reviews provided by the learners was relatively constant, although in the 2nd assignment it was a bit lower; overall, a large number of the reviews were submitted (80%).

Table 3 depicts the number of reviews submitted in the first review phase compared with the number of reviews submitted in the extra review phase. Interestingly, while the number of submitted mandatory reviews did not see an increase along the

[2] Percentages are rounded to the nearest integer throughout the paper.

Table 3 Number of reviews submitted by the students in the first review phase (P1) and extra review phase (P2)

	Assignment I			Assignment II			Assignment III			Overall		
Reviews	P1	P2	Total	P1	P2	Total	P1	P2	Total	P1	P2	Total
Count	88	37	125	81	38	119	90	46	136	259	121	380
%	70	30	100	68	32	100	66	34	100	68	32	100
P1/P2	2.38			2.13			1.96			2.14		

Table 4 Number of reviews submitted by the students in 2019-2020

	Assignment I	Assignment II	Assignment III	Overall
Reviews submitted	47 (87%)	46 (70%)	40 (63%)	133 (73%)

Table 5 Number of submissions that received one, two, and respectively three reviews for each of the peer assessment sessions

	Assignment I	Assignment II	Assignment III	Overall
Solutions submitted	18	22	21	61
Solutions with 1 review	0	5 (23%)	5 (24%)	10 (16%)
Solutions with 2 reviews	7 (39%)	10 (45%)	13 (62%)	30 (49%)
Solutions with 3 reviews	11 (61%)	7 (32%)	3 (14%)	21 (34%)

sessions (recording a minimum in the 2nd assignment), the number of submitted extra reviews slightly increased session by session, and in the last assignment the share of extra reviews represented more than a third of the total number of reviews submitted by the students. These results unveil that an important portion of the students wanted to provide additional reviews, on top of the required ones.

4.2 Results

4.2.1 Overview of the review allocation results from 2019-2020 academic year

In what follows we start with presenting the results obtained in the previous year (Badea & Popescu, 2022), to provide a baseline for comparison in terms of reviewing involvement and review allocation fairness. 29 students registered in LearnEval, out of 30 learners enrolled in the course. Table 4 summarizes the number of reviews provided by the students in 2019-2020 academic year. It can be noticed that the number of reviews decreased over time. Overall, a relatively large share of the reviews were submitted (73%), but lower than the 2020-2021 academic year (80%).

Table 5 depicts the number of submissions that received one, two, and respectively three reviews, in the peer assessment sessions. The high reviewing rate of the

Table 6 Distribution of the number of reviews submitted by the students

Number of reviews	0	1	2	3	4	5	6	7	8
Number of reviewers	4 (14%)	2 (7%)	2 (7%)	1 (3%)	2 (7%)	2 (7%)	7 (24%)	8 (28%)	1 (3%)

Table 7 Number of submissions that received one, two, and respectively three reviews for each of the peer assessment sessions after the first review phase

	Assignment I	Assignment II	Assignment III	Overall
Solutions submitted	36	35	37	108
Solutions with 1 review	0	0	0	0
Solutions with 2 reviews	20 (56%)	25 (71%)	21 (57%)	66 (61%)
Solutions with 3 reviews	16 (44%)	10 (29%)	16 (43%)	42 (39%)

students during the 1st assignment meant that all the submissions received two or three assessments. However, the lower reviewing rates during the 2nd and 3rd assignment (70% and 63% respectively) caused the share of solutions that received only one review to be relatively large (23% and 24% respectively). Therefore, a high variation can be noticed in the number of evaluations a submission received, especially in the 2nd assignment. Furthermore, this variation is highlighted by considering all the submissions together (see *Overall* column).

Table 6 illustrates the distribution of the number of reviews submitted by the students. The maximum number of assessments a student could submit was eight, although there were cases when the limit was seven (depending on the number of solutions submitted). The number of reviews provided by the learners varied considerably. Four students (14%) did not submit any reviews, while two students (7%) submitted only one evaluation. However, around half of the students submitted six (24%) or seven (28%) reviews.

As can be seen, significant issues related to the static review allocation approach were encountered: non-uniform distribution of the number of reviews submitted per solution in the 2nd and 3rd assignment, many solutions received only one or two evaluations, and an important share of the students (38%) performed less than five reviews. Hence, the new review allocation mechanism attempts to solve these issues, as described next.

4.2.2 Overview of the review allocation results from 2020-2021 academic year

In the next delivery of the MTEL course the number of registered students in LearnEval was somewhat higher (40 learners), as the total number of students enrolled in the course was also higher (44 students). Table 7 depicts the number of submissions that received one, two, and respectively three reviews, after the first review phase in the peer assessment sessions. Noteworthy, the reviewing process was fair as all the submissions received two or three assessments, thus the difference between the most and the least reviewed solution was one.

Table 8 Number of submissions that received two, three and respectively four reviews for each of the peer assessment sessions after the extra review phase

	Assignment I	Assignment II	Assignment III	Overall
Solutions submitted	36	35	37	108
Solutions with 2 reviews	2 (6%)	0	0	2 (2%)
Solutions with 3 reviews	15 (43%)	22 (63%)	12 (32%)	49 (45%)
Solutions with 4 reviews	19 (53%)	13 (37%)	25 (68%)	57 (53%)

Table 9 Distribution of the number of reviews submitted by the students in the mandatory review phase and extra review phase

Number of reviews	0	1-4	5-8	9
Number of reviewers (mandatory review phase)	7 (18%)	4 (10%)	5 (13%)	24 (60%)
Number of reviewers (extra review phase)	19 (48%)	9 (23%)	5 (13%)	7 (18%)

Table 8 summarizes the number of submissions that received two, three and respectively four reviews (no solution received less than two assessments), after the extra review phase in the peer assessment sessions. Noteworthy, all the submissions received three or four assessments, with the exception of two solutions in the 1st assignment. This was due to the fact that a student did not submit the two assigned extra reviews, hence the submissions were left with only the two evaluations received during the first review phase. Such cases when students request extra reviews but do not submit them are very rare, as can be noticed from the rest of the assignments where no such situations occurred. Furthermore, as an incentive to submit the assessments, the student loses the bidding points offered for the assigned extra reviews (in our context 100 bidding points for each extra review) in case they do not submit them.

Table 9 illustrates the distribution of the number of reviews submitted by the students in the mandatory review phase (second row). The maximum number of assessments a student could submit was nine. The number of reviews provided by the learners varied less compared with the previous year. More than half of the students (60%) performed all the required assessments; however, there were still several students who performed no reviews (7, or 18%). The last row of Table 9 depicts the distribution of the number of reviews submitted by the students in the extra review phase. The maximum number of extra assessments a student could request and submit was nine. Approximately half of the students decided to perform extra reviews (21, or 53%); this percentage is higher than the one reported in Kulkarni et al. (2015) where 23% of the students reviewed more than the required number of two submissions. Almost a third of the students (30%) submitted at least five extra reviews. Furthermore, a small part of the students performed all the possible extra assessments (7, or 18%).

Table 10 illustrates the distribution of the amount of bidding points gathered by the students at the end of the peer assessment activity. Almost two thirds of the

Table 10 Distribution of the amount of bidding points gathered by the students

Number of bidding points	[0, 900)	[900, 1800)	[1800, 2700)	[2700, 3600]
Number of students	15 (38%)	13 (33%)	6 (15%)	6 (15%)

students (25, or 63%) acquired more than 900 points, the amount of bidding points gathered in case all the required nine reviews were performed; this denotes a high level of involvement in the reviewing activity.

The results presented in this section highlight the advantages of applying the new dynamic review allocation mechanism with extra bidding compared with a static approach. The allocation is fairer as each submission received relatively the same number of evaluations after the first review phase and after the extra review phase; the difference between the most and the least reviewed solutions is one in the vast majority of cases (there are only two exceptions after the extra review phase). By contrast, in the previous year this metric varied to a larger extent, many solutions receiving three evaluations while others received only one. In addition, we noticed that the bidding points stimulated learners to provide more reviews. Many students decided to attend the extra review phase, earning a significant amount of bidding points. Furthermore, the additional reviews resulted in a higher number of assessments per submission compared to the previous year.

4.2.3 Learners' perception regarding the dynamic review allocation mechanism with extra bidding

We conceived and disseminated a dedicated questionnaire at the end of the semester to investigate the learners' experience with the peer assessment activity and LearnEval platform. Twenty-two of the students (i.e., 55%) filled in this questionnaire. In the following, we report on the items related to the new dynamic review allocation mechanism.

1) *Involvement in the extra review phase*

In terms of involvement, most respondents (20, or 91%), attended at least one of the extra review phases, with 17 students (77%) declaring they usually performed all the three extra reviews. Many students said they attended the extra review phase to gain bidding points, but also to "*help other colleagues*" or "*view other solutions*".

2) *Usefulness of the extra review phase*

More than half of the students (12, or 55%), considered the extra review phase as being useful. But there were also some cases in which the helpfulness of this extra review phase was not very clear for the students (e.g.: "*I do not consider it really necessary. The time for it could be allocated for something else.*"). However, most of the students (15, or 68%), agreed that the reviewing module was intuitive to use.

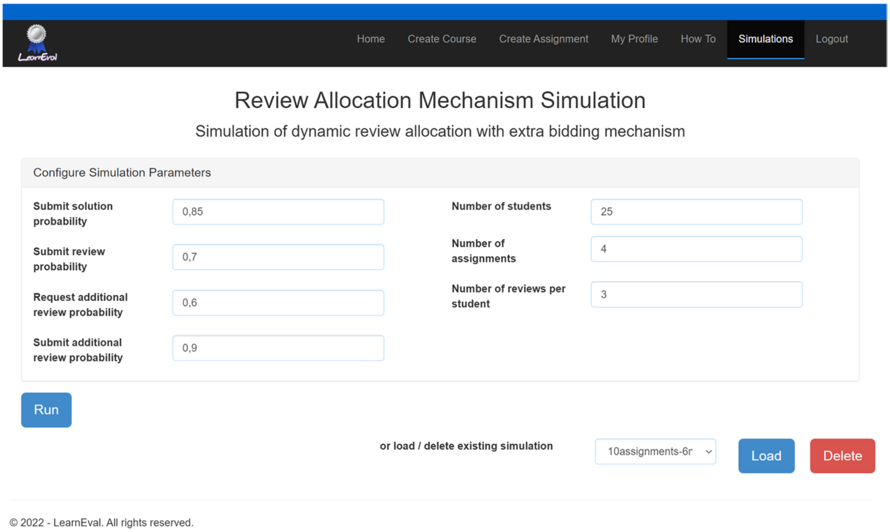

Fig. 6 Dynamic review allocation mechanism simulation module

3) *Bidding points*

Almost half of the students (10, or 45%), declared that they monitored the bidding points gathered on a regular basis. Most of the students (20, or 91%), agreed that the bidding points they collected were enough to request all the extra reviews they wished to perform.

4) *Satisfaction with LearnEval platform*

Overall, most of the students (15, or 68%), were satisfied or very satisfied with the peer assessment platform. The comments were primarily positive, e.g.: "*everything was fine and there were no problems*", "*it was easy to use*", "*I was content with the platform*", or "*good platform for reviewing others*".

5 Simulation module for the dynamic review allocation mechanism

A dedicated simulation module was developed, as part of LearnEval, to assess the new dynamic review allocation mechanism's applicability and scalability. The simulation module takes as input several configuration parameters, executes the review allocation algorithm, and displays the allocation results using various intuitive data visualization components.

The configuration parameters supported by the simulation module (Fig. 6) are the following:

- *Submit solution probability (SSP)* – the probability for a student to submit a solution

- *Submit review probability (SRP)* – the probability for a student to submit a review
- *Request additional review probability* – the probability for a student to request an extra review
- *Submit additional review probability* – the probability for a student to submit the assigned extra review
- *Number of students* – the number of students enrolled in the course
- *Number of assignments* – the number of assignments (peer assessment sessions) to be simulated
- *Number of reviews per student* – the number of reviews required to be performed by a student in the first review phase.

At the beginning of the simulation, each student is randomly assigned a reviewing score that depicts their initial assessment skills. Next, the students are allocated to one of the three reviewing categories. They are ordered descendant according to their reviewing score and the first third are allotted to HRS, the second third are allotted to MRS, and the rest are allotted to LRS. The simulation starts with the configured number of assignments and for each one, it creates and simulates a peer assessment session. An assignment contains three stages: submission phase, first review phase, and extra review phase.

1) *Submission phase simulation*

In the submission phase, the algorithm iterates over each student and based on the *SSP* configuration parameter a mock-up submission is created or not. At the end of this phase, the first review phase simulation starts.

2) *First review phase simulation*

In this phase, the algorithm iterates over each student and computes the probability for them to submit the reviews. However, based on the observed behavior from previous courses, reviewers commonly follow two trends: students that submit every required review, and students that submit no reviews at all. The cases in which students submit only a part of the assessments are more rare. Therefore, we have two core probabilities inferred by the system based on the *SRP* defined by the instructor: the probability to submit all the required reviews, and the probability to submit none of them. A third one, the probability to provide only a part of the assessments, is lower than the first two and from our observations it is typically around 5% (i.e., about 1 in 20 students decides to perform only some of the evaluations). Furthermore, we observed that students having lower review competence and belonging to LRS category usually submit a lower number of assessments compared with the ones from MRS and HRS categories. Thus, the probability for a LRS student to submit an assessment is half compared to the rest. However, in the 1st assignment the actual assessment skills of the peers are not known, thus, the probability is the same regardless of the category. The allocation of the submissions to reviewers uses

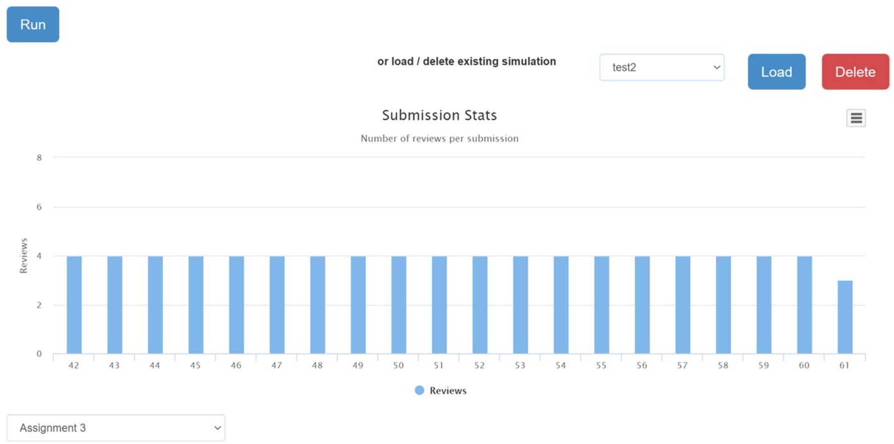

Fig. 7 Dynamic review allocation mechanism simulation module – number of reviews per submission

the new dynamic review allocation approach: it considers the number of submitted reviews for the deliverable and the review category of the requester. At the end of this phase, the extra review phase simulation starts.

3) *Extra review phase simulation*

In this phase, the algorithm iterates over each student and computes the probability for them to request extra reviews, and then further computes the probability to submit them. The student behavior is again modeled based on our observations from previous courses where peer assessment activity was employed. Thus the student behavior in the extra review phase is similar in many aspects with the one exhibited in the first review phase: the probability for a student to request all or none extra assessments is higher than the probability to request only a few of them; the probability of a LRS student to request extra assessments is lower compared to the rest of the students. However, the probability to request and perform all the reviews in this phase is lower compared to performing all the reviews in the first review phase. Again, the allocation of the submissions to the reviewers uses the new dynamic review allocation approach: it considers the number of submitted reviews for the deliverable and the review category of the requester.

For illustration purposes we performed the simulation with the parameters shown in Fig. 6. The results are shown in Fig. 7, 8 and 9. Figure 7 depicts the number of assessments each solution received at the end of the extra review phase after the simulation of the 3rd assignment. All the submissions, with only one exception, received the same number of evaluations (i.e., four). The simulation results highlight the fairness of the allocation, as the difference between the highest and the lowest number of evaluations a solution receives is one, both after the first review phase (Fig. 8) and after the extra review phase (Fig. 7). Furthermore, each solution receives relatively the same number of evaluations from each review category. Comparing Fig. 9 with Fig. 8 we can notice that in the extra review phase the additional

Fig. 8 Dynamic review allocation mechanism simulation module - first review phase allocation results

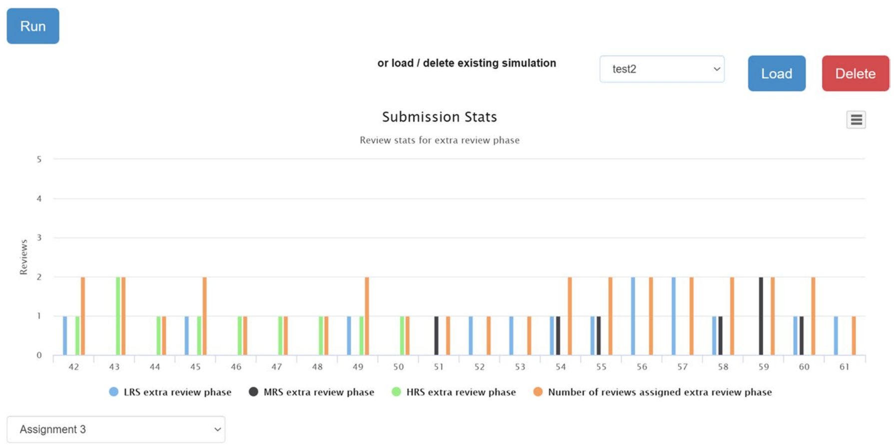

Fig. 9 Dynamic review allocation mechanism simulation module - extra review phase allocation results

assessments were allotted to the solutions which received less reviews in the first review phase.

The visualization component of the simulation module is comprehensive, displaying many aspects related to the allocation results, such as: number of *HRS/MRS/LRS* students allotted per solution, confidence in the grade assigned to the submission computed based on the assessors' competences, number of evaluations submitted by the students after the first review phase for every assignment, number of evaluations submitted by the students after the extra review phase for every assignment, overall number of evaluations submitted by the students for every assignment, total number of evaluations submitted by each student, as well as the amount of

Table 11 Estimated probabilities based on students' behavior exhibited in MTEL 2020-2021 course

Simulation configuration parameter	Parameter value
Submit solution probability	90%
Submit review probability	72%
Request additional review probability	39%
Submit additional review probability	90%
Number of students	40
Number of assignments	3
Number of reviews per student	3

bidding points gathered. The simulation results can be saved for later visualization and examination.

In a subsequent step, we ran the review allocation simulation module with configuration parameters' values inferred from students' observed behavior in the MTEL course (see Table 11). The simulation results are very close to the actual results recorded in the peer assessment activity. Thus, we can attempt to extrapolate some parameters' values, such as the number of students enrolled, number of peer assessment sessions or evaluations per student, and visualize the allocation results.

The estimated probabilities were used to simulate the allocation process for 10 assignments, 6 required reviews per student and 3 extra optional reviews per student. The outcomes were promising as every solution received relatively the same number of evaluations for each assignment (the difference between the most and the least reviewed solution was one). Furthermore, the distribution of assessors per solution in terms of review category was relatively balanced (Fig. 10).

6 Discussion and conclusion

According to the literature on peer assessment, the allocation of the solutions to reviewers is done mainly randomly, without any consideration regarding the competence of the assessor, and statically, without considering the fact that some students do not submit the assigned reviews. The random and static approaches have shown their limitations: some solutions receive more reviews while others receive less, some submissions are solely evaluated by highly competent assessors while others are solely evaluated by less competent reviewers. In recent years, novel approaches have been proposed with promising and satisfying results regarding the fairness of the allocation process (Abrache et al., 2021; Estévez-Ayres et al., 2013; Staubitz et al., 2016). The results of our work support the idea that the review allocation procedure should be done dynamically, at student's request, in order to consider the ongoing assessment state of the solutions.

The current paper presents several static review allocation approaches supported by our LearnEval peer assessment platform and highlights the challenges they exhibit. These approaches can create inequalities which affect the reliability of the

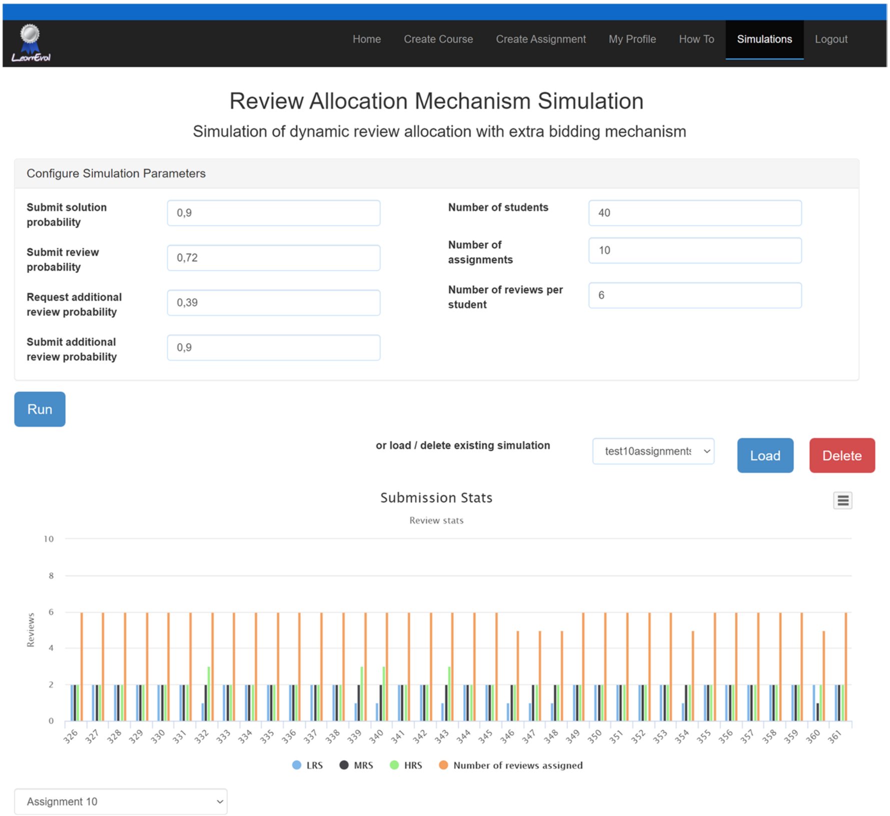

Fig. 10 Dynamic review allocation mechanism simulation module - allocation results with 10 assignments and 6 reviews required per student

peer assessment process, hence, an innovative dynamic review allocation mechanism is proposed. This novel mechanism builds on existing dynamic review allocation approaches, such as the one proposed by Staubitz et al. (2016), but integrates substantial innovative enhancements. The mechanism splits the reviewing activity into two stages: a mandatory first review phase and an optional extra review phase. The mandatory review phase uses features such as the dynamic allocation and review expiration deadline to increase the likelihood for a solution to be evaluated. On the other hand, the extra review phase allows peers to perform additional assessments, increasing the number of reviews each submission receives, and implicitly, the entire process reliability. A new concept is introduced, bidding points, that allows peers to request extra reviews. The new mechanism was employed in 2020-2021 academic year in the context of a Multimedia Technologies in E-Learning course. At the end of the first review phase each solution had two or three evaluations. Furthermore, at the end of the extra review phase each solution had three or four evaluations (with the exception of two submissions), denoting a relatively high interest of the students in the additional reviewing activity. The outcomes show that the current proposal offers a viable, effective and fair mechanism where each solution receives relatively the same number of evaluations. Comparisons with a previously applied static review allocation approach show an increased number of assessments per solution and a more equitable allocation. Furthermore, students' answers to an opinion survey regarding the new mechanism were mainly positive and appreciative.

A dedicated configurable simulation module that emulates the execution of the proposed mechanism and presents the outcomes using various intuitive visualization components was also integrated as part of LearnEval. The module allows researchers to run the allocation process using different configuration parameters, such as: high number of enrolled students, high number of assignments, diverse number of reviews required per student, or various involvement levels. The simulation was executed with the parameters' values recorded from students' behavior in an actual course, but using a higher number of assignments and a higher number of reviews required per student. Although employing students' behavior from a single course offers limited insights, the allocation results showed that the allotment is fair irrespective of the number of assignments or number of reviews required. In the future we plan to study more in-depth the various behaviors exhibited by the students in different courses and perform extensive simulations.

Furthermore, various enhancements could be envisioned for the proposed review allocation mechanism. In the current version, a student can request additional reviews in the extra review phase only once – a multiple request functionality could be included. Moreover, the extra reviews could have an expiration deadline (just like the mandatory reviews), allowing solutions to be returned to the allocation pool in case the assessments are not submitted on time. Another limitation of the current mechanism is that the bidding points can be visualized only as a numeric value. Adding various badges, score levels and a graphical visualization component could provide an enhanced gamification experience for the students.

Finally, a shortcoming of the current study is represented by the limited number of students enrolled in the course and the low number of assignments. Although we simulated the unfolding of the allocation mechanism with different configuration parameters, the actual distribution of the reviews could differ in a real scenario where more students and assignments are involved. Therefore, as future work we plan to perform more experiments in different contexts and more comprehensive data analyses in order to investigate the fairness and reliability of the allocation process and peer assessment activity.

Data availability The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Conflict of Interest None.

References

Abrache, M.-A., Bendou, A., & Cherkaoui, C. (2021). Clustering and Combinatorial Optimization Based Approach for Learner Matching in the Context of Peer Assessment. *J Educ Comput Res, 59*(6), 1135–1168. https://doi.org/10.1177/0735633121992411

Abrache, M., Megder, K., & Cherkaoui, C. (2018). Feedback-Oriented Assessor Model - Application: Allocation of Submissions in Online Peer Assessment. In E. Abdelwahed, L. Bellatreche, M. Golfarelli, D. Méry, & C. Ordonez (Eds.), *Model and Data Engineering, MEDI 2018, Lecture Notes in Computer Science, V11163* (pp. 137–149). Springer. https://doi.org/10.1007/978-3-030-00856-7_9

Badea, G., & Popescu, E. (2019a). A web-based platform for peer assessment in technology enhanced learning: student module prototype. In *Proc. of ICALT 2019* (pp. 372–374). IEEE Computer Society Press. https://doi.org/10.1109/ICALT.2019.00115

Badea, G., & Popescu, E. (2019b). Instructor support module in a web-based peer assessment platform. In *Proc. of ICSTCC 2019* (pp. 691–696). IEEE Computer Society Press. https://doi.org/10.1109/ICSTCC.2019.8885687

Badea, G., & Popescu, E. (2019c). Using LearnEval peer assessment platform in project-based learning settings: A first experience report. In *Proc. of ITHET 2019* (pp. 1–5). IEEE Computer Society Press. https://doi.org/10.1109/ITHET46829.2019.8937359

Badea, G., & Popescu, E. (2020a). Supporting students by integrating an open learner model in a peer assessment platform. In *Proc. of ITS 2020. Lecture Notes in Computer Science* (Vol. 12149, pp. 105–114). Springer. https://doi.org/10.1007/978-3-030-49663-0_14

Badea, G., & Popescu, E. (2020b). Using peer assessment in conjunction with project-based learning: A comparative study. In *Proc. of ICALT 2020* (pp. 376–380). IEEE Computer Society Press. https://doi.org/10.1109/ICALT49669.2020.00119

Badea, G., & Popescu, E. (2020c). Analyzing the validity of the peer assessment process in a project-based learning scenario: Preliminary results. In *Proc. of ICSTCC 2020* (pp. 831–836). IEEE Computer Society Press. https://doi.org/10.1109/ICSTCC50638.2020.9259671

Badea, G., & Popescu, E. (2022). LearnEval peer assessment platform: Iterative development process and evaluation. *IEEE Transactions on Learning Technologies* (in press).

Crespo García, R. M., Pardo, A., Pérez, J. P. S., & Delgado-Kloos, C. (2005). An Algorithm for Peer Review Matching Using Student Profiles Based on Fuzzy Classification and Genetic Algorithms. In M. Ali & F. Esposito (Eds.), *Innovations in Applied Artificial Intelligence, IEA/AIE 2005, Lecture Notes in Computer Science, V3533* (pp. 685–694). Springer. https://doi.org/10.1007/11504894_95

Davies, P. (2000). Computerized peer assessment. *Innov Educ Train Int, 37*(4), 346–355. https://doi.org/10.1080/135580000750052955

Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From game design elements to gamefulness: Defining gamification. In *Proc. MindTrek 2011* (Vol. 11, pp. 9–15). Association for Computing Machinery. https://doi.org/10.1145/2181037.2181040

Estévez-Ayres, I., Crespo García, R. M., Fisteus, J. A., & Delgado-Kloos, C. (2013). An Algorithm for Peer Review Matching in Massive Courses for Minimising Students' Frustration. *J Univ Comput Sci, 19*(15), 2173–2197. https://doi.org/10.3217/jucs-019-15-2173

Freeman, M., & McKenzie, J. (2002). SPARK, a confidential web–based template for self and peer assessment of student teamwork: benefits of evaluating across different subjects. *Br J Educ Technol, 33*(5), 551–569. https://doi.org/10.1111/1467-8535.00291

Gamage, D., Staubitz, T., & Whiting, M. E. (2021). Peer assessment in MOOCs: Systematic literature review. *Distance Education, 42*(2), 268–289. https://doi.org/10.1080/01587919.2021.1911626

Gamage, D., Whiting, M. E., Rajapakshe, T., Thilakarathne, H., Perera, I., & Fernando, S. (2017). Improving assessment on MOOCs through peer identification and aligned incentives. In *Proc. ACM Conference on Learning@ Scale 2017* (pp. 315–318). Association for Computing Machinery. https://doi.org/10.1145/3051457.3054013

Gehringer, E. F. (2001). Assignment and quality control of peer reviewers. In *Proc. of the 2001 American Society for Engineering Education Annual Conference and Exposition* (pp. 6.224.1–6.224.8). https://doi.org/10.18260/1-2%2D%2D8941

Gehringer, E. F., & Cui, Y. (2002). An effective strategy for dynamic mapping of peer reviewers. In *Proc. of the 2002 American Society for Engineering Education Annual Conference and Exposition* (pp. 7.168.1–7.168.16). ASEE. https://doi.org/10.18260/1-2%2D%2D10645

Giannoukos, I., Lykourentzou, I., Mpardis, G., Nikolopoulos, V., Loumos, V., & Kayafas, E. (2010). An adaptive mechanism for author-reviewer matching in online peer assessment. In M. Wallace, I.E. Anagnostopoulos, P. Mylonas & M. Bielikova (Eds.), *Semantics in adaptive and personalized services. Studies in Computational Intelligence* (V279, pp. 109–126). Springer, . https://doi.org/10.1007/978-3-642-11684-1_7

Kulkarni, C. E., Bernstein, M. S., & Klemmer, S. R. (2015). PeerStudio: Rapid peer feedback emphasizes revision and improves performance. In *Proc. ACM Conference on Learning@ Scale 2015* (pp. 75–84). https://doi.org/10.1145/2724660.2724670

Muravyeva, E., Janssen, J., Specht, M., & Custers, B. (2020). Exploring solutions to the privacy paradox in the context of e-assessment: informed consent revisited. *Ethics Inf Technol, 22*, 223–238. https://doi.org/10.1007/s10676-020-09531-5

Lin, S. S. J., Liu, E. Z. F., & Yuan, S. M. (2001). Web-based peer assessment: Feedback for students with various thinking styles. *J Comput Assist Learn, 17*(4), 420–432. https://doi.org/10.1046/j.0266-4909.2001.00198.x

Loddington, S., Pond, K., Wilkinson, N., & Willmot, P. (2009). A case study of the development of WebPA: An online peer-moderated marking tool. *Br J Educ Technol, 40*(2), 329–341. https://doi.org/10.1111/j.1467-8535.2008.00922.x

Mader, S., & Bry, F. (2019). Towards an annotation system for collaborative peer review. In R. Gennari et al. (Eds.), *Methodologies and intelligent systems for technology enhanced learning, 9th International Conference, MIS4TEL 2019. Advances in Intelligent Systems and Computing* (Vol. 1007, pp. 1–10). Cham: Springer. https://doi.org/10.1007/978-3-030-23990-9_1

Murakami, C., Valvona, C., & Broudy, D. (2012). Turning apathy into activeness in oral communication classes: Regular self- and peer-assessment in a TBLT programme. *System, 40*(3), 407–420. https://doi.org/10.1016/j.system.2012.07.003

Politz, J. G., Patterson, D., Krishnamurthi, S., & Fisler, K. (2014). CaptainTeach: Multi-stage, in-flow peer review for programming assignments. In *Proc. ITICSE'14* (pp. 267–272). Association for Computing Machinery. https://doi.org/10.1145/2591708.2591738

Rubin, R. F., & Turner, T. (2012). Student performance on and attitudes toward peer assessments on Advanced Pharmacy Practice Experience assignments. *Curr Pharm Teach Learn, 4*(2), 113–121. https://doi.org/10.1016/j.cptl.2012.01.011

Staubitz, T., Petrick, D., Bauer, M., Renz, J., & Meinel, C. (2016). Improving the peer assessment experience on MOOC platforms. In *Proc. ACM Conference on Learning@ Scale 2016* (pp. 389–398). https://doi.org/10.1145/2876034.2876043

Sukstrienwong, A. (2017). A web-based peer assessment system for assigning student scores in cooperative learning. *TEM J, 6*(4), 717–725. https://doi.org/10.18421/TEM64-10

Tenório, T., Bittencourt, I. I., Isotani, S., Pedro, A., & Ospina, P. (2016). A gamified peer assessment model for on-line learning environments in a competitive context. *Comput Hum Behav, 64*, 247–263. https://doi.org/10.1016/j.chb.2016.06.049

Topping, K. (1998). Peer Assessment between Students in Colleges and Universities. *Rev Educ Res, 68*(3), 249–276. https://doi.org/10.2307/1170598

Topping, K. J. (2009). Peer assessment. *Theory Pract, 48*(1), 20–27. https://doi.org/10.1080/00405840802577569

Verleger, M., Diefes-Dux, H., Ohland, M. W., Besterfield-Sacre, M., & Brophy, S. (2010). Challenges to Informed Peer Review Matching Algorithms. *J Eng Educ, 99*(4), 397–408. https://doi.org/10.1002/j.2168-9830.2010.tb01070.x

Wang, Y., & Sun, F. (2018). How to choose an appropriate reviewer assignment strategy in peer assessment system? Considering fairness and incentive. In *Proc. of the 4th Annual International Conference on Management, Economics and Social Development (ICMESD 2018)* (pp. 603–608). Atlantis Press. https://doi.org/10.2991/icmesd-18.2018.107

Wang, Y. Q., Wang, C., Gao, Y., & Sun, F. Q. (2018). Inappropriate reviewer assignment that causes education inequality in peer assessment. In *Proc. of the 3rd Annual International Conference on Education and Development (ICED 2018)* (pp. 36–42). Atlantis Press. https://doi.org/10.2991/iced-18.2018.6