AsPeer: Method to Self and Peer Assessment in Large Online Design Classes

https://doi.org/10.3991/ijet.v18i10.38617

Hariesh K. Sankaran[1], Raju Shanmugam[1], Bhavesh Shah[1], Sultan Ahmad[2](✉), Hikmat A. M. Abdeljaber[3], Jabeen Nazeer[2]
[1]Karnavati University, Gandhinagar, Gujrat, India
[2]Department of Computer Science, College of Computer Engineering and Sciences, Prince Sattam Bin Abdulaziz University, Alkharj, Saudi Arabia
[3]Department of Computer Science, Faculty of Information Technology, Applied Science Private University, Amman, Jordan
s.alisher@psau.edu.sa

Abstract—Peer and self-assessment open opportunities to scale assessments in online classrooms. This article reports our experiences of using AsPeer-peer assessment system, with two iterations of a university online class. We observed that peer grades highly correlated with staff assigned grades. It was recorded that, the peer grade of all student submissions within range of instructor grade averaged to 21.0% and that within the next 2 ranges was 49.0%. We performed three experiments to improve accuracy of peer grading. First, we observed grading bias and introduced a data driven feedback mechanism to inform peers of it. Students aided by feedback were mindful and performed grading with better accuracy. Second, we observed that the rubric lacked efficiency in translating intent to students. Simplified guiding questions improved accuracy in assessment by 89% students. Third, we encouraged peers to provide personalized qualitative feedback along with rating. We provided them with feedback snippets that addressed common issues. 64% students responded that the snippets helped them to critically look at submissions before rating.

Keywords—peer assessment, self-assessment, online education, design studio, qualitative feedback

1 Introduction

1.1 Evolving role of teacher, student, and classroom

In today's day and age, Massive Open Online Courses, MOOCs, are popular up and coming mode of dissemination of knowledge in the education community [1][2]. Various think tanks and academic institutions believe MOOC may be a leading element in the future of education. These classes provide learners with on-demand video instructional content, automated quizzes, problem sets and discussion forums that allow students to interact with one another. Many open online courses use typical automated assessment methods which makes it impractical to assess open-ended skill based

work that is integral to fields of creative education like design [3]. Donald A. Schön [4] through his 'reflection of action' theory established that the studio method is the standard classroom model for design education [5][6]. The design studio plays a key pedagogical role where one can view, examine and critique others work [7]. However, students in a global online classroom scenario lack the possibility of having to see and share feedback on others work making it difficult to translate essential values and norms in design practice [8]. To improve evaluation in large classrooms, one method would be for students to evaluate their peers work and hence encourage self-assessment and peer-learning [9].

Peer assessment radically relieves staff from large assessment responsibilities and provides them "time"; time that they can invest to enhance their class experience and delivery [10]. Peer assessment reduces the staffs' role as assessors and shifts their emphasis on coaching [11]. Students no more worry about misconceptions that teachers grade on personal taste and focus as they are involved in assessment themselves. Further, when teachers coach but do not grade, students tend to focus more on conceptual understanding and application [12]. Rubrics and assessment criteria provided in advance, projects to students, a fair and consistent grading system based solely on the quality of ones work. Peer assessment activates students to be active learners and assessors with ability to critically reflect on subject matter. Students eventually become participants in their educational process than being a product in a factory line. Empowered with being a participant in the education process it is but natural for students to contribute to peer learning through actionable feedback. Though providing for constructive feedback might not be every student's forte; with guiding snippets on concepts discussed, we have noticed that students can provide for effective feedback.

These online classes or platforms also allow us to study some data regarding students, their engagement, staff, their engagement, course material, submission quality, assessment metrics and a lot more; which can be used to analyze and design better curricula and material. Unlike course materials in in-person classrooms; where apart from presentations; staff can engage in personal and impromptu examples and situations; it is difficult to that in online classes. It is ideal to have a well-structured, comprehendible and fine grained course material [13].

1.2 The traditional design studio

The design studio can be seen as the prototype of design education, particularly for architects, fashion, product, human-centered interactions [14,15] and digital designers [16], and its culture is exemplary.

This studio model of education first developed at the École des Beaux-Arts in Paris [17] in the seventeenth century for the promotion of neoclassical aesthetic values has continued ever since to be used, even by the Bauhaus in Germany in the early twentieth century after function had replaced form as the primary architectural value [18]. Studios provide a shared environment for students to work together, uplifts social motivation, initiates discussions and enhances peer learning while examining one another's work [19]. Critiques and feedbacks both formal and informal help students considerably improve their work [20]. The studio is also the place which facilitates professional socialization where the ethos of a profession is born [21].

Students examine and understand various methods of design when they get to see peer students work along with their work. A typical open studio facilitates display of student work for discussion. In particular, it encourages better self-reflection on their work compared to their peers, looking at the process in each ones design and understands the decision and tradeoffs made [22].

Professional education of architecture and design students have advantaged from this method and laid the importance of formative feedbacks. Informal formative feedback is often through oral critique or 'crit' sessions with the entire group by teachers or other experts [23]. As crits are delivered in public and are open for all students to participate, it encourages them to learn from peer work to explore diverse possibilities [24], and also reflect better on their own design process [25].

Good design has varied criteria and is not explicitly defined for all contexts [26]. During a design jury session expert critiques provide summative assessment based on their training, experience and exposure [27] with no specific rubric to follow [28]. This method does not ask the experts to justify their rating. A more reliable method provides a rubric and assessment process to observe, interpret and evaluate creative work [29].

A good design studio online should encourage the following four requirements. First, it must support open ended creative design work where multiple ideas & perspectives [30] and multiple successful approaches are encouraged to be explored [31]. Second, qualitative feedback, both formative and summative feedback, to be provided at milestone stages defined. Third, peer assessment must aide students to examine and understand how good design principles can be applied in different contexts. Fourth, assessment method should empower students to self-reflect and assess their own strengths, weakness and creative methods.

Prototype Wireframes of Student Projects in the online class

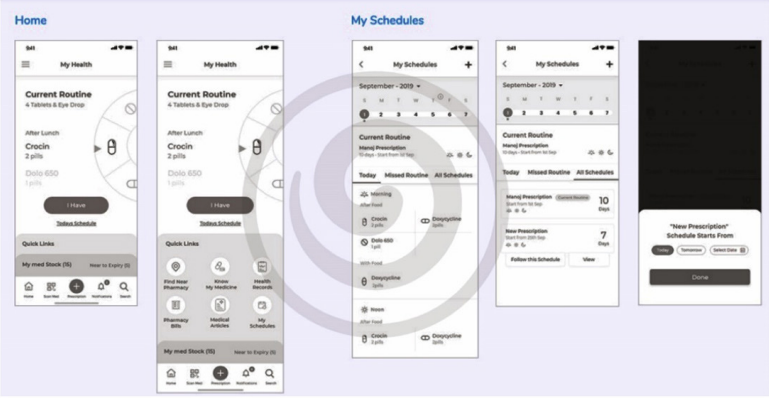

Fig. 1. Prototype from student project in the online class
(low fidelity wireframe stage of a health app)

1.3 Peer assessment – physical and virtual classrooms

Assessment in the large online courses typically relies on automated assessment methods. The variability in design and other open-ended project based education and lack of defined evaluation criteria for such projects makes the implementation of automated assessment challenging [32]. Adding to that automated systems cannot capture relevant constructive feedback that could help students improve on their work [33]. Therefore, open-ended assignments generally rely on human graders. The human grader method is a time intensive exercise where one personally grades, sketches, models, renders and prototypes [2][34]. Such a method is relevant and possible only with a small student-to-grader ratio and makes it ineffective in large online classrooms.

2 AsPeer: the system

AsPeer, is a peer assessment system for large online open-ended classes as shown in Figures 1 and 2. It consists of two stages; where in stage 1 guides peers to range fix peer submissions using a few staff assessed submissions to calibrate. Stage 2 encourages peers to rate submission within a scale of 1–10 comparing submissions within an automatically chosen range. Overall, peer assessed submissions highly correlated with staff assessed submissions.

2.1 Method of the virtual class

This online class is an introduction to user experience and interface design for digital products. All students were enrolled in a University and are offered as an open elective to students from all disciplines ranging from business, fashion, interior and product design. The teaching materials and structure are based on the eight week UX/UI course for Sixth Semester students at Karnavati University, India. During the course, lectures were provided to students as videos to watch along with some short quizzes to reinforce concepts. They also have to complete an assignment per week. Students watch six videos in a week with each video spanning 12 to 15 minutes. In all students completed five design assignments, one at each stage of the design project, where in each covered a milestone. The students had to design a digital product as shown in Figure 1 as a response to one of the given design briefs. In iteration one, 1393 students enrolled for the course while 1274 completed it. Whereas in iteration two, 1678 students enrolled with 1548 going on to complete the course.

2.2 Number and characteristics of students

Typical to a University provided online open elective, the course attracted diverse participants. Both iterations of the course were taken by graduate, post-graduate students, and research scholars both part time and full time. The age range of students was between 18–45 years with the median being 22 years of age. Ten percentage of the students enrolled in the course were from the graduate or postgraduate program in design while the rest were from business, liberal arts, humanities, journalism, medicine and engineering (see Figure 3). In all 1274 completed the course in the first iteration and 1548 completed in the second. The student questions were collated through a survey that students

had to fill at the mid of every week and were answered exclusively through an online live interactive session once at the end of every week. Furthermore, students were encouraged to interact though the online class forum to discuss and clarify immediate doubts.

2.3 Design project and assignments

The course was designed such that the students had to develop one design project to present at the end of the course. The design project had 5 stages with an assignment to be submitted at the end of each. These assignments required students to create sketches, user maps, paper prototypes and physical artifacts and upload them as photographs on their submission drive. Specific templates to compose work for each assignment were provided by the tutors in advance. All assignments were submitted online and graded using AsPeers two stage method. Assignment at each stage of the design project included a unique rubric that explicitly described the criteria for judging student work [11]. These unique rubrics were shared with the students at the beginning of each assignment so students familiarize and refer them to understand the desired milestone to be achieved. Students were informed that peers could see all submitted work before and while grading. Some students' works were used by staff as examples in class announcements and lectures.

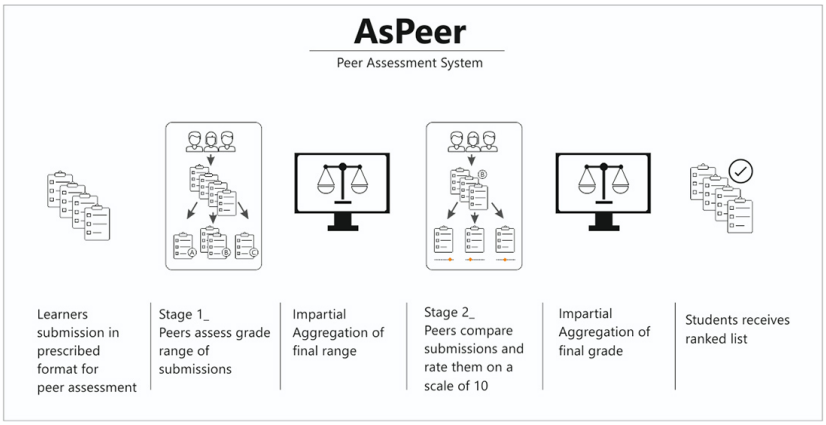

Fig. 2. Process structure of the peer assessment system – AsPeer

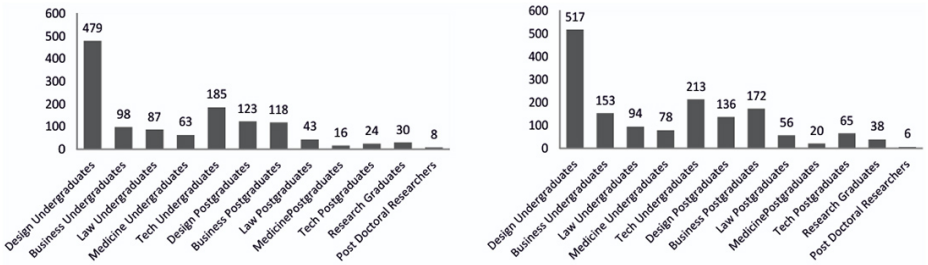

Fig. 3. Occupation and level of students in iteration 1 (left) and iteration 2 (right) of the online course

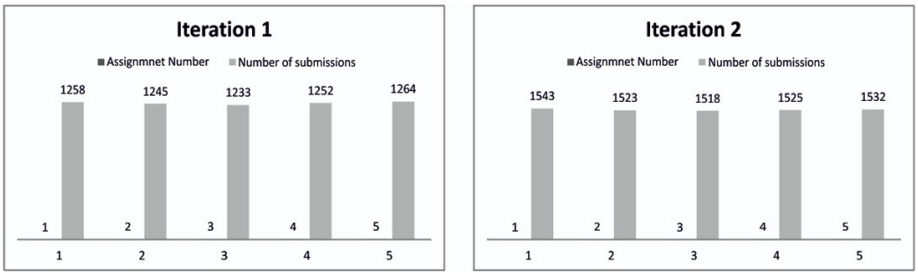

Fig. 4. Number of students who submitted each assignment

2.4 Peer assessment method

AsPeer-peer assessment system consists of two stages. The number of students who submitted each assignment is shown in Figure 4. Stage 1 guides peers to range fix peer submissions using a few staff assessed submissions to calibrate. Stage 2 encourages peers to rate submission within a scale of 1–10 comparing submissions within an automatically chosen range.

Stage 1 – Assessing for grade range. Students as peer reviewers need basic guidance on how to review and grade. A calibrated peer review system [35], guides students in learning to grade by first practicing to grade using sample submissions. After each submission deadline, all three lead staffs evaluated 50 submissions. The mean of all three staff grades were taken as final. Staff grades with higher deviation were re-graded by all three staff for congruence. The staff graded submissions were used to both train students as well as estimate accuracy of grading. Student, only on submission of their assignment, could access the peer assessment page, and had two days to assess for stage 1 and two days to assess for stage 2 to complete the assessment process.

Each student undergoes a training phase where they assess 5 peer works alongside staff graded peer works. Students see staff assessed submissions with an explanation which guides them to calibrate their assessment. They move on to the actual assessment page, once their range of assessment matches with that of staffs, or complete assessing at least five submissions. In the assessment phase, students assess the range of 10 peer submissions. Anonymous to students two submissions as part of the ten assessed by each student are staff graded ones. This establishes a measure of accuracy. Immediately after assessing peers, students assessed their own work in the same method. Some students assessed more than one staff graded submission per assignment as the system would give them fresh ones for calibration when they logged out before finishing assessment or returned to the website after a long time.

An algorithm assigns the median grade range for a student by consolidating assessments from a pool of five randomly selected peer graders. To ascertain the credibility of the process, 20 staff graded submissions were used. This method facilitates the system to estimate the grade range of every student submission, using a small set of randomly selected staff evaluated submissions (since all peers see at least two staff-graded submissions). Thus an agreeable distribution of range is reached between staff and peer.

Stage 2 – Assessing for final grade. Post assessing for grade range of peer submissions in stage 1, the students graded 10 groups of two way compared submissions

shown to them in a randomized manner but belonging to the same range. On completing peer rating, each student rates his work with another submission from the same range. Self-assessment and peer assessment used identical interfaces. The self-assessment stage is only for students to estimate the appropriateness of their grading compared to other students and has no influence on his grades. Stage 2 of the peer assessment was open only for students who had completed stage 1 of the peer assessment processes.

The two-ways comparative grading is recorded on a visual scale of 1–10 (1 being the least effective and 10 being the most effective attempt on the assignment). The comparative grading groups are chosen in random but within the range already established in stage 1 of AsPeer. An algorithm assigns the final grade for a student by calculating the median from five randomly selected peers grading. No staff grading is taken for calibration in this stage though they rate submission similar to how peers rate. The median of the comparative peer grading is taken as the final grade for the assignment making it entirely peer participated assessment (close to real life scenarios). This stage of peer grading is to be completed within two days after notification from the application admin.

2.5 Experiments

We performed three experiments to improve the accuracy of grading among peers;

First, we observed grading bias in stage 1 of the peer assessment and introduced a data driven feedback mechanism to inform peers of it during their subsequent assessments. These guiding feedback were introduced in iteration 2 of the course. It was observed that students were more mindful and performed with better accuracy in assessing range of peer submissions.

Second, first iteration of the course used the traditional rubric interpreted to formulate guiding questions were used to assess submission range for stage1 of assessment method. We observed that the rubric lacked efficiency in translating intent of each dimension to student raters. Many students failed to understand academic vocabulary which was meant for staff and expert graders. In the second iteration of the course, the rubrics and guiding questions were simplified to be direct and uncomplicated. Further it broke complex guiding sentences with varied possible perspectives to separate ones to gather focused responses. Overall, 89% of students found it easier and quicker to assess submission range in the second iteration than the first.

Through the above experiments and related observations, we have come to understand the plentiful possibilities in the areas of peer assessment, formative feedback, and developing readable rubrics for design and other open-ended courses in online classes.

3 Accuracy of peer assessment

3.1 Stage 1-grade range

Results – range agreement between staff. Student submissions were randomly assigned to three staff members but from range of students in the top 10%, mid 10% and the bottom 10% established from earlier available data. All Staffs rated 50 submissions during each iteration of the course. It was recorded that the average disagreement

between staff raters on a submission for any dimension on the rubric was within 2 ranges on the grading scale.

Usually, differing judgments and interpretations contributed to grading differences among staff. Such differences were limited in both iterations due to a consensus-based mechanism [28] among all three staff in developing the rubrics for each assignment.

Results – range agreement between peer and staff. All assignments in this class were graded on a range scale of 5 points each (10 ranges) based on the rubric. It was recorded that in the first iteration of the course, 21.0% of submissions assessed fell within range on the grading scale while 49.0% within 2 ranges. The second iteration improved with 25.0% assessments falling within range, and 58.0% within the next 2 ranges. It was observed that the peer grade range improved over time thus exhibiting better sense of evaluation of peer work (see Figures 8 and 9).

It was observed that in the first iteration of the course, 16.0% of self-assessed submissions fell within their range on the grading scale while 70.0% on the next 2 ranges (Figure 5). The second iteration improved with 36.0% of self-assessed submissions falling within their range while 54.0% on the next 2 ranges (Table 1). It was also noticed that on average the self-assessed grades were within 2 ranges, but mostly on the higher side (Figure 6).

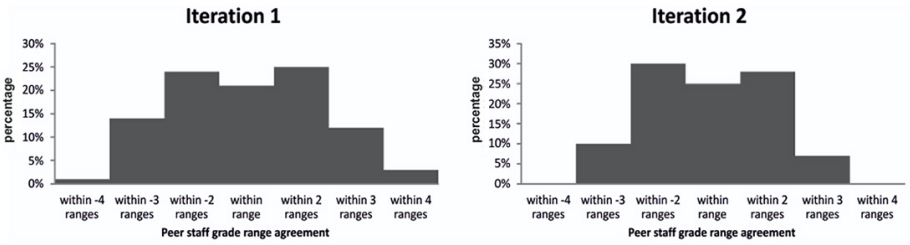

Fig. 5. Accuracy of grade range agreement of submissions by peers to that of staff in iteration 1 (left) and iteration 2 (right) for all five assignments in the course

Table 1. Summary of grade range agreement in iteration 1 and 2 between staff and peers

Peer staff grade range agreement	Iteration 1	Iteration 2
within -4 ranges	1%	0%
within -3 ranges	14%	10%
within -2 ranges	24%	30%
within range	21%	25%
within 2 ranges	25%	28%
within 3 ranges	12%	7%
within 4 ranges	3%	0%

Table 2. Rubrics used in iteration 1 of the study where students had several doubts on the meaning of the words and description used

	OUTSTANDING	EXCELLENT	SUPERB	GOOD	BETTER	ABOVE AVERAGE	AVERAGE	BELOW AVERAGE
	85-90 %	80-85%	75-80%	70-75%	65-70%	60-65%	55-60%	50-55%
Cognition Conceptual Studio	Work evidences **exceptional** breadth/depth of knowledge and understanding, strong and independent insight, critical awareness at forefront of discipline		Work evidences **good** breadth/depth of knowledge and understanding, reasonably independent insight and critical awareness at forefront of discipline		Work evidences **basic** demonstration of breadth/depth of knowledge with limited insight and critical awareness at forefront of discipline		Work demonstrates **little** breadth/depth of knowledge, critical awareness /insight at forefront of discipline	
Process Exploration Contextual thinking	Evidence of a wide range of strategic and manipulative understanding and resolution of appropriate inquiry with complex methods,mature attributes,evidences exceptional and mature attributes, competencies and practices commensurate with resource and time management, achievement and completion of deliverables.		Focused and innovative resolve of appropriate inquiry with complex methods for problem solution, practices commensurate with resource management, time management, achievement of plan of work, completion of deliverables.		Satisfactory evidence of appropriate inquiry with considerable degree of help required with complex methods and techniques for problem solution in unfamiliar, unstructured contexts,demostrates satisfictary attributes, competencieswith resource management, time management, achievement of plan of work, completion of deliverables.		Limited evidence of appropriate inquiry, able to proceed only if helped with complex methodologies and techniques for problem solution in unfamiliar, unstructured contexts. demonstrate basic attributes ,practices commensurate with resource management, time management, achievement of plan of work, completion of deliverables.	
Articulation Presentation Documentation	Exceptional ability to cogently and coherently communicate orally/ textually/ graphically makes advanced arguments, defends work done in professional manner.		Competent ability to make cogent and coherent communication in oral/ textual/ graphic terms, makes little advanced arguments and defends own work just about adequately.		Satisfactory ability to make coherent oral/ textual/ graphic communication, makes intermidiate arguments and , limited defense of own work.		Incoherent communication in oral/textual/graphic terms, basic arguments , limited defense of own work.	
Motivation Interactions Consistency	Exhibits an extremely concientious and spirited desire to learn and enhance the learning of others in the class.		Seems interested in learning, makes an above average effort to gain the most out of learning experience.		A willing participant in the project, and exhibits willingness to learn concepts and course material.		Shows litttle evidence of wanting to be in the to learn the material. The motive for the individual is somewhat questionalble.	

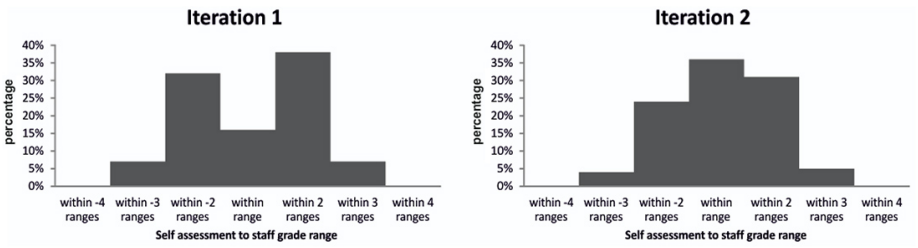

Fig. 6. Accuracy of grade range agreement of self-assessed submissions to that of staff in iteration 1 (left) and iteration 2 (right) for all five assignments in the course

Results – Time taken for peer assessment. Peers spent varied amount of time to assess different assignments. On average, 75% of peers took 8.6 minutes to 20.3 minutes to complete each assessment. The peer assessment system maintains anonymity where; raters don't see students' name while grading and students don't see raters' name while all scores given to them by different raters are seen. Providing time bound feedback to students is imperative as each assignment is built on its previous assignment. It took 4 days to complete the assessment process and share grades and feedback for each assignment giving students a good five days to act of the feedback. Some students who volunteered to assess more submissions (which usually are randomly chosen) were allowed to do so.

Providing feedback on grading accuracy. Thus far we discussed the efficiency of calibrated peer assessment in large online classes. Now we explore if, during stage 1 of AsPeer, accuracy in grading can be improved by providing feedback on a peers previous grading pattern. Earlier research has exhibited that feedback to crowd work boosts productivity [14]. From the third assignment of the course we experimented by

providing students with feedback on whether their grading on the previous assignment was "on range", "range lower" or "range higher" compared to staff grades as shown in Figure 7. It was noticed that 60% students who previously graded low, and 80% of students who graded high, graded on range compared to staff grades. Thus with this experiment we record that by providing students feedback on their grading pattern we could help improve accuracy.

Results – students reaction. AsPeer, peer assessment system received mostly positive responses from students. Fifteen percentage of students found the interface engaging and completed more assessments than required. Students found the activity enriching and believed their peers gained as much value as they gained by grading submissions. Following are some of the responses collated from the survey conducted at the end of stage 1 of the peer assessment system;

- 68% of students mentioned viewing other students' work as being beneficial.
- 43% reported learning new ways to communicate their ideas and improve presentation methods.
- Majority of students reflected that evaluating peers work was inspiring, motivating
- Most students were amazed to see alternate perspective explored to the same assignment, evoking curiosity.
- Students responded that self-assessment enabled them to relook at their own work and compare methods.

Overall, students reacted that the grade range method help them understand the rubrics and the intended outcomes of an assignment better, for a specific range (see Table 2). Students also reflected that the method though consumed more effort and time helped enrich their learning experience (see Figure 10).

3.2 Stage 2-comparative grading

Results – students reaction. It was largely reported that stage 2 of the assessment process was lot simpler and enjoyable. Comparing two students work side-by-side gave them an opportunity to identify shortcomings and opportunities for improvement in submissions. Students had reported that the two stage peer assessment method provided more confidence in the system though consumed more effort and time. The mean rating was 5.34 for confidence in two stage assessment method (6 point Likert scale, where 6: "highly reliable"). The students responded with a mean rating of 4.06 for time consumed and effort involved (6 point Likert scale, where 6: "highly worthy of time and effort"). In continuation students also reflected that stage 1 took more time and understanding of rubric than stage 2 of the assessment method. The mean rating was 4.34 for time and understanding of stage 1 of assessment method (6 point Likert scale, where 6: "extremely easy to understand and assess") whereas the mean was 5.12 for stage 2 measured on the same scale. Future work could focus on the design of effective interfaces that help peers discover, engage, and complete tasks swiftly.

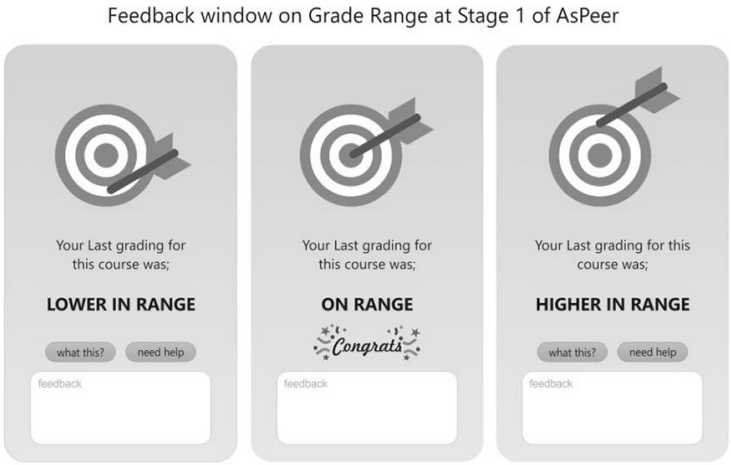

Fig. 7. Interface design of how students received feedback on their performance in their earlier assessment to grade range

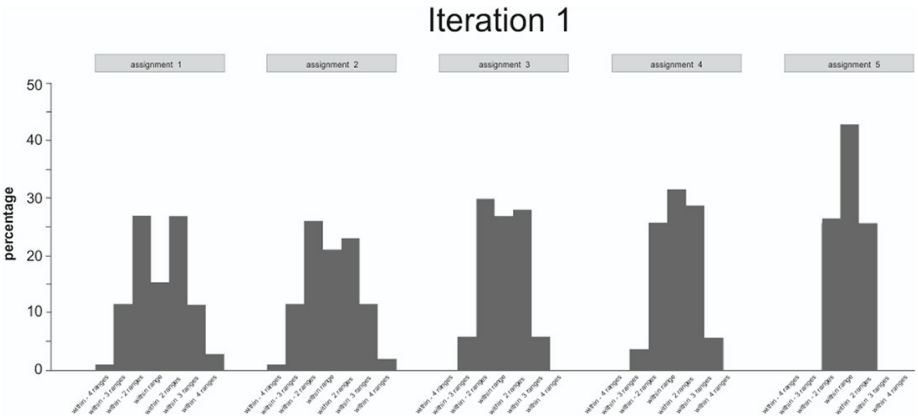

Fig. 8. Grade range agreement between peer and staff in all 5 assignments of iteration 1

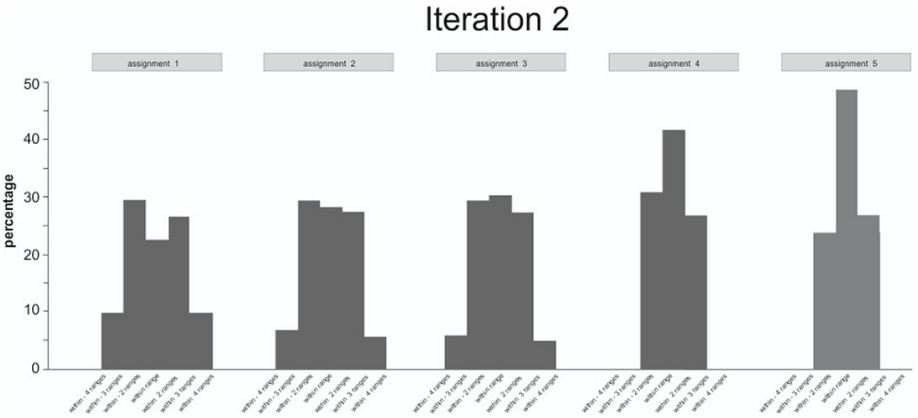

Fig. 9. Grade range agreement between peer and staff in all 5 assignments of iteration 2

Does peer assessment enhance critical thinking? Prompts that are pre-engineered encourage students to provide feedback beyond the snippet text. Three reasons why feedback snippets/prompts improve the standard of peer feedback are discussed here. First, giving peers with a list of potential prompts helps them examine and identify attributes. Moreover, snippets aides in converting peer effort in assessment from a task of recall and identification to a task of recognition. This encourages the act of giving feedback [36,37]. Second, providing a list of common, assignment-specific issues or opportunities that the submission could have essentially reduces inhibition and prompts peers to think critically [38]. Third, because feedback snippets mostly used terminology learned in class, they may trigger cued recall of these concepts [39] leading to more conceptual and actionable comments. This article demonstrates that insightful snippets improves peer feedback and marginally and contributes to making peers better thinkers.

Future works in this area can explore varied methods of providing peer feedback and also identify appropriate nodes during the process of working on the assignment rather than wait till the assessment stage.

4 Findings

4.1 Rubric design-driven by data

Iterations are imperative to design processes and often pays substantial dividends due to constant alignment to change [40] and the design of assessment systems are no exception. Data-driven analysis of teaching methods could enable tutors make improvements to lectures and other input materials in large online classes. To elaborate, we discuss some data-driven changes we made during the two iterations. We asked students and peers to rate the leading questions provided (reflecting varied dimensions of the rubric) during the assessment process in stage 1 (range fixing) as shown in Table 3. It was recorded that peers and staff found some questions as more relevant and helpful than others. We reviewed questions with low staff and peer rating and revised them in iteration two with feedback from the forum and weekly online discussions.

Most rubric revisions revolved around making rubrics more easily understandable and comprehendible.

4.2 Comprehension and word choice

The initial rubric used in the first iteration was similar to that shared with expert graders with an understanding of how they comprehend it. These rubrics sometimes used complex sentence structure and repetitive verb or adjective for multiple dimensions. This is not uncommon: some examples from previously developed rubrics reflect similar issues [41]. We hypothesized that to understand conceptual differences the rubrics should use a parallel sentence structure that would aide students better [42][43]. We recorded in the first iteration that rubric items with parallel sentence structure had lower disagreement scores; thus revising all rubrics to use parallel sentence structure. Also, when the initial rubric, were shared for peer grading, many students did not understand what a few words meant; like "trivial," "exceptional," "functional," and "critical". Most of these doubts were clarified on forum discussion and during the weekly live sessions. The revised version replaces such words with more specific ones to evoke pointed responses. In the second iteration of the class, a revised rubric was used with better readability. Overall, the agreement on rubrics between peer and staff was 3% higher than the first iteration.

4.3 Limitations of AsPeer – peer assessment method

Peer and self-assessment though has many advantages; it also has its share of limitations. The response from students on the efficacy of the peer assessment system is shown in Figure 10. First, staff and peers understand and evaluate work differently especially in open-ended creative courses. Though rubrics helps guide experts and novices reach consensus about creativity, their consensual judgments differ remarkably [44] due to their implied understanding of value [45]. Second, peer assessment imposes strict time lines for students due to logistical reasons in providing feedback to a class with a large number of students. Some students complained of not being able to complete assessments within the time given. Finally, AsPeer-peer and self-assessment method was found to be engaging and enriching for most students. Some students who weren't as involved or did not understand the rubrics lost motivation. Also from surveys, we have noticed that the students are generally satisfied with the overall grade but were unhappy with inaccurate and non-actionable feedback from peers. Addressing aspects for providing qualitative peer feedback remains future work.

5 Conclusion and future work

This article reports our experience with the use of AsPeer-peer assessment in two iterations of online class for university enrolled students of varied disciplines. We demonstrated that providing students' feedback about their rating bias, while assessing grade range in stage 1, improves subsequent accuracy. Some exciting opportunities for future work are discussed below;

First, the range fixing stage of the assessment process has peer graders not just have to assess if a submission is good to fit within a rubric range but in the first place to check if the submission is complete. Considerable time and effort from the student end goes into this check. Is it possible for students to check the completeness of the assignment before submission? Would students reveal the status truthfully? Future work could explore; introduction to submission templates and how word check could aide in the process. *Second*, the design of the rubrics contributes greatly in the range fixing stage of the peer assessment method discussed in this article. We identified that the traditional rubric used in iteration 1 was not clearly understood by peers; due to its vocabulary. With a varied forms of communication, is it possible to explore vocal and visual rubrics or a combination of them to improve peer understand? It might better aide students from different countries and languages also to contribute better. *Third*, once the mean grade range is fixed for each submission aided by a calibrated grading method using a small number of staff assessed submissions; each submission is graded along with 2 other submissions from the same range on a scale of 1–10; based on multiple comparative grading. This grading method is democratic and reflects the peers creative inclination to each submission and does not follow a rubric. Comparing two submissions side-by-side is not an easy task, unless there are some submission guidelines or templates introduced. Future work could explore how the interface could aide quick peer grading.

Several actionable feedbacks are provided by peers during their grading process. These feedbacks when collated provides for some healthy data to be analyzed. Algorithms could help us segregate feedbacks based on student performance and their grade ranges. Specific group of focused, repetitive but actionable feedbacks for each grade range identified can help understand how to trigger automated snippets even before the submission and during the process of the assignment. These feedbacks could be provided as pop-up prompt during the process of the assignment to trigger curiosity and better explorations. Future work can explore the efficacy of the in process peer feedback and how it can be provided in a timely manner.

These future works will transform design education in ways previously unimagined and unexplored.

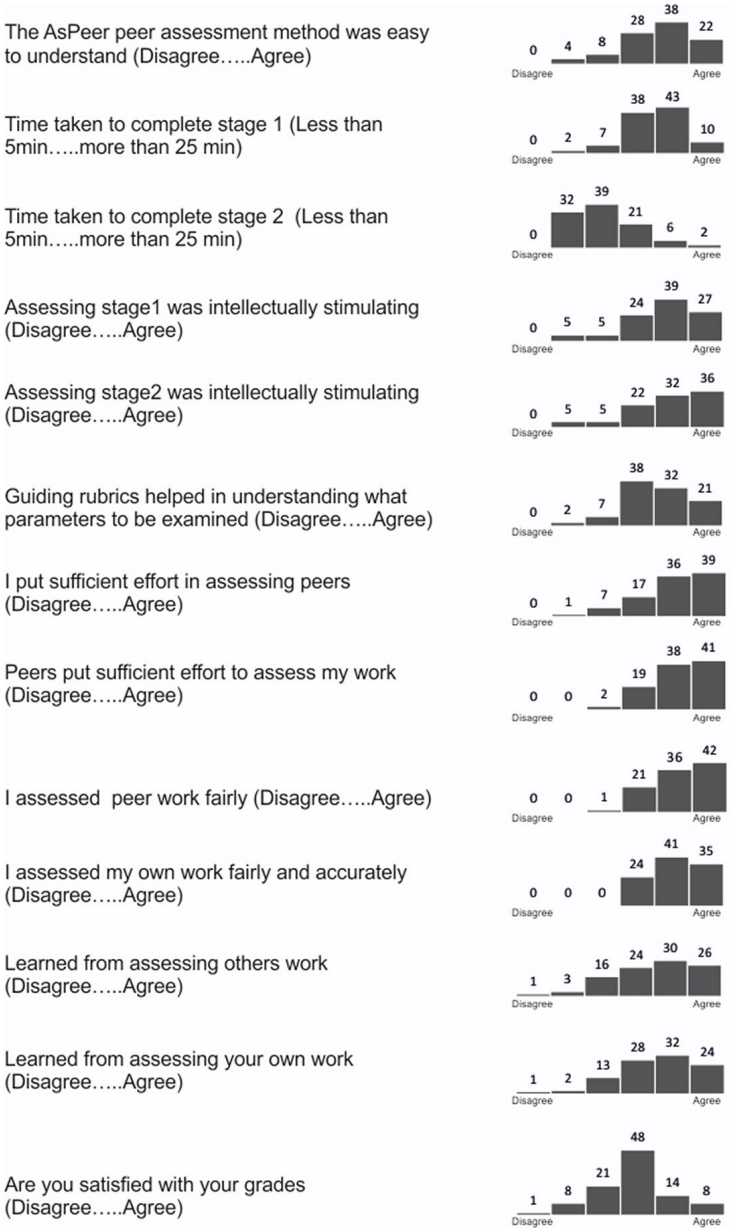

Fig. 10. End of course survey result of student's perception on the AsPeer peer assessment system

Table 3. Rubrics provided for the peer review during iteration 2 of the study showed better understanding by peers and also helped in constructive feedback

Metrics	Learning Outcomes	Outstanding 85-90%	Excellent 80-85%	Superb 75-80%	Good 70-75%	Better 65-70%	Above Average 60-65%	Average 55-60%	Below Average 50-55%	Just Pass 45-50%	Fail 40-45%
Ideation	KU 1: Explain research techniques, methodologies and discipline.										
	KU 2: Relate an individual design philosophy and approach to the development of design solutions.										
Research	KU 2: Integrate critical, analytical and creative skills in the origination and analysis of design concepts, proposals and solutions										
	KU 3: Integrate research and scholarship skills in the structure of a strategy for learning.										
	KU 4: Evaluate, select and apply appropriate structural, technical and material considerations within design development.										
	SQA 1: Identify areas of interest and create a map of future projects										
	SQA 2: Apply a synthesis of overall diagnostic, analytical, and creative abilities in research.										
	SQA 3: Independently identify, plan, sustain research into a chosen area of study.										
	SQA 4: Analyze problems and synthesize solutions, through the use of innovation, flexibility, adaptability, logical and lateral thinking.										
Execution	KU 1: Demonstrate knowledge of how to consider consumer behavior, cultural, social and technological contexts and differences in built environments appropriately in your research project.										
	KU 2: Create persuasive professional and business communication and presentations for projects/ internship/ research projects										
	KU 3: Manage own time and work to deadlines.										

6 Acknowledgment

This study is supported via funding from Prince Sattam Bin Abdulaziz University project number (PSAU/2023/R/1444). We thank the Deanship of Scientific Research, Prince Sattam Bin Abdulaziz University, Alkharj, Saudi Arabia for help and support.

7 References

[1] G. Martin. 2012. Will massive open online courses change how we teach? Communications of the ACM 55, 8 (2012), 26–28. https://doi.org/10.1145/2240236.2240246

[2] A. Stanley and M. E. Porter. 2002. Engaging large classes: Strategies and techniques for college faculty. ERIC.

[3] M. Tohidi, W. Buxton, R. Baecker, and A. Sellen. 2006. Getting the right design and the design right. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM, 1243–1252. https://doi.org/10.1145/1124772.1124960

[4] Schön. 1985. The design studio: An exploration of its traditions and potential. London: Royal Institute of British Architects (1985).
[5] S. Dow, A. Kulkarni, S. Klemmer, and B. Hartmann. 2012. Shepherding the crowd yields better work. In Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work. ACM, 1013–1022. https://doi.org/10.1145/2145204.2145355
[6] T. Gallien and J. Oomen-Early. 2008. Personalized versus collective instructor feedback in the online courseroom: Does type of feedback affect student satisfaction, academic performance and perceived connectedness with the instructor? International Journal on E-Learning 7, 3 (2008), 463–476. Retrieved February 3, 2022 from https://www.learntechlib.org/primary/p/23582/
[7] Venables and R. Summit. 2003. Enhancing scientific essay writing using peer assessment. Innovations in Education and Teaching International 40, 3 (2003), 281–290. https://doi.org/10.1080/1470329032000103816
[8] K. Topping. 1998. Peer assessment between students in colleges and universities. Review of Educational Research 68, 3 (1998), 249–276. https://doi.org/10.3102/00346543068003249
[9] B. J. Zimmerman and D. H. Schunk. (2001). Reflections on theories of self-regulated learning and academic achievement. In B. J. Zimmerman & D. H. Schunk (Eds.), Self-regulated learning and academic achievement: Theoretical perspectives (pp. 289–307). Lawrence Erlbaum Associates Publishers.
[10] S. W. Huang and W. T. Fu. 2013. Enhancing reliability using peer consistency evaluation in human computation. In Proceedings of ACM 2013 Conference on Computer Supported Collaborative Work. ACM. https://doi.org/10.1145/2441776.2441847
[11] Kulkarni, K. Papadopoulos, J. Cheng, D. Koller, and S. R. Klemmer. 2013. Peer and self-assessment in massive online classes. 20, 6 (2013). https://doi.org/10.1007/978-3-319-19641-1
[12] P. Pintrich and A. Zusho. 2007. Student motivation and self-regulated learning in the college classroom. In R. P. Perry and J. C. Smart, Eds. The Scholarship of Teaching and Learning in Higher Education: An Evidence-based Perspective. Springer, 731–810. https://doi.org/10.1007/1-4020-5742-3_16
[13] J. C. Kaufman, J. Baer, J. C. Cole, and J. D. Sexton. 2008. A comparison of expert and nonexpert raters using the consensual assessment technique. Creativity Research Journal 20, 2 (2008), 171–178. https://doi.org/10.1080/10400410802059929
[14] T. Winograd. 1990. What can we teach about human-computer interaction? (plenary address). In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM, 443–448. https://doi.org/10.1145/97243.97322
[15] S. Greenberg. 2009. Embedding a design studio course in a conventional computer science program. In Creativity and HCI: From Experience to Design in Education. Springer, 23–41. https://doi.org/10.1007/978-0-387-89022-7_3
[16] J. E. Tomayko. 1991. Teaching software development in a studio environment. ACM SIGCSE Bulletin 23, 1 (1991), 300–303. https://doi.org/10.1145/107004.107070
[17] Drexler, R. Chafee, and others. 1977. The architecture of the ecole des beaux-arts. MIT Press, Cambridge, MA. https://courseworks2.columbia.edu/files/596897/download?download_frd=1
[18] N. K. Farooqui, A. Almazyad, S. Ahmad. (2011). Knowledge management in education. 2011 International Conference on Information and Computer Applications (ICICA 2011). https://www.researchgate.net/publication/307888797_Knowledge_Management_in_Education
[19] Y. J. Reimer and S. A. Douglas. 2003. Teaching HCI design with the studio approach. Computer Science Education 13, 3 (2003), 191–205. https://doi.org/10.1076/csed.13.3.191.14945

[20] H. A. M. Abdeljaber and S. Ahmad. 2017. Program outcomes assessment method for multi-academic accreditation bodies: Computer science program as a case study. International Journal of Emerging Technologies in Learning (iJET) 12, 05 (2017), 23–35. https://doi.org/10.3991/ijet.v12i05.6410

[21] B. De La Harpe, J. F. Peterson, N. Frankham, R. Zehner, D. Neale, E. Musgrave, and R. McDermott. 2009. Assessment focus in studio: What is most prominent in architecture, art and design? International Journal of Art & Design Education 28, 1 (2009), 37–51. https://doi.org/10.1111/j.1476-8070.2009.01591.x

[22] Tinapple, L. Olson, and John Sadauskas. 2013. CritViz: Web-based software supporting peer critique in large creative classrooms. Bulletin of the IEEE Technical Committee on Learning Technology 15, 1 (2013), 29.

[23] Uluoglu. 2000. Design knowledge communicated in studio critiques. Design Studies 21, 1 (2000), 33–58. https://doi.org/10.1016/S0142-694X(99)00002-2

[24] K. Cennamo, S. A. Douglas, M. Vernon, C. Brandt, B. Scott, Y. Reimer, and M. McGrath. 2011. Promoting creativity in the computer science design studio. In Proceedings of the 42nd ACM Technical Symposium on Computer Science Education. ACM, 649–654. https://doi.org/10.1145/1953163.1953344

[25] D. P. Dannels and K. N. Martin. 2008. Critiquing critiques a genre analysis of feedback across novice to expert design studios. Journal of Business and Technical Communication 22, 2 (2008), 135–159. https://doi.org/10.1177/1050651907311923

[26] J. Forlizzi and K. Battarbee. 2004. Understanding experience in interactive systems. In Proceedings of the 5th Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques. ACM, 261–268. https://doi.org/10.1145/1013115.1013152

[27] Snodgrass and R. Coyne. 2006. Interpretation in architecture: Design as a Way of Thinking. Routledge.

[28] Amabile, T. M. (1982). Social psychology of creativity: A consensual assessment technique. Journal of Personality and Social Psychology 43, 5 (1982), 997–1013. https://doi.org/10.1037/0022-3514.43.5.997

[29] B. Feldman. 1994. Practical art criticism. Prentice Hall New York. Retrieved March 6 2022 from http://books.google.com/books?id=x6lPAAAAMAAJ

[30] Fallman. 2003. Design-oriented human-computer interaction. In Proceedings of the SIGCHI conference on Human factors in computing systems. ACM, 225–232. https://doi.org/10.1145/642611.642652

[31] S. Jha, S. Ahmad, H. A. Abdeljaber, A. A. Hamad, and M. B. Alazzam. 2021. A post COVID machine learning approach in teaching and learning methodology to alleviate drawbacks of the e-whiteboards. Journal of Applied Science and Engineering 25, 2 (2021), 285–294. https://doi.org/10.6180/jase.202204_25(2).0014

[32] S. P. Dow, A. Glassco, J. Kass, M. Schwarz, D. L. Schwartz, and S. R. Klemmer. 2010. Parallel prototyping leads to better design results, more divergence, and increased self-efficacy. ACM Transactions on Computer-Human Interaction 17, 4 (2010), 18. https://doi.org/10.1145/1879831.1879836

[33] M. A. Hearst. 2000. The debate on automated essay grading. Intelligent Systems and Their Applications, IEEE 15, 5 (2000), 22–37. https://doi.org/10.1109/5254.889104

[34] S. Hsi and A. M. Agogino. 1995. Scaffolding knowledge integration through designing multimedia case studies of engineering design. In Proceedings of the 1995 Frontiers in Education Conference. Vol. 2. IEEE, 4d1–1. https://doi.org/10.5555/1253523.1253768

[35] Han, "Using Calibrated Peer Review to Encourage Writing," in Proceedings of the 2012 Association of Small Computer Users in Education (ASCUE) Summer Conference, 2012, 23–32. Retrieved February 3, 2022 from https://ascue.org/wp-content/uploads/2014/11/2012-final.pdf

[36] J. Nielsen. 1993. Iterative user-interface design. Computer 26, 11 (1993), 32–41. https://doi.org/10.1109/2.241424

[37] J. Nielsen. 1994. Enhancing the explanatory power of usability heuristics. In Proceedings of the SIGCHI conference on Human Factors in Computing Systems. ACM 152–158. https://doi.org/10.1145/191666.191729

[38] D. Galinsky and G. B. Moskowitz. 2000. Counterfactuals as behavioral primes: Priming the simulation heuristic and consideration of alternatives. Journal of Experimental Social Psychology 36, 4 (2000), 384–409. https://doi.org/10.1006/jesp.1999.1409

[39] J. L. Little and E. L. Bjork. 2012. Pretesting with multiple-choice questions facilitates learning. In Proceedings of the Annual Meeting of the Cognitive Science Society. Retrieved March 26, 2022 from https://www.researchgate.net/profile/Jeri-Little/publication/265883438_Pretesting_with_Multiple-choice_Questions_Facilitates_Learning/links/5441da910cf2e6f0c0f66224/Pretesting-with-Multiple-choice-Questions-Facilitates-Learning.pdf

[40] Y.-A. Bachiri and H. Mouncif. 2023. Artificial intelligence system in aid of pedagogical engineering for knowledge assessment on MOOC platforms: Open EdX and Moodle. International Journal of Emerging Technologies in Learning (iJET) 18, 05, 144–160. https://doi.org/10.3991/ijet.v18i05.36589

[41] M. A. Haque, S. Haque, S. Zeba, K. Kumar, S. Ahmad, M. Rahman, S. Marisennayya, and L. Ahmed. 2023. Sustainable and efficient E-learning internet of things system through blockchain technology. E-Learning and Digital Media 0, 0 (2023). https://doi.org/10.1177/20427530231156711

[42] B. Markman and D. Gentner. 1993. Splitting the differences: A structural alignment view of similarity. Journal of Memory and Language 32 (1993), 517–517. https://doi.org/10.1006/jmla.1993.1027

[43] Sahni, J. (2023). Is learning analytics the future of online education? Assessing student engagement and academic performance in the online learning environment. International Journal of Emerging Technologies in Learning (iJET) 18, 02 (2023), 33–49. https://doi.org/10.3991/ijet.v18i02.32167

[44] C. Cheshire and J. Antin. 2008. The social psychological effects of feedback on the production of Internet information pools. Journal of Computer-Mediated Communication 13, 3 (2008), 705–727. https://doi.org/10.1111/j.1083-6101.2008.00416.x

[45] D. Chinn. 2005. Peer assessment in the algorithms course. ACM SIGCSE Bulletin 37, 3 (2005), 69–73. https://doi.org/10.1145/1151954.1067468

8 Authors

Hariesh K. Sankaran (0000-0001-5818-199X) is a research scholar at Karnavati University. He is an architect, product designer and urbanist presently working on domains concerned with the future of tech enabled online design education. His email is hariesh@gmail.com.

Raju Shanmugam (0000-0002-8613-9518) is a doctoral mentor and supervisor at Karnavati University. He is the dean of the School of Computational Intelligence at Karnavati University. His email is srajuhere@gmail.com.

Bhavesh Shah (0009-0003-3906-9048) is a doctoral mentor and supervisor at Karnavati University. He is a practicing architect and research mentor at the Karnavati School of Research. His email is bhaveshpriti@gmail.com.

Sultan Ahmad (0000-0002-3198-7974) is an IEEE member and currently working as Senior faculty in Department of Computer Science, College of Computer Engineering and Sciences, Prince Sattam Bin Abdulaziz University, Al-Kharj, Saudi Arabia. He has more than 15 years of teaching and research experience. He has around 75 accepted and published research papers and book chapters in reputed SCI, SCIE, ESCI and SCOPUS indexed journals and conferences. He has an Australian Patent in his name also. He has authored 4 books that are available on Amazon. His research area includes Intelligence computing, big data, machine learning and Internet of Things. He has presented his research papers in many national and international conferences. He is a Member of IEEE, IACSIT and Computer Society of India. His email is s.alisher@psau.edu.sa.

Hikmat A. M. Abdeljaber (0000-0001-9557-3933) is Asst. Professor in Department of Computer Science, Faculty of Information Technology, Applied Science Private University, Amman, Jordan. He was born in Kuwait, in 1967. He received the Ph.D. degree in information sciences and technology in 2010 from the Universiti Kebangsaan Malaysia, UKM, Malaysia. He has published papers in the area of information retrieval and artificial intelligence. His research interests include information retrieval, semantic web technology, data mining and machine learning. His email is h_abdeljaber@asu.edu.jo

Jabeen Nazeer (0000-0002-9242-6230) is lecturer in Department of Computer Science, College of Computer Engineering and Sciences, Prince Sattam Bin Abdulaziz University, Al- Kharj, Saudi Arabia. He has more than 20 years of teaching and research experience. Her email is j.hussain@psau.edu.sa.

Article submitted 2023-02-05. Resubmitted 2023-03-08. Final acceptance 2023-03-09. Final version published as submitted by the authors.