*International Journal of Engineering Education* Vol. 34, No. 4, pp. 1299–1313, 2018
Printed in Great Britain

0949-149X/91 $3.00+0.00


# Impact of Collaborative Team Peer Review on the Quality of Feedback in Engineering Design Projects*

MAHENDER MANDALA[1,4], CHRISTIAN SCHUNN[2], STEVEN DOW[3], MARY GOLDBERG[4], JON PEARLMAN[4], WILLIAM CLARK[5] and IRENE MENA[5]

[1] Department of Bioengineering, University of Pittsburgh, 3700 O'Hara St, Benedum Hall, Pittsburgh, PA 15261, USA.
E-mail: m.mandala@pitt.edu

[2] Learning Research and Development Center, University of Pittsburgh, LRDC Rm 821, 3939 O'Hara St, Pittsburgh, PA 15260, USA.
E-mail: schunn@pitt.edu

[3] Cognitive Science Department, University of California, San Diego, CSB Rm 131, 9500 Gilman Drive, La Jolla, CA 92093, USA.
E-mail: spdow@ucsd.edu

[4] Human Engineering Research Laboratories, University of Pittsburgh, 6425 Penn Ave Ste 400, Pittsburgh, PA 15206,USA.
E-mail: mgoldberg@pitt.edu, jpearlman@pitt.edu

[5] Mechanical Engineering and Material Science, University of Pittsburgh, 3700 O'Hara St, Benedum Hall, Pittsburgh, PA 15260, USA.
E-mail: wclark@pitt.edu, imena@pitt.edu

Increasing classroom sizes and decreasing financial and human resources have encouraged educators to seek innovative strategies to manage large classrooms. Several instructors have begun using web-based peer reviews as a way to increase open-ended feedback. Recent work in design-based classes has revealed that students struggle to provide meaningful peer feedback. Furthermore, it remains unclear how best to increase student motivation and engagement with the process. In a sophomore mechanical engineering class, we investigated the effect of a collaborative team of reviewers (a team of reviewers generating a single review) on the quality of feedback generated and on student perception of the process. Feedback generated by 117 students on their peers' design projects over two assignments was analyzed using a mixed-methods approach. We found that collaborative team of reviewers produced higher quality feedback than did individual reviewers. Students spent more time on reviews in teams but found the process engaging and more fun than with individual reviews. Furthermore, students perceived individual and team review tasks as requiring similar levels of effort. Our findings indicate that team review approach could help reviewers provide better feedback in engineering design reviews. Additionally, collaboration improved student engagement in the process. Over the past two decades, peer reviews have remained a solitary endeavor—this study is the first group process implementation of peer review and provides a basis for future exploration of the topic.

**Keywords:** peer review; team based learning; design; cooperative/collaborative learning; design evaluation

## 1. Introduction

A core element of design and design education is situated and frequent feedback [1, 2]. As design instructors deal with a faculty rewards system that does not incentivize high-effort teaching [3, 4], constraining budgets and widening class room sizes, a common casualty is feedback provision. It simply is not considered feasible to provide careful feedback to large numbers of design teams under an increasing temporal, human and financial resource pressure. Increasingly, web-based peer-to-peer feedback has emerged as an alternative to instructor feedback with a potential to scale well and keep in pace with growing class size, while creating newer avenues for student learning [2, 5–8].

However, relatively little is known about how to effectively structure peer review for design projects, increase student participation and engagement with the process, and more importantly how to mimic and maintain the natural learning environment of traditional design reviews and critiques [9]. Most saliently, the rich interactions of studio critique are not replicated in the current standard of web-based peer feedback, where individual reviewers typically provide feedback in isolation from their peers' opinions. Consequently, individual web-based reviewers miss out on opportunities to discuss with their peers their misunderstandings or support their technical limitations, and also learn from others' critiques as they concurrently evaluate the work. In addition, as equal status learners in class, individual student peers may not possess all the necessary skills or the design experience to effectively review open-ended creative problems that comprise most design-based learning classes.

Teams of reviewers working together could potentially overcome these problems: a team review could combine a range of technical skills and subject knowledge just as they do in achieving their own project design goals. However, teams are inherently plagued by group coordination costs [10] and prone to be biased towards polarized views of a few vocal members [11], which may be particularly

* Accepted 23 February 2018.

problematic for team reviewing. On the other hand, the simpler task of providing reviews may be easier to schedule than whole design projects. Biases may be less problematic when discussing outsiders work rather than their team's own work due to inherently lower investment in the task. The relative benefits of team reviews versus traditionally individual peer reviews has not been studied.

Responding to this gap in the literature, this article examines the merits and impact of review structuring utilizing individual and collaborative team of reviewers. In an introductory engineering design class, students designing a physical product conducted peer reviews of design logbooks and project videos under two review structures: one group conducted individual reviews while another completed collaborative team review. In the collaborative team review condition, students worked with their project teams in a collocated setting to review peers' work and collectively generate a single team review. Individual reviewers followed the current standard of review, with each student working independently on their assigned reviews.

Contributions of this article include detailing the peer review strategy of using a collaborative team of reviewers, providing empirical data that compares peer reviews conducted by collaborative teams and individuals, showcasing novel methodological approaches to study feedback using coding scheme to measure impact on performance and NASA TLX survey tool to measure peer review task related student effort, and implications on future computer technologies aimed at peer or crowd-sourced reviewing.

## 2. Background and related work

Research on formal peer review in classrooms has been conducted for more than three decades [12, 13], and the last decade has especially focused on web-based peer review, with its affordances of structure, ease of delivery, and anonymity. In these years, studies have explored the impact of peer review on student learning [14, 15], effectiveness of the peer feedback generated [6, 8, 16], and the student experience of participating in peer reviews [17–19]. Moreover, researchers have also compared students and instructors on their scoring and feedback [7, 20, 21]. These studies have shown that peer review is generally reliable, generates more feedback for the students, and has a beneficial impact on student learning for both reviewers and receivers of peer feedback.

The recent interest and growth in peer review research are largely associated with advancements in technology that have made facilitation of peer reviews far easier than before [22, 23] and have made them scalable to even large MOOCs with tens of thousands of students in one class [5]. Moreover, the call for improvement in assessment with increased inclusion of students in the process have further made the case for making peer review an integral part of the pedagogy [24, 25]. In the following sections we examine peer review within the context of design-based classrooms and conclude with the aims of our study.

### *2.1 Web-based peer review in the domain of design*

In design, peer review is not a novel or uncommon activity. Studio based peer and instructor critique have been central to design students' training for over a century. As a primary pedagogical tool in design, studios provide a natural multifaceted learning environment, where students not only develop their design, communication, and reflexive skills [26], but also socialize into the professional values, culture, and expectations of the field [27]. In addition to several benefits, studios serve a dual purpose of providing immediate situated feedback to the designer and supporting assessment of their work. Recently, studio-based pedagogy impacted creative fields outside the traditional arts and architecture [28, 29]. The peer interaction component of studio model was utilized by Hurst and colleagues, who implemented a structured face-to-face peer review in an engineering design classroom finding positive benefits on teams' progress, peer to peer idea sharing, and design communication [30].

The engaging environment of a studio and face-to-face peer feedback is sustainable in small class sizes, where such an interaction can be easily managed by the instructor. As class size increases, the facilitation of studio critique becomes a constraining factor in the process. In order to remedy the scale issue, researchers have looked for inspiration in the peer review research and tools developed in the fields of writing and computer science. Tinapple and colleagues developed and implemented a peer review tool for ''large creative classroom'', using peer-based public ranking of student projects in class and revealing the identities of the anonymous reviewers and authors at the end of review phase [2]. They found students socialized into a tighter community and supported each other's work when using this peer review tool. Similarly, Kulkarni and colleagues, scaled peer reviewing to a massive open online course involving design projects with thousands of students, and found rapid feedback, greater iteration, and improved grades [31].

### *2.2 Strategies for organizing the web-based review process*

Despite several benefits of peer feedback, there are several outstanding issues that require further examination by researchers. One primary issue is student engagement and participation in the process. Students are often apprehensive of peer feedback [18]. Their apprehension largely stems from ambiguity in feedback received [9] and perception of lack of expertise of reviewers [32]. These circumstances create a negative cycle where students provide less helpful or low-quality feedback, strengthening their notion that peer reviewers are unreliable, and thus reducing their engagement and participation in future peer review cycles. Researchers in the field of writing have grappled with improving student engagement in the process for over a decade. Studies have looked at instructing peers on providing feedback [33] including showing exemplar snippets of feedback [34], using a training module to calibrate their marking [23], using carefully crafted rubrics [35], or creating a course environment that is conducive to feedback [36].

Recent work in the domain of design found that nearly half of the freeform feedback from peers contained only praise or encouragement and lacked any suggestions for improvement or refinement [31]. This mimics similar outcomes in a study collecting peer feedback on engineering projects [37]. Further, design educators often dig deeper into design problems, but students often focus on pointing out communication problems in the documents [38].

In engineering design, Krause and Neeley [39], found written peer feedback gathered more information than simple verbal feedback interaction in face-to-face setting. Written design communication requires the designer to articulate their ideas and process in both a visual and descriptive fashion that enables the reviewer to form a coherent understanding of the designers' intentions. Additionally, designers and reviewers need to be well versed with the many languages of design used in its communication [25, 40]. We posit some of the ambiguity instilled in peer feedback in design stems from peers' lack of understanding of the designers' intent. Furthermore, engineering design crosses multiple domains of knowledge and skills, which at any given instance an individual peer reviewer may not fully possess. So how can one structure the reviews to engage reviewers more deeply into design issues?

### *2.3 Our study*

We look at the studio critique model for inspiration. An often-overlooked aspect of studio critique is the collaborative atmosphere of review generation. Peers do not review the work in a vacuum, working instead collaboratively with others in constructing their feedback. Such a collaborative team review process may be particularly beneficial in a web-based peer review set up, where peers often work with a passive design document and attempt to construct an understanding of the design intent. Zhu et al., [41] explored the use of a collaborative team of reviewers in an online crowd-sourced environment along with individual reviewers and an aggregate of individual reviewers. They found that such collaborative teams of reviewers working synchronously and collaboratively produced more useful feedback than individual reviewers, which aligned closely with expert feedback, and had increased internal consistency. Additionally, the aggregate feedback from individual reviewers outperformed the collaborative team of reviewers by a nominal margin. These results are promising in that they make the case for exploring collaborative reviewing strategies. However, it is not yet clear how these issues will tradeoff in a real course context, evaluating complex objects and also involving social issues that may be less prevalent in an anonymous online research study. Thus, we test the hypothesis that:

*H1: Collaborative teams of reviewers will generate better quality feedback (defined by accuracy of feedback and its impact on grade) than individual reviewers.*

However, collaboration in a team also has its limitations. In addition to increased coordination costs, teams may exhibit "group polarization," a phenomenon in which groups exhibit judgment closely resembling their individual biases rather than the "truth" [11]. These issues were detected in the work done by Zhu et al., (2014); however, the collaborative team reviewers nonetheless outperformed the individual reviewer in all measures. And although, collaboration may yield better quality feedback, if students find this collective process requires increased effort on their part, they may not fully accept or engage in the process. Thus, we test the hypothesis that:

*H2: Student perception of the effort required to generate feedback in collaborative teams will be significantly greater than feedback generation as an individual.*

Coordination costs in teams could be lowered but may not be completely eliminated. As such, if H1 holds true, in order to improve the quality of feedback, the increased effort could be an acceptable tradeoff for utilizing team reviews. Besides, the way the peer review process is implemented and utilized has been shown to significantly impact the way students both perceive the process and engage with it. For example, by improving the speed at which feedback was generated and received, Kulkarni et al. [31] found students to better engage in the process. Similarly, by creating a virtual community environment, Tinapple et al. [2] found students increasingly appreciated their peer feedback. In evaluating collaboration in peer reviews, a secondary goal of this article is to determine student perception of the process, which if negative, may impede adoption even if it is found to improve feedback quality.

## 3. Methods

The study described below was conducted in a classroom within the School of Engineering at a large public university, where web-based peer review was used in the past. The classroom allowed us to experience an authentic implementation of peer review process and its integration into the syllabus.

### *3.1 Course structure*

The study took place within a course titled Introduction to Mechanical Design, a sophomore-level introductory course on basic mechanical engineering design and product development process. The course consisted of several in-class lectures, computer-aided design labs and assignments, two team-based design projects carried out through the duration of the course, with no final examination. The two design projects were assigned 40% of the total grade in the class and were conducted in sequence. Peer reviews and the current research work were part of the first design project, named Design Project 1 that the teams worked on for the majority of the semester. Students elected their own teams, which were constrained to contain exactly five members (with a few exceptions of four-member teams). Students with no team preference were randomly assigned to instructor-generated teams. This team membership remained fixed for both the projects and through to the end of the course.

### *3.2 The design project*

Student-elected projects, which focused on new product development or improvement of a physical product, were vetted and approved by the instructors. These projects ranged from design and development of a novel single-handed bottle opener to re-designing of shaving blades for ease of cleaning (see Fig. 1).

Teams were required to document their work in a design log book, specifically including client statement, their hypotheses, initial user discovery, idea generation, preliminary designs, initial prototype, final design and communication. Teams were expected to ideate and design the product or system to a level where their designs could be readily fabricated, however, a physical prototype was optional. Teams participated in peer review by submitting a logbook after completing preliminary designs, and one additional time after completing final design.

### *3.3 Participants*

The peer review process was a class requirement for all 287 students. To incentivize responding to the survey questionnaires and being willing to be observed during collaborative peer review, students who participated in all aspects of the research were awarded 2 bonus points out of 100 over the final grade. All but 25 students agreed to participate in research (and received the 2-point bonus by completing instructor assigned alternative task). As a result, there were 58 project teams (with 4 or 5 students per team) randomly assigned to condition.

Students affiliated with the department of Mechanical Engineering & Material Science made up the largest major in class (79%), followed by students from Bioengineering (19%) and Electrical & Computer Engineering (2%). A majority of students were sophomores (61%) followed by juniors (21%), seniors (14%), and 5th year seniors (4%). A large portion of the class were male (78%), followed by female (22%).

### *3.4 Review structuring*

Collaborative peer reviews were performed by already existing project teams. While the assignment submission deadlines and content remained the same across all student teams, the collaborative teams were instructed to meet together, and discuss and generate a single peer review. These teams were given the option to reserve a multimedia room within the school of engineering as a meeting location for their reviewing tasks. Individual reviewers conducted the peer reviews independently.

All reviewers, individual and collaborative team as a whole, were assigned two randomly selected projects for review in each peer review assignment and used the same instructor developed reviewing rubrics. Furthermore, the whole review process was double blinded, with both the providers and receivers of feedback remaining anonymous throughout the process.

Each project received an average of 7.2 reviews. These reviews included at least one team review.

### *3.5 Peer review management*

Peerceptiv (Panther Learning Systems Inc., Pittsburgh, PA), also known as SWoRD, is a web-based peer review tool (Fig. 2) and largely used in the writing assessment field [42, 43]. A research version of SWoRD, which allows for increased customization and access to experimental features, was used in this class. Research personnel assisted in setting up,

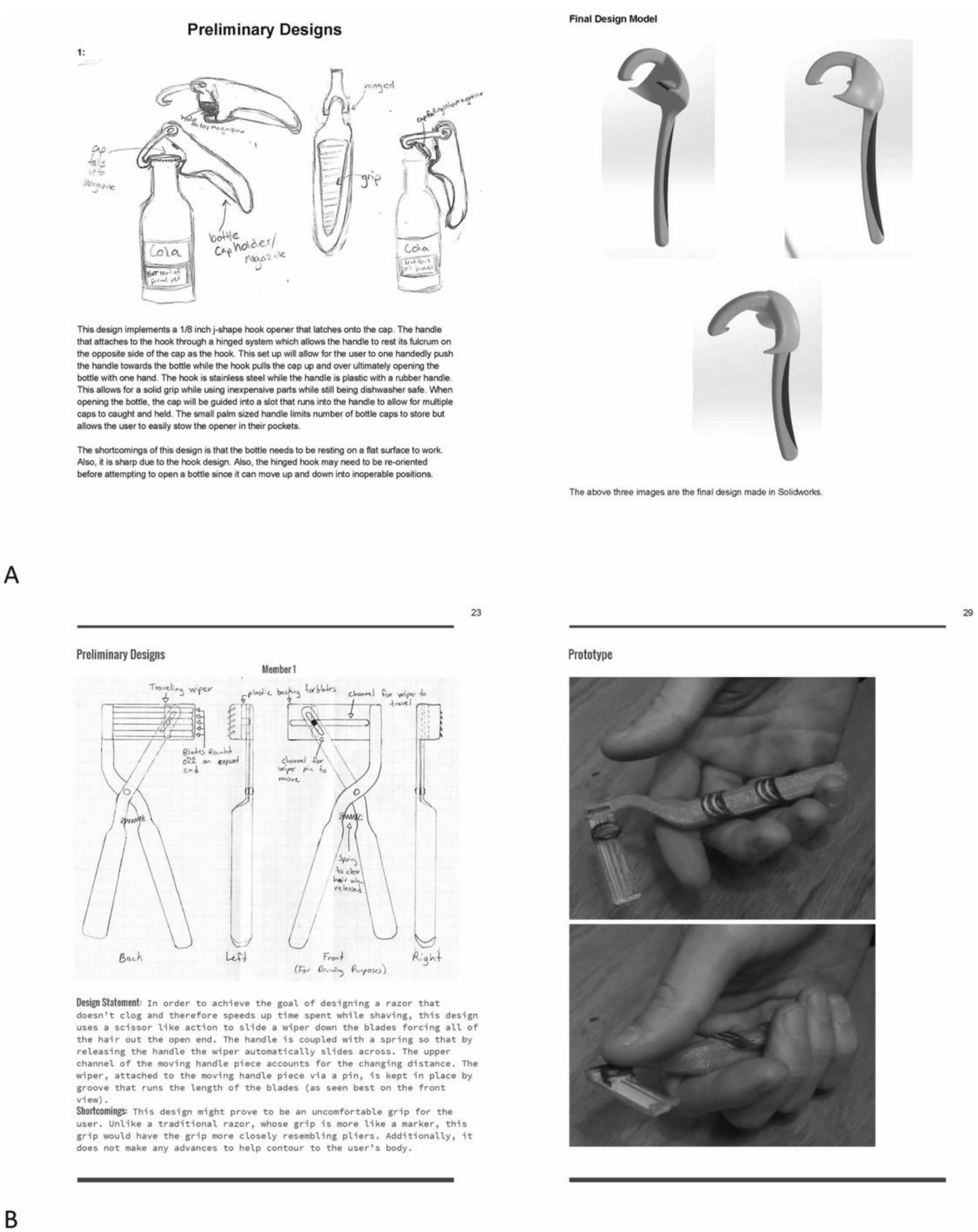


**Fig. 1.** Sample of the student projects from the classroom showcasing preliminary designs and final models for two different teams, A (top row) and B (bottom row).

managing and troubleshooting the system for the entire class.

In order to support collaborative team reviews, custom randomization code was implemented outside the system. The randomization code was written in MATLAB (The MathWorks, Inc., Natick, MA) and assigned the same two random projects to review to each member of the team if they were in a team review condition, while allocating the rest individually. When students conducted a collabora-

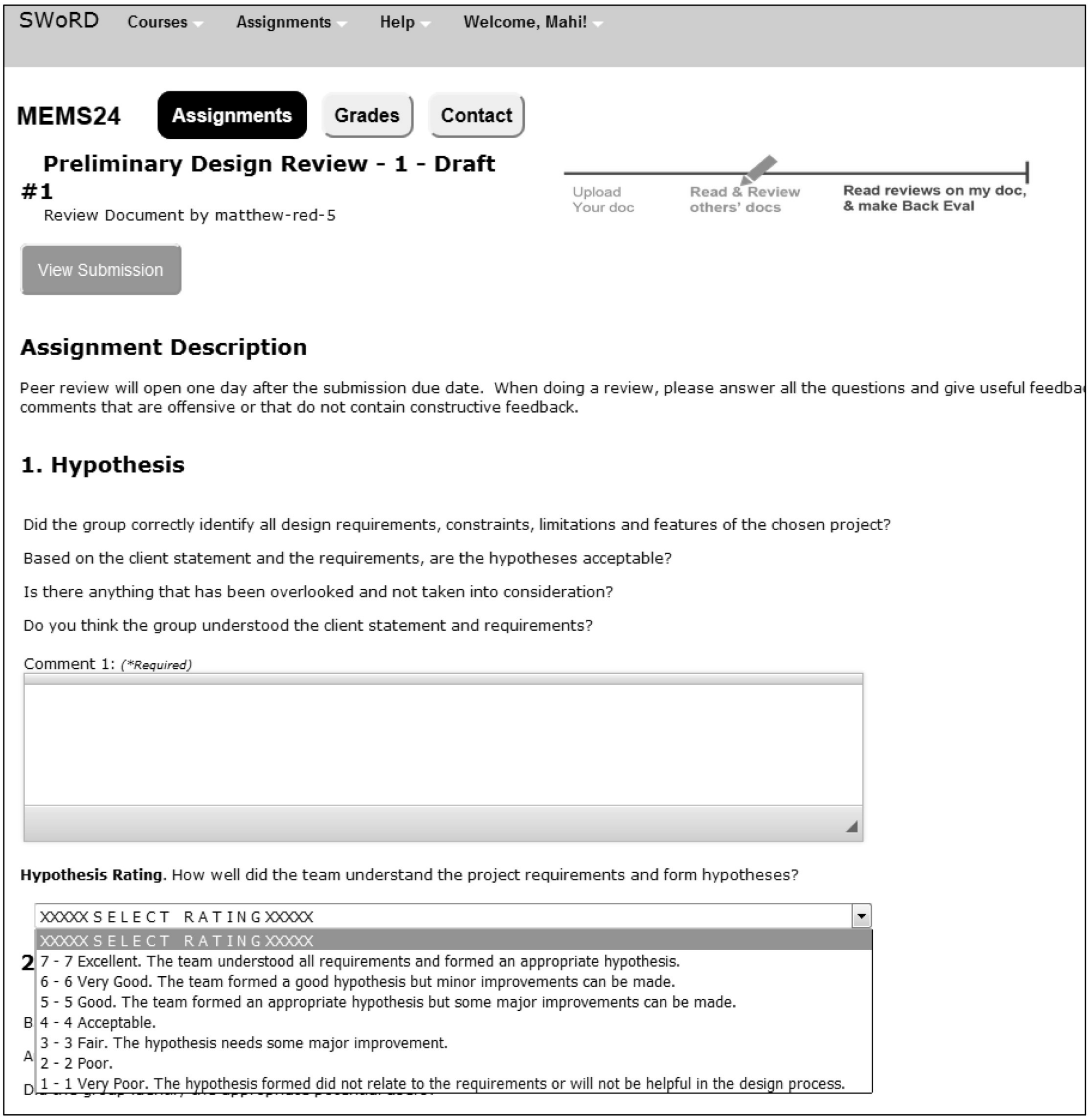


**Fig. 2.** The Peerceptiv/SWoRD user interface for reviewing documents (labelled by author pseudonym). Students enter freeform feedback under open text edit boxes (e.g., 'Comment 1'), with guidance provided for each reviewing dimension (e.g., Hypothesis dimension). Following each dimension's freeform comments, reviewers choose a rating from a dropdown menu using a 1–7 Likert rating with text anchors for each rating level that are specific to the dimension.

tive team review, each student in the team was given access to the same two projects to review, with any one member of the team providing the actual feedback and scoring.

### 3.6 Study design

Student teams were randomly grouped into either an individual review group (38 teams) or a team review group (20 teams). Individual review group member reviewed peer work independently, while members of the collaborative team review worked together on their reviews. Data collected from first peer review assignment was used to compare and contrast the two groups.

Participants were flipped between individual and team reviews for the second peer review assignment. However, due to increased noise generated by the crossing over (learning effect), lack of independence of data (group data crossover to individual data and vice-versa), and change in rubrics used, quantitative contrasts focused on the first peer review cycle. However, qualitative data contrasting the student experience within individual and team reviews included data from the second peer review cycle.

### *3.7 Dependent measure and data collection*

There were three primary sources of data: survey questionnaires, peer review feedback, and field notes from observations of teams conducting collaborative team reviews. Students conducting collaborative team reviews were asked to allow study personnel to passively observe their reviewing process. A convenience sample of 13 teams out of the 40 teams completing collaborative team reviews were observed (6 teams in the first peer review assignment, 7 teams in the second). No individual reviewers were observed. All observations were carried out by the primary author, who described the intent of the observations along with explicit statement on confidentiality of the record. No audio recordings or images were captured during these observations.

### *3.8 Feedback quality and sentiment*

Feedback quality was measured by accuracy and appropriateness of feedback. Two independent raters (authors 5 and 6, also the instructors of the class) rated the feedback on a gradient scale (see code book in Table 1) referencing the project logbooks and videos. Raters assigned a code based on whether feedback when implemented in the associated projects would yield a grade change. Feedback was rated per dimension (following the rubrics; assignment 1 rubric had 5 dimensions). A mean of these scores was used to reflect the overall quality of feedback per reviewer (be it an individual or a team). Additionally, proportion of high quality feedback (feedback with a score of +2 points) per reviewer was calculated.

Feedback sentiment was characterized as positive, negative or neutral and coded in a similar fashion as quality (Table 1) per dimension. A net sentiment score was then calculated and converted once again into an ordinal score reflecting the net sentiment per reviewer.

Instructor rating depended on whether feedback, if implemented, would improve the project grade (see Table 2 for an example of coded data). To reduce effects of noise from coding, we analyzed the data at the level of comment quality aggregated across dimensions. During the initial training, raters had a relatively high reliability in their ratings (Cronbach's $\alpha$ = 0.76). These values remained similar post training (Cronbach's $\alpha$ = 0.72). Disagreements between raters were resolved through in-person discussions moderated by the first author.

### *3.9 Time spent and effort required*

Survey questions inquired about the time and effort individuals spent to complete the reviews. Time was self-reported in units of minutes. Effort was calculated using the NASA TLX [44, 45], a multidimensional subjective workload assessment questionnaire of end-user workload on a given human-machine interaction. It uses six sub-scales: mental demands, physical demands, temporal demands, own performance, effort, and frustration" [44].

### *3.10 Statistical methods and analyses*

Group comparisons were conducted using *t*-test or Mann-Whitney *U* test for parametric and non-parametric data, respectively. Effect sizes are represented by standardized mean difference in the form of Cohen's *d*, and were appropriately adjusted for the type of analysis [46]. The measure of variability is reported with the mean in terms of either standard deviation (denoted by SD) or standard error (denoted by SE) as appropriate.

Descriptive analyses were conducted on quantitative survey data. Binomial tests were used to compare the observed frequencies of dichotomous variables, with the default probability parameter set at 0.5. All statistical tests were conducted using SPSS 23.0 (IBM Corp., Armonk, NY).

### *3.11 Analysis of field notes and open-ended survey items*

Field notes captured team behavior in the collaborative team review setting. Along with open ended responses from surveys, these data were

**Table 1.** Coding schema used to code the open-ended feedback

| Quality | Code | Score |
|---|---|---|
| If followed, would increases grade by a grade point or more. | 4 | +2 |
| Would improve the work but not by a whole grade point. | 3 | +1 |
| Would not impact the score. | 2 | 0 |
| Would negatively impact the score. | 1 | –1 |
| **Sentiment** | **Code** | **Score** |
| Positive | A | +1 |
| Neutral | B | 0 |
| Negative | C | –1 |

Note. Rater assigned 'code' which was converted to corresponding 'score' for analyses.

**Table 2.** Exemplar peer feedback, coded by raters

| **Sample raw feedback text from a peer reviewer** | **Rater assigned codes** |
|---|---|
| Yes, the group asked appropriate questions. However, they asked three questions total and for 2 out of the three questions they only had 3 answers! This is not nearly enough intital [sic] user discovery to get an accurate depiction of what the public could want. The questions, although they get the job done, are very basic and are asked in a biased way. The group never asked people about a positive experience they've [sic] had with bottle openers, so the group doesnt [sic] know what these potential users finds positive at all about bottle openers. The group identified the correct potential users but I am unsure of how many people they actually interviewed. They only identified three different users, but for the first question there are 10 answers, and for the next two there are only three answers. From personal experience, I find it very hard to believe that everyone had the same answers for the last two questions! Next time I would like to see more questions being asked to a larger group of potential users. | 4C |
| The questions are clear and well-thought out as noted above. Overall, the group identified a lot of student users (which was one of their main potential consumers), but it may have been more useful to also talk to people in other sectors (such as travelers and office workers noted in the client statement). They may have provided even more useful information from a different perspective than a student. | 3A |
| The hypothesis identifys [sic] all the main design requirements that must be included in theri [sic] design. I think the hypothesis could be improved if they include more of the design constaints [sic] (such as being able to hold 30lbs or be compact). However, I feel that overall they understood the client statement and that it is an acceptable hypothesis. | 2A |
| The names of group members should not be in the logbook. I feel the expected users should exclude professional players, because usually pro players have a good grip and don't need this product. The hypotheses are overall acceptable. | 1B |

Note. Numerical score represents quality rating, letter score represents overall sentiment score.

analyzed using an inductive framework that groups information into themes and consequently assists in drawing conclusions [47]. All field notes and survey data were coded by primary author.

## 4. Results

To account for missing data, only students who completed all three surveys (pre-course, post-assignment 1, and post-assignment 2), N = 117 (41% of class total), were selected for review quality coding and for data analyses. The selected sample had a similar grade make up (M = 94, SD = 4) as the overall class population (M = 92, SD = 7). These 117 students represented 50 unique teams, and in many instances not all members of the team were present in the dataset. Furthermore, in assignment 1, the focus of the quantitative analyses, 39 students participated in team reviews (representing 16 teams), while 78 (including 2 missing) students participated in individual review. In assignment 2, 41 students participated in team reviews (representing 15 teams), while 76 (including 4 missing) students participated in individual review.

Note: To help distinguish sample size representing teams from individual, we use the following notation for sample size representing number of teams: $N_t$.

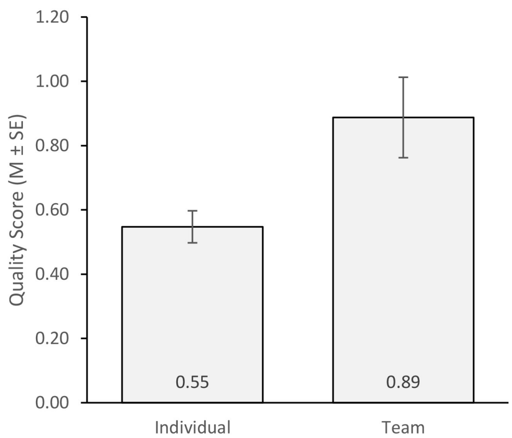


**Fig. 3.** The mean quality scores (and SE bars) of feedback generated by individual and collaborative team reviewers. Team reviews resulted in significantly higher (p = 0.015) mean quality score compared to individual reviews.

### *4.1 Collaborative team reviewers produced better quality feedback than individual reviewers*

Reviewers in a collaborative team (See Fig. 3) generated significantly higher quality feedback ($N_t$ = 16, M = 0.89, SD = 0.5, proportion of high quality feedback per reviewer = 24%) compared to individual reviewers (N = 76, M = 0.55, SD = 0.43, proportion of high quality feedback per reviewer = 13%), $U$ = 374.00, $z$ = –2.44, $p$ = 0.015, Cohen's $d$ = 0.53.

### *4.2 Individual reviewers were more positive in their reviews than collaborative team of reviewers*

Individual reviewers (See Fig. 4) generated slightly more positive feedback (67% positive, 7% negative)

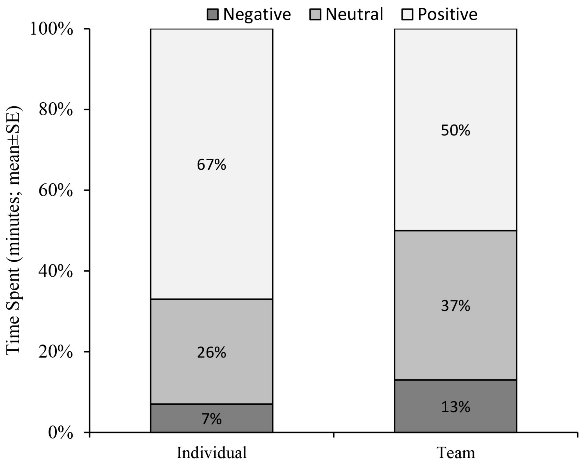


**Fig. 4.** Mean feedback sentiment per reviewer type. Individuals were generally more positive.

compared to collaborative team reviewers (50% positive, 13% negative). However, this difference was not statistically significant, $U$ = 504.00, $z$ = –1.44, $p$ = 0.151, Cohen's $d$ = 0.30.

### 4.3 Collaborative team reviewers spent more time than individual reviewers on the reviewing tasks

Students in the collaborative review teams reportedly spent significantly more time (N = 39, M = 116, SD = 50 min) compared to individual reviewers (N = 73, M = 93, SD = 46 min), $U$ = 1035.500, $z$ = –2.392, $p$ = 0.017, Cohen's $d$ = 0.46. See Fig. 5.

### 4.4 Students perceived similar task effort needed to complete reviews in both individual and collaborative team reviews.

Contrary to hypothesis H2, students perceived similar task effort required in completing both the individual and collaborative team-based reviews.

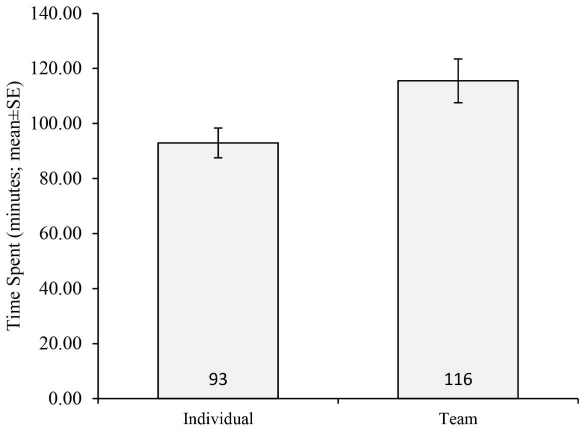


**Fig. 5.** The mean time spent (and SE Bars) on generating feedback by individual and collaborative team reviewers. Team reviewers spent significantly (p = 0.017) more time on reviews than individual reviewers.

There were no statistical differences detected in the perceived effort required to complete the reviews as reported by students in individual review group (TLX, M = 57%, SD = 12%), and collaborative team review group (TLX, M = 54%, SD = 12%), $t(113) = 1.05$, $p = 0.296$.

### 4.5 Students preferred collaborative team reviewers and found them engaging and beneficial

When students were asked to contrast the experience, they had conducting reviews as a team and as individuals, contrary to our expectations, students were split on which structure made the review easier to complete (39% agreed individual review were easy, while 38% disagreed, N = 68; see Fig. 6). However, a majority perceived that the feedback generated from team reviews was of better quality (70% agree, 12% disagree) than that from individual reviews. Student opinion expressed in open-ended survey questions, captured this sentiment and the logistical challenges insightfully. For example, comments like, "*The only thing I didn't like about peer reviewing together was having to get the group in all one spot* [sic]*, at one time to review. It was logistically challenging*", and, "*It took a longer time to do the team review than the individual review, and it was also harder to pick a time to do it since we all needed to be together in the same room . . .*" highlight the reasons for the increased effort required to complete team reviews. On the other hand, responses such as, "*Team was better to see other's points of view that you may not have thought of on your own. It also incorporated more perspectives to provide better overall feedback.*", and "*. . . it* [team review] *was helpful to ask questions about the reports intentions and it was more enjoyable*" demonstrate why students perceived team reviews generated higher quality feedback. Students largely agreed that providing feedback as part of a collaborative team was fun (66% agree, 11% disagree; Fig. 6), as is well described by student comments, "*Team reviews were more fun because you could socialize in between making reviews . . .*", "*. . . I do think we did a better job overall reviewing as a group. It was definitely more fun.*" Observations of teams in the process of the reviewing supported the assertion that team reviews were fun and engaging. As teams continued working on the reviews, they spent more time than they allotted for the review, resulting in some members of the team having to skip a part of the review to make it to their next appointment. In several such instances, students seemed reluctant to leave and miss out on the "fun."

Another notable result was that students perceived providing feedback as a team made them better reviewers (57% agree, 18% disagree, Fig. 6). Student comments provide a rationale for their

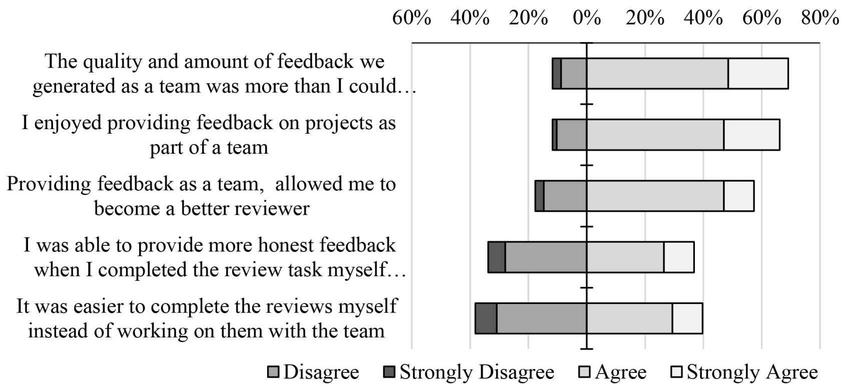


**Fig. 6.** Students' opinion contrasting their experiences in collaborative team reviews and individual reviews.

perception. For example, one student reported, *"Reviewing as a team allowed me to consider other aspects that my teammates brought up that I didn't necessarily think of"*, while another reported, *"I like working as a team. We can bounce ideas off each other, they can find things that I couldn't find that was missing in others work, helped me learn about how to be a better reviewer."* Furthermore, team observations captured that teams spent a significant amount of the review time on determining appropriateness and accuracy of work done, and in explaining the projects or aspects of the projects to each other. They used their own team knowledge to construct an understanding of the projects, with some teams using web-searches, or textbook and class notes to supplement their analysis.

Overall, students seemed to prefer collaborative team reviews over individual reviews over a variety of reasons. In students' view, *"Working as a team was better because we could all collaborate on the responses instead of just having one point of view."*, and, *"Team reviewing is a more genuine review approach. It adds a degree of accountability not achievable with solo* [individual] *reviewing."*

A few students note the issues that arise when working in a team, e.g., "*Team review takes way too long. Too much arguing/making decisions on minor details. Individual review takes significantly less time (one hour individually vs 2.5 hours as a group)"*, and, "*We stressfully made time and arranged an inconvenient meeting to do the team review . . . the same results would've come from the individual review and in a quicker and more efficient manner."*

Finally, the survey enquired student preference between individual and collaborative team review. They were asked to base their choice on their own personal learning from completing the review, their anticipated performance in the review task, and effort required to complete the review. In all three cases, students overwhelmingly picked the collaborative team reviews (see Fig. 7). The percentage of students who picked between the two choices (team or individual) were significantly different from 50% for performance (binomial test $p = 0.002$, $N = 68$) and effort (binomial test $p = 0.005$, $N = 68$). The results trended towards students picking collaborative team over individual review based on their anticipated learning from the task, however, this result was not significantly different from the random probability of picking either set at 50% (binomial test $p = 0.114$, $N = 68$). While these results support the aforementioned student opinion, one surprising outcome here relates to the students' choice based on effort required to complete the review, as seen in Fig. 7. Despite the fact that students noted the extra time and coordination needed to complete the collaborative team review, they preferred it over individual reviewing (68% team review, 32% individual review).

Note: The sample size ($N = 68$) above represents

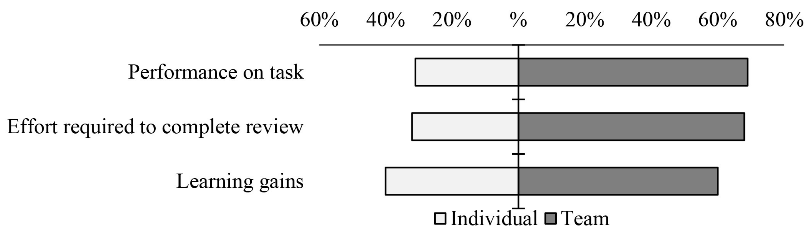


**Fig. 7.** Poll showing the type of review (individual or collaborative team) students would like to be part of based on their performance, effort needed of them, and their anticipated learning.

students who completed reviews as individuals and teams across the two assignments, excluding students who only participated in individual reviews. The comparison of review structure could only be possible in the crossover sequences.

## 5. Discussion

In this study, we explored the benefit of structuring peer reviews to include a collaborative team review element in them. We found that collaborative teams of reviewers generally produced higher quality feedback compared to feedback generated by individuals. Furthermore, individual reviewers seemed slightly more positive than collaborative team reviewers in their reviews. Surprisingly, there were no differences in student's perceived effort needed to complete tasks across individual and collaborative team review conditions, although objectively they spent more time in the team review condition. It is likely that the perception of effort depends on the motivation students have for the tasks rather than on the actual number of hours invested in the task [48]. And combined with the fact that students found collaborative teams to be more engaging and fun, it could explain the lack of additional effort students felt while spending more time and expending coordination costs on collaborative team reviews.

### *5.1 Why did collaborative teams generate higher quality feedback?*

We offer two explanations for why collective generation of feedback as a team yielded improved feedback quality. First, students in collaborative teams were able to clarify their understanding of the project, the assessment requirements, and expectations of their role as reviewers, resulting in increased accuracy of the feedback. Second, collaborative team reviewers seemed engaged, found peer review fun, and spent more time than individual reviewers, further enhancing their feedback generations efforts.

Individual reviewers on the other hand, had to independently form an understanding of all aspects of the review process and the design intent of their peers, and given their novice status in the field of design, it might have made reviewing these open-ended problems accurately, a challenge. Nevertheless, individual reviews cannot be completely dismissed. Because collaborative teams are sometimes plagued with issues such as group production losses and potential "group polarization", individual reviewers can examine a larger quantity of projects (i.e., three individual reviewers, each reviewing two random projects would generate six reviews, to generate the same amount of reviews, a team of three individuals would need to review six projects). The current study did not examine the effects of production losses as a result of working in a team. However, it found that the perceived effort was similar in both collaborative team and individual review condition, i.e., students found both activities similarly taxing.

As another concern, an aggregate of individual reviews could still yield better feedback [41, 49]. However, this aggregation may not work well when individual reviewers do not accurately understand the project at hand or when the project evokes contrasting views. Assuming that collaborative team of reviewers produce higher quality feedback, it may be that an aggregate feedback compiled from a combination of individual and collaborative team feedback will further improve overall feedback quality.

### *5.2 Do collaborative team reviews impact student learning?*

This study did not explicitly measure the learning impact of the peer review structure. However, students reported that they became better reviewers after conducting a collaborative review, in addition to electing collaborative reviews on the basis of increased learning gains from participation. Furthermore, observation of the collaborative team reviewers revealed a cooperative learning atmosphere, where students freely exchanged their thoughts and ideas about their peers' work, while also engaging in self-assessment. In certain cases, teams worked out the design problem of the projects under review, framing the problem, scoping out issues and ideas, and considering what the outcomes would be if they were to work on the problem. Such a level of active engagement with their peers' projects may provide additional design experiences to the team, and since the review teams were also project teams, it creates opportunities for the teams to work together on a new task, socialize, and collectively understand the assessment used in the class.

Furthermore, it could be inferred from team observations that some teams delegated work within their project and only had a higher-level understanding of the activities their teammates carried out. As teams reviewed their peers' work and assessed their own work, many within the team gained a deeper perspective of different tasks their team members completed and their impact on assessment. This level of team assessment may perhaps help students recognize the positive interdependence that should exist within the team, and improve their collective efficacy (a team's belief about its own capabilities to work together; see Lent, Schmidt [50]). In addition to learning, the quality of feedback generated has been shown to exert a significant positive influence on reviewers' own performance on subsequent iteration of the assignment [51] as well as impact student perception of the peer review process [18]. Nonetheless, formally addressing the learning component of collaborative team peer review remains future work.

### 5.3 Why did students prefer collaborative team reviews?

We expected collaborative team reviews to increase student burden, and some of the results indicate this to be true. However, we also found that the reviewing process in collaborative teams was considered more engaging and fun. It is possible that although students spent more time and effort on the review, the increased engagement and association of fun with the task, may have skewed their preference to collaborative team reviews. Additionally, students in collaborative team reviews were able to share their understanding of the work under review and their role as reviewers, share their frustration and excitement with their teammates, and view multiple perspectives on the same work, most of which they were unable to do when conducting the review independently. Furthermore, reviewing together in project teams, allowed team members to socialize under reduced pressure, on a non-critical non-project task. This reduced pressure on the team may have negated any experience the teams had with free riders who are typically encountered in teams. Students also seemed to recognize the impact collaborative team review had on their learning, which they opined also made them better reviewers. Therefore, on the whole, given the contrast students picked working together as a team on these reviews.

An interesting observation from collaborative team reviews was its impact on reviewer accountability. We posit that the social aspect of reviewing in a team creates a new expectation that each member of the team should carefully consider and fully participate in the review discussion and feedback generation. This is perhaps an important outcome of using collaborative team reviews, in that, it ensures increased participation in the peer review process across the class undermining the digital distancing that is often created in an anonymous review process.

### 5.4 Generalizability and impact beyond engineering design

Results from this study indicate that collaborative team reviews have the potential to improve student engagement and motivation, and consequently yield better quality feedback. The current work focused on using peer reviews in the context of design-based classrooms. However, there are significant implications of the work that have the potential to impact the use of peer reviews beyond the current context. Design poses complex challenges, mixing fact-based domain knowledge and creative conceptualization that necessitates the use of diverse perspectives and collaboration in teams. It seems a natural fit to the use of collaborative team of reviewers in any peer review process utilized in evaluating design. Contemporary web-based peer review processes, in general, often end up adding complexities to any review task by reducing communication to a passive written document, using confusing or unclear assessment criteria, and using hesitant students attempting to unravel the expectations of the process and their role. Therefore, a collaborative review approach could enhance the review experience in other domains as well, especially when reviewers have little to no experience in reviewing. Early research work by Zhu and colleagues [41], in which a virtual crowd of reviewers working together performed better than individual reviewers on several reviewing tasks including mathematics problems and writing, adds support to this hypothesis.

### 5.5 Limitations

The study was conducted in an authentic peer review implementation within a large engineering class, and the contextual details could have influenced the obtained results. For example, properly devised rubrics play an important role in helping novice reviewers create meaningful feedback [35].

The quality of projects was not controlled in this study, and thus some variation in feedback that projects received may have influenced their quality rating. The large sample of reviewed projects should potentially balance out the quality of projects reviewed.

In the current study, only collaborative team reviews were observed. With no corresponding data from individual reviewers, there were fewer opportunities for contrast.

Finally, the role of study personnel in data collection and analysis could induce potential bias in reporting the results in this study. Primary author, moderated rater disagreements when rating the feedback comments on quality and sentiment, conducted and analyzed passive observations of collaborative teams reviewing, and analyzed all survey data. Nevertheless, care has been taken to systematically analyze data and report the results in an objective fashion.

### 5.6 Future work

This study builds a basis for future work to explore strategies that can improve the peer review processes and outcomes. There exist several directions for future work. To begin with, the temporal effect of team reviews on individual reviewing capabilities need to be thoroughly examined. This examination could help educators and researchers determine the applicability of team reviews within certain contexts (help train novice reviewers, utilize only in the first instance in a class etc.).

Recent work by Carberry and colleagues [52] found adding verbal feedback gathered more information than simple written feedback in a web environment. Future work, could combine and examine the use of verbal feedback and written feedback in a collaborative setting to further improve the quality of feedback generated and increase student engagement.

The coding methodology used in this study for quantifying the quality of feedback focused on only one aspect of the feedback: accuracy in terms of predicted impact. Although this approach focused on a critical aspect of feedback, there exists several other criteria that could be assigned to feedback, to create a thorough understanding of the various ways feedback generated by a collaborative team and individual reviewer differs. Dannels and Martin [53] explored the typology (categorization) of feedback in studio critique ranging from novice to experts. This work suggests that feedback, at least in the studio setting, is composed of following types in decreasing order of frequency, "*judgement, process oriented, brainstorming, interpretation, direct recommendation, investigation, free association, comparison, and identity invoking.*" Thus, future work could evaluate feedback based on design-oriented feedback typology.

A minority who preferred individual reviews, brought out some of the issues that could plague collaborative team review. The most notable issue was the potential for teams to engage in prolonged discussion over minor details in the project under review. Although the increased time on task could benefit the team and its members in developing a better understanding of the assessment and their roles both as a reviewer and as members of the team, it can also distract reviewers from providing crucial feedback. Future work could examine aspects of the assignment where students engage in meaningful discussions and build on developing guides and rubrics specific to collaborative team reviews. Such initiatives could further improve the feedback generated.

As mentioned earlier, future work could quantify the benefits of collaborative team peer review structure on the reviewers' learning, motivation, and performance. In the current work, project teams were utilized to conduct the collaborative team reviews as well, subsequent research could look at the benefits of such structuring on the team and explore the use of review-only teams that are created independent of the project teams.

## 6. Conclusion

In this study, we investigated the impact of structuring peer reviews using a team of reviewers instead of individual reviewers on the quality of feedback generated and student perception of effort required. In the context of a large engineering design classroom, we found that collaborative teams have the potential to generate better feedback, and students find them no different than individual reviews in the effort required to accomplish the reviewing task. Furthermore, we found that students preferred conducting the reviews in collaborative teams because the process was more engaging as well as perceived to be more beneficial to their learning. The positive effects observed in this classroom study provides a detailed, concrete example of an approach to peer review that can be easily implemented by engineering instructors and further studied by engineering education researchers in the future research.

*Acknowledgments*—The authors would like to thank the Human Engineering Research Laboratories, Pittsburgh, for providing research computers for data analyses, and the Swanson School of Engineering, University of Pittsburgh, Pittsburgh, for research space to conduct research observations.

## References

1. S. Fitch, Art Critiques: A Guide, *Studies in Art Education*, **57**(2), 2016, pp. 185–187.
2. D. Tinapple, L. Olson and J. Sadauskas, CritViz: Web-based software supporting peer critique in large creative classrooms, *Bulletin of the IEEE Technical Committee on Learning Technology*, **15**(1), 2013, p. 29.
3. R. H. Todd and S. P. Magleby, Evaluation and rewards for faculty involved in engineering design education, *International Journal of Engineering Education*, **20**(3), 2004, pp. 333–340.
4. C. L. Dym, J. W. Wesner and L. Winner, Social dimensions of engineering design: Observations from Mudd Design Workshop III, *Journal of Engineering Education*, **92**(1), 2003, pp. 105–107.
5. C. Kulkarni, K. P. Wei, H. Le, D. Chia, K. Papadopoulos, J. Cheng, D. Koller and S. R. Klemmer, Peer and self assessment in massive online classes, *ACM Transactions on Computer-Human Interaction (TOCHI)*, **9**(4), 2013, pp. 131–168.
6. K. Topping, Peer assessment between students in colleges and universities, *Review of Educational Research*, **68**(3), 1998, pp. 249–276.
7. N. Falchikov and J. Goldfinch, Student peer assessment in higher education: A meta-analysis comparing peer and teacher marks, *Review of educational research*, **70**(3), 2000, pp. 287–322.
8. K. Cho, When multi-peers give better advice than an expert: The type and impact of feedback given by students and an expert on student writing (Doctoral Dissertation), University of Pittsburgh, *Retrieved from ProQuest Dissertations & Theses Full Text*, 2004.
9. M. E. Cardella, P. Buzzanell, A. Cummings, D. Tolbert and C. B. Zoltowski, A tale of two design contexts: Quantitative and qualitative explorations of student-instructor interac-

tions amidst ambiguity, *Design Thinking Research Symposium*, West Lafayette IN, 2014.
10. J. M. Levine and R. L. Moreland, Progress in small group research, *Annual Review of Psychology*, **41**(1), 1990, p. 585.
11. D. G. Myers and H. Lamm, The group polarization phenomenon, *Psychological Bulletin*, **83**(4), 1976, p. 602.
12. N. Falchikov, Product comparisons and process benefits of collaborative peer group and self assessments, *Assessment and Evaluation in Higher Education*, **11**(2), 1986, pp. 146–166.
13. D. J. Boud and W. H. Holmes, Self and peer marking in an undergraduate engineering course, *IEEE Transactions on Education*, **24**(4), 1981, pp. 267–274.
14. L. Li, X. Liu and A.L. Steckelberg, Assessor or assessee: How student learning improves by giving and receiving peer feedback, *British Journal of Educational Technology*, **41**(3), 2010, pp. 525–536.
15. N. Falchikov and M. Blythman, *Learning Together: Peer Tutoring in Higher Education*, Routledge-Falmer, London, 2001.
16. M. Ekoniak, M. J. Scanlon and M. J. Mohammadi-Aragh, Improving student writing through multiple peer feedback, *IEEE Frontiers in Education Conference*, Oklahoma City OK, 2013, pp. 626–628.
17. D. Nicol, A. Thomson and C. Breslin, Rethinking feedback practices in higher education: a peer review perspective, *Assessment & Evaluation in Higher Education*, **39**(1), 2014, pp. 102–122.
18. J. H. Kaufman and C. D. Schunn, Students' perceptions about peer assessment for writing: their origin and impact on revision work, *Instructional Science*, **39**(3), 2011, pp. 387–406.
19. M. Conde, L. Sanchez-Gonzalez, V. Matellan-Olivera, and F.J. Rodríguez-Lera, Application of Peer Review Techniques in Engineering Education, *International Journal of Engineering Education*, **33**(2(B)), 2017, pp. 918–926.
20. M. M. Patchan, D. Charney and C. D. Schunn, A validation study of students' end comments: Comparing comments by students, a writing instructor, and a content instructor, *Journal of Writing Research*, **1**(2), 2009, pp. 124–152.
21. T. Hovardas, O. E. Tsivitanidou and Z. C. Zacharia, Peer versus expert feedback: An investigation of the quality of peer feedback among secondary school students, *Computers & Education*, **71**(0), 2014, pp. 133–152.
22. K. Cho and C. D. Schunn, The SWoRD is mightier than the pen: Scaffolded writing and rewriting in the discipline, *IEEE International Conference on Advanced Learning Technologies*, Joensuu Finland, 2004, pp. 545–549.
23. R. Robinson, Calibrated Peer Review$^{TM}$: an application to increase student reading & writing skills, *The American Biology Teacher*, **63**(7), 2001, pp. 474–480.
24. K. P. Cross and M. H. Steadman, *Classroom research: Implementing the scholarship of teaching*, Jossey-Bass San Francisco CA, 1996.
25. C. L. Dym, A. M. Agogino, O. Eris, D. D. Frey and L. J. Leifer, Engineering design thinking, teaching, and learning, *Journal of Engineering Education*, **94**(1), 2005, pp. 103–120.
26. D. A. Schön, *Educating the Reflective Practitioner: Toward a New Design for Teaching and Learning in the Professions*, Jossey-Bass, San Francisco, 1987.
27. D. P. Dannels, Performing tribal rituals: A genre analysis of "crits" in design studios, *Communication Education*, **54**(2), 2005, pp. 136–160.
28. J. E. Tomayko, Teaching software development in a studio environment, *ACM SIGCSE Bulletin*, 1991, pp. 300–303.
29. Y. J. Reimer and S. A. Douglas, Teaching HCI design with the studio approach, *Computer Science Education*, **13**(3), 2003, pp. 191–205.
30. A. Hurst and O. G. Nespoli, Peer review in capstone design courses: An implementation using progress update meetings, *International Journal of Engineering Education*, **31**(6(B)), 2015, pp. 1799–1809.
31. C. E. Kulkarni, M. S. Bernstein and S. R. Klemmer, PeerStudio: Rapid peer feedback emphasizes revision and improves performance, *Proceedings of the Second ACM Conference on Learning@Scale*, 2015, pp. 75–84.
32. N.-F. Liu and D. Carless, Peer feedback: the learning element of peer assessment, *Teaching in Higher education*, **11**(3), 2006, pp. 279–290.
33. S. Gielen, E. Peeters, F. Dochy, P. Onghena and K. Struyven, Improving the effectiveness of peer feedback for learning, *Learning and Instruction*, **20**(4), 2010, pp. 304–315.
34. R. Sadler, Ah!... So that's 'quality', in P. Schwarts and G. Webb (eds), *Assessment: case studies, experience and practice from higher education*, Kogan Page, London, 2002.
35. A. Yuan, K. Luther, M. Krause, S. I. Vennix, S. P. Dow and B. Hartmann, Almost an Expert: The Effects of Rubrics and Expertise on Perceived Value of Crowdsourced Design Critiques, *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, San Francisco, 2016, pp. 1005–1017.
36. D. Boud, Sustainable assessment: rethinking assessment for the learning society, *Studies in continuing education*, **22**(2), 2000, pp. 151–167.
37. S. Vasana and A. D. Ritzhaupt, A case study of a method for hybrid peer-evaluation in engineering education, *World Transactions on Engineering and Technology Education*, **7**(1), 2009, pp. 34–40.
38. F. Marbouti, H. Diefes-Dux and M. Cardella, Students' and engineering educators' feedback on design, *122nd American Society for Engineering Education Annual Conference & Exposition*, Seattle, WA, 2015.
39. G. G. Krauss and L. Neeley, Peer Review Feedback in an Introductory Design Course: Increasing Student Comments and Questions through the use of Written Feedback, *International Journal of Engineering Education*, **32**(3(B)), 2016, pp. 1445–1457.
40. C. J. Atman, D. Kilgore and A. Mckenna, Characterizing Design Learning: A Mixed-Methods Study of Engineering Designers' Use of Language, *Journal of Engineering Education*, **97**(3), 2008, pp. 309–326.
41. H. Zhu, S. P. Dow, R. E. Kraut and A. Kittur, Reviewing versus doing: learning and performance in crowd assessment, *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*, Baltimore, 2014, pp. 1445–1455.
42. K. Cho and C. D. Schunn, Scaffolded writing and rewriting in the discipline: A web-based reciprocal peer review system, *Computers & Education*, **48**(3), 2007, pp. 409–426.
43. C. Schunn, A. Godley and S. Demartino, The Reliability and Validity of Peer Review of Writing in High School AP English Classes, *Journal of Adolescent & Adult Literacy*, 2016.
44. S. G. Hart and L. E. Staveland, Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research, *Advances in Psychology*, **52**, 1988, pp. 139–183.
45. S. G. Hart, NASA-task load index (NASA-TLX); 20 years later, *Proceedings of the human factors and ergonomics society annual meeting*, 2006, pp. 904–908.
46. A. Ivarsson, M. B. Andersen, U. Johnson and M. Lindwall, To adjust or not adjust: Nonparametric effect sizes, confidence intervals, and real-world meaning, *Psychology of Sport and Exercise*, **14**(1), 2013, pp. 97–102.
47. J. P. Goetz and M. D. Lecompte, *Ethnography and qualitative design in educational research*, Academic Press Orlando, FL, 1984.
48. D. J. Nicol and D. Macfarlane-Dick, Formative assessment and self-regulated learning: A model and seven principles of good feedback practice, *Studies in higher education*, **31**(2), 2006, pp. 199–218.
49. K. Reily, P. L. Finnerty and L. Terveen, Two peers are better than one: aggregating peer reviews for computing assignments is surprisingly accurate, *Proceedings of the ACM International Conference on Supporting Group Work*, Sanibel Island FL, 2009, pp. 115–124.
50. R.W. Lent, L. Schmidt, J. Schmidt and G. Pertmer, Exploration of collective efficacy beliefs in student project teams: implications for student and team outcomes, *American Society for Engineering Education Annual Conference & Exposition*, Montreal CA, 2002.
51. R. Althauser and K. Darnall, Enhancing critical reading and writing through peer reviews: An exploration of assisted performance, *Teaching Sociology*, 2001, pp. 23–35.

52. A. R. Carberry, S. R. Brunhaver, K. R. Csavina and A. F. Mckenna, Comparison of written versus verbal peer feedback for design projects, *International Journal of Engineering Education*, **32**(3(B)), 2016, pp. 1458–1471.

53. D. P. Dannels and K. N. Martin, Critiquing critiques a genre analysis of feedback across novice to expert design studios, *Journal of Business and Technical Communication*, **22**(2), 2008, pp. 135–159.

**Mahender Mandala** received his PhD from the University of Pittsburgh in 2016. He is currently Visiting Assistant Professor at the Department of Bioengineering, Swanson School of Engineering, University of Pittsburgh. His research interests include Assistive Technology, Rehabilitation Engineering, Design Research, Medical Product Design, and Engineering Education. Mahender has been involved in the development of a digital design critique tool for use within engineering design education and has several years of experience working on new product development.

**Christian Schunn** obtained his PhD from Carnegie Mellon in 1995. He currently is a Senior Scientist at the Learning Research and Development Center and a Professor of Psychology, Learning Sciences and Policy, and Intelligent Systems at the University of Pittsburgh. Most recently he became Co-Director of the Institute for Learning. He has led many research and design projects in science, mathematics, engineering, technology, and writing education. His current research interests include STEM reasoning (particularly studying practicing scientists and engineers) and learning (developing and studying integrations of science & engineering or science & math), neuroscience of complex learning (in science and math), peer interaction and instruction (especially for writing instruction), and engagement and learning (especially in science). He is a Fellow of several scientific societies (AAAS, APA, APS) as well as a Fellow and Executive member of the International Society for Design & Development in Education. He has served on two National Academy of Engineering committees, K-12 Engineering Education and K-12 Engineering Education Standards.

**Steven Dow** is an Assistant Professor of Cognitive Science and a member of the Design Lab at UC San Diego. Steven received the National Science Foundation CAREER Award in 2015 for research advancing collective innovation. He was co-PI on three other National Science Foundation grants, a Google Faculty Grant, Stanford's Postdoctoral Research Award, and the Hasso Plattner Design Thinking Research Grant. Steven was on the faculty in the HCI Institute at CMU from 2011–2015. He holds an MS and PhD in Human-Centered Computing from the Georgia Institute of Technology, and a BS in Industrial Engineering from University of Iowa.

**Mary Goldberg**, PhD serves as the Education & Outreach Project Director at the Human Engineering Research Laboratories, Assistant Professor in the Department of Rehabilitation Science and Technology, and Innovation Track Director of MS in Clinical Research at the University of Pittsburgh. She has a background in education with a concentration in rehabilitation science; psychology; and Spanish. She has served as Co-PI on several training grants in the field of assistive technology for undergraduates, veterans, and graduate students, with a particular emphasis on students with disabilities. Dr. Goldberg has also been involved in the development and evaluation of peer evaluation and training management software systems. Dr. Goldberg received her PhD in Administrative and Policy Studies of Education with a focus on online learning in assistive technology and her additional research interests include program evaluation, STEM education, and international capacity building in assistive technology.

**Jon Pearlman**, PhD received his BS in Mechanical Engineering from the University of California at Berkeley and his M.Sc. in Mechanical Engineering from Cornell University with a focus in Biomechanics and his PhD in Rehabilitation Science and Technology at the University of Pittsburgh. His research centers on Assistive Technology Product development and quality control. Related to these topics, Dr. Pearlman has several assistive technology design and development projects which are funded through the VA and the University of Pittsburgh; many of these projects are in collaboration with clinician and commercialization partners to ensure they are clinically relevant and are commercially available. Jon is inventor or co-inventor on nine issued or pending patents.

**William Clark**, PhD is a Professor of Mechanical Engineering and Materials Science at the University of Pittsburgh. He has directed an active research program in dynamic systems and controls throughout his career, resulting in over 110 technical papers on the subjects of smart structures, dynamic systems, sensors, and control systems. He has been an active participant in the ASME, having chaired the Technical Committee on Vibration and Sound and the Design Engineering Division. Dr. Clark's teaching has recently focused more on innovation, with courses in lean innovation, entrepreneurship and the creative design process in order to equip students with skills and tools to become innovative and creative engineers.

**Irene B. Mena**, PhD has a BS and MS in industrial engineering, and a PhD in engineering education. Her research interests include first-year engineering and graduate student professional development.