Are Many Heads Better Than One?  
Using Peer Reviews in Engineering Design Courses

Vahid Garousi  
Software Quality Engineering Research Group (SoftQual)  
Department of Electrical and Computer Engineering, University of Calgary  
2500 University Drive NW, Calgary, AB Canada T2N 1N4  
vgarousi@ucalgary.ca

Abstract

It is important for engineering students to peer review each other’s work during design projects. Based on the demonstrated value of peer reviews in engineering (e.g., the software industry), numerous industry experts have listed it at the top of the list of desirable engineering skills and practices. However, surprisingly, not many engineering courses in Canadian or even non-Canadian universities and colleges include peer review activities in their design courses. The author thus decided to apply peer reviews to the design project of a senior software engineering course. The purpose of this article is to present our experimental findings, lessons learned, possible challenges and recommendations that may be used to promote learning and also the usage of peer review activities in teaching other engineering courses. The results of our experiment show promising signs of using peer review in a design project.

1. Introduction

Peer review is an important step in the design and implementation of engineering projects which provides an evaluation of design decisions and documents to meet quality objectives. Peer reviews are usually performed by independent teams or individuals not associated with the original design team [1].

Peer review has also been found to be one of the most effective ways to promote quality and productivity of design processes in various engineering disciplines (e.g., software [2], electrical [3], civil [4], mechanical [5], and fire [6]). The primary purpose of peer review activities in engineering design projects is to identify defects existing within the work products developed throughout the development life-cycle.

Data collected during the peer review process is used not only to correct defects, but also to evaluate and improve the development process itself. Reports from the software industry, for example, indicate that inspections (one type of peer reviews) have gained wide acceptance as a development tactic and can take up to 15 percent of the time allotted to a software project [7]. Based on the demonstrated value of peer reviews in engineering, numerous industry experts have listed it at the top of the list of desirable software development skills and practices (e.g., [7]).

However, surprisingly, not many design-oriented engineering courses in Canadian universities and colleges currently include peer review activities in their design courses. This is while the engineering industry heavily relies on different types of peer review activities (e.g., team review, inspection, and walkthrough). We cannot afford to graduate engineering students who are unable to effectively perform peer reviews when they need to critically examine the quality of products (or documents) generated by their peers.

In addition to helping students learn more about the peer review process, several published articles (e.g., [8]) have also indicated that the peer review process may offer other educational benefits for students such as improved analytical and defect detection abilities.

There have been other studies in the literature (e.g., [9-12]) which have empirically studied the impact of different types of peer review activities in design-oriented engineering courses (discussed in Section 2). To further assess the effectiveness of peer reviews in a Canadian design engineering classroom setting, we applied peer reviews during the design project of a 4th year course in the software engineering program of the University of Calgary. The outcomes from the instructor and students’ experience with peer reviews are reported in this paper. The author believes that the recommendations derived from the data analysis are of general interest and he hopes the findings will encourage more use of peer reviews in other
engineering courses as engineering is by nature a peer-review-based discipline.

The rest of this article is structured as follows. Related works are discussed in Section 2. The experiment setup details are explained in Section 3. The outcomes of the experiment are reported in Section 4. Finally, conclusions and future works are discussed in Section 5.

The author is aware that the experiment outcomes and data presented in this paper are based on one design course project and that a word of caution is required before generalizing the conclusions.

2. Related Works

There have been only a few engineering design courses which involve students in peer review processes (e.g., [9-15]). Due to space constraint, we discuss in this paper only a few of the related works. The reader is referred to [16] for a wider survey of related works on the topic. Interestingly, most of the known usages of peer reviews so far have been in the software engineering courses.

The work in [9] is a comparative study of academic and industry software projects with the goal of improving academic software engineering projects. Controlled peer reviews were incorporated in two academic projects (in two Canadian universities: University of British Columbia, and École Polytechnique de Montréal) and in an industrial project to compare its effectiveness. The experiment was reported to be very useful in teaching real-life engineering peer review skills to students.

Based on using peer reviews in courses offered at the Worcester Polytechnic Institute, the article in [10] discusses ways to better prepare information system professionals for a rapidly changing world in designing such systems, and identifies the challenge for educators. Authors argue that students also need opportunities to develop the skills to define criteria, set standards, and evaluate both products and processes. They recommend that students should review the work of teams from prior semesters before embarking on their own projects. In doing so, they can learn from the mistakes of others and practice continuous improvement. It is discussed that many faculty members who have attempted peer review claim that students are not very good at it because they are not accustomed to it. But the authors of [10] propose that, with practice, student could get good at it.

An experiment report with the use of reciprocal peer reviews in a human-computer interaction design laboratory at Indiana University is reported in [11]. Sullivan, the author and also the instructor, reported that intensive peer review empowered students to master the course concepts, hone their teamwork and communication skills, master the peer review process, and learn to learn from each other. As an interesting and useful experience, Sullivan reported that for students who are not accustomed to doing peer reviews, inspections (as a more formal type of review) can be too big of a step. For students, the prospect of maintaining inspection statistics can be so intimidating as to thwart mastery of the peer review basics. He thus suggested that walkthroughs can be used as a stepping stone toward building the skills needed for inspections. For this reason, we also applied the walkthrough paradigm as the formality level of peer-review in our project and experiment.

Tyran’s quantitative experimental study in [12] is a similar work to ours in which the impact of a software inspection exercise (a type of peer review) is studied in a software analysis and design course in the Western Washington University in 2006. Overall, Tyran found the software inspection exercise to be a worthwhile component to the course. The exercise provided an experiential approach to help students gain a better feel for the software inspection process. It was observed that once students have completed the exercise, they will often apply their new inspection skills to other courses such as team-based design projects.

The work in [13] describes techniques that were used in different competency-based engineering design courses at Washington State University, as part of the Transferable Integrated Design in Engineering Education (TIbDEE) coalition. Activities performed in the courses included peer evaluation of team deliverables, and led to developing ways to measure performance, reflecting on team processes and setting realistic goals that can be achieved. Feedback for the team came from a written peer evaluation and a self-evaluation. The overall outcome and student perceptions of the use of peer review were reported to be outstanding.

The outline for the 4th year design course (ECE 445) [14] at the University of Illinois at Urbana-Champaign includes a “team review” phase. The followings are excerpts from the course outline: “Each team must peer review another”. “Sponsors are invited to participate in reviews, as they wish, and to make input during those reviews”.

The outline of the course CON 492 (Construction Contracts) [15] at the Bradley University’s Department of Construction Engineering includes a topic on the role of peer review and inspection in construction operations. However, it is not clear how detailed the review process is taught in that course.

It is very evident that the modern engineering industry relies extensively on different types of peer
4. Experiment Results

We now report the findings of our experiment, lessons learned, possible challenges and recommendations.

4.1. Questionnaire Results

Similar to a few other empirical studies on peer reviews (e.g., [12]), a questionnaire was prepared to quantitatively measure students’ perception and satisfaction. The questionnaire and its statistical data are shown in Table 1. The questionnaire format was adapted from works by Tyran [12] and Wiegers [19]. A 7-point Likert scale was used for the questionnaire: (1) Strongly disagree….., (7) Strongly agree.

The questionnaire was distributed to the students in the class after the project had finished. Answers were anonymous. 8 students filled out the questionnaire. The average and standard deviations (STD) of the questionnaire results are reported in Table 1. The histogram of questionnaire results is visualized as a stacked-column chart in Figure 1. The most interesting observations based on the histogram values are discussed next.

Out of 64 question answers (8 students * 8 questions), there was no “strongly disagree” answer for any question.

Ideally, as an outcome of doing the exercise, the author wanted the students to build their confidence regarding the peer review process. It appears that the exercise was effective in this regard. By comparing question 1 and 2 answers, we can clearly see that students’ confidence in conducting a successful peer review increased after the exercise. Hopefully, this increased level of confidence may encourage students to conduct effective peer reviews in their future careers as it is the norm in industry.

In addition, the author hoped that students would gain other educational benefits, e.g., improved analytical abilities and defect detection abilities (e.g., [8]). By looking at question 3 and 4 answers, we can notice a noticeable knowledge gains in preparing high quality design documents, and being more effective in identifying defects, as most students agreed or strongly agreed to these questions.

By looking at question 5 answers, one can notice that only 2 out of 8 students were not satisfied with the process (the number of “somewhat disagree” answers in question 5). According to question 6 answers, only 2 of 8 students thought performing the exercise was not a very pleasant experience. This issue will be synthesized in our discussions on the lessons learned (Section 4.3). Only one student answered...
Table 1. Questionnaire and it statistical data.

<table>
<thead>
<tr>
<th>Question number</th>
<th>Questionnaire Item</th>
<th>Average</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre vs. Post Peer Review Exercise: Change in Confidence</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.</td>
<td>Before I did the exercise, I was confident that I could organize and conduct a successful software peer review.</td>
<td>4.37</td>
<td>1.77</td>
</tr>
<tr>
<td>2.</td>
<td>After I did the exercise, I am confident that I could organize and conduct a successful software peer review.</td>
<td>5.75</td>
<td>0.71</td>
</tr>
<tr>
<td>Perception of Knowledge Gains: As a result of participating in this exercise, I gained knowledge that will help me:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.</td>
<td>To prepare higher quality design documents in the future (i.e., documents similar to the ones used in this exercise)</td>
<td>5.62</td>
<td>0.52</td>
</tr>
<tr>
<td>4.</td>
<td>To be more effective in identifying defects in design documents in the future</td>
<td>5.5</td>
<td>0.93</td>
</tr>
<tr>
<td>Satisfaction with the Peer Review Process and Perception of the Exercise</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.</td>
<td>I was satisfied with the process that was used in this exercise to identify defects and generate actions (feedback) list.</td>
<td>4.75</td>
<td>1.16</td>
</tr>
<tr>
<td>6.</td>
<td>Performing this exercise was a pleasant experience.</td>
<td>4.62</td>
<td>1.51</td>
</tr>
<tr>
<td>7.</td>
<td>I think the peer review and inspections met my team’s objectives.</td>
<td>5.37</td>
<td>0.74</td>
</tr>
<tr>
<td>8.</td>
<td>I felt I was able to significantly improve the quality of the work product through the inspections.</td>
<td>5.37</td>
<td>0.92</td>
</tr>
</tbody>
</table>

Figure 1- Likert-scale data of students’ perception and satisfaction. “neither” for questions 7 and 8, while others had positive answers on those questions.

4.2. Thoroughness of Peer Reviews

To analyze the thoroughness of peer reviews in different project milestones and by different teams, we measured the length of peer reviews (in words) made by each team. Although, this might be seen as a naïve thoroughness/complexity metric, however it proved to be a simple and not very time-consuming tool in measuring thoroughness.

The length of peer reviews performed by each team for each team (as per the peer review assignment) was measured during each of the three project milestones. Due to space shortage, only the lengths (in words) of peer review documents made in milestone 1 is depicted as a 3-D surface plot in Figure 2(a). The results for milestones 2 and 3 can be found in [16]. The average and standard deviations (STD) of the values per each milestone are shown in Figure 2:(b). The most interesting observations based on these charts are discussed next.

As the 3-D plots show, the thoroughness of peer reviews was not homogenous, i.e., some teams wrote more extensive reviews than others. For example, in milestone 1, the analysis document of team 5 was reviewed by team 2 and they provided a review document with 856 words, while team 2 provided a review document of 231 words for team 1.

The reviews done by each team were not homogenous in length either, e.g., in milestone 1, team 2’s review of team 5’s work was 856 words in length, while the same team (#2) wrote a shorter review document (231 words) for team 1’s analysis document. This can be explained according to at least three possible causes: (1) team 5’s document had many issues compared to that of team 2, (2) team 2 reviewed team 5’s document earlier and they had enough time to do it before deadline, while they had time shortage (due to the milestone 1 deadline) to thoroughly review team 1 document, or (3) there were personal/friendship ties (between team 5 and 2) which led to more thorough review by team 2 for team 5. These are all realistic issues and challenges which should be taken into account by the peer review process organizer.

As the average and the STD values chart shows, peer reviews’ lengths decreased as the project progressed from milestone 1 to milestones 2 and 3. This can be explained according to last least two reasons: (1) the average quality of work artefacts increased by project’s progress, thus there were less issues in the artefacts, and that led to shorter review documents, (2) students’ interest in the peer review process decreased, so they were not motivated anymore to write lengthy peer reviews. The author’s inspection and observation implied the first reason above which is a positive sign. By a careful analysis on the projects, there was a noticeable improvement in the average quality of work artefacts as the project progressed.
4.3. Lessons Learned and Recommendations

As discussed in Section 3, we intentionally applied a less formal peer review process in this project. As Sullivan discussed in [11], for students, the prospect of maintaining inspection statistics can be so intimidating as to prevent mastery of the peer review basics. Thus, Sullivan [11] believes that walkthroughs can be a less complex activity. We also experienced similar findings in this regard.

Also, the author did not micro manage the peer review process w.r.t. timing details, thoroughness of reviews. The students were given the freedom to choose their own defect check-list and degree of thoroughness. We noticed from student comments that they faced some challenges and some of them wished they were given full details of peer review expectations and also had formal training on how to do peer review.

Given the time constraints for the course (13 weeks) and also extensive contents of the course (Component-Based Software Engineering), the instructor chose the above settings and he believes the outcome was satisfactory. Similar to every new experiment in engineering, and especially software engineering, there are certainly new things and experiences to learn from every experiment.

At the end, a suitable trade-off analysis should be used by each instructor to determine the choice of formality level for the peer review process. If there is enough time in the beginning of an engineering project to briefly train students on peer review formally, it will certainly lead to better outcome and satisfaction at the end. However, as there are always many things to learn than we have time for, it usually remains as a challenge to find the best balance between how much formal peer review and how much the details of the engineering project at hand to teach to students.

4.4. Threat to Validity

The class size is the main threat to the validity of this experiment. The course under study had a low enrolment (only 9 students). Perhaps, further studies (following the steps reported in this article) should be done on the use of peer reviews in other design-oriented engineering courses with more student enrolment. However, as we discussed in Section 4, our experiment results seem be rational and a good indicator of the benefits and also possible challenges of the use of peer reviews in engineering courses.

5. Conclusions and Future Works

Based on the demonstrated value of peer reviews in engineering (e.g., the software industry), numerous industry experts have listed it at the top of the list of desirable engineering development practices. Experience has shown that problems are eliminated earlier this way and reviews are as effective as or even more effective than testing. Thus, the peer review practice has become the norm and part of the culture in virtually all successful engineering companies, and it is increasing added to various engineering codes (e.g., in civil engineering).

However, surprisingly, not many design-oriented engineering courses in Canadian or even non-Canadian universities and colleges include peer review activities in their design projects. The author thus applied a peer review exercise during the design project of a senior software engineering course. In our experiment, we observed noticeable knowledge gains in students with respect to preparing high quality design documents and also being more effective in identifying defects. Also we found out that the students’ confidence in conducting a successful peer review (in future projects) increased after the exercise,
something which will for sure help them in their future engineering careers and will directly help build better engineering products for Canada.

As future works, the author is thinking of the following directions. We did not provide a formal training to students on peer review details itself. It would be worthwhile in a future project setting to briefly teach peer review skills and tools and assess the extent to which it impacts the students’ success in effective peer reviewing.

We would also like to perform more formal peer review processes (e.g., inspection) and also to use peer review support software tools (e.g., Microsoft Visual Studio Team System, and IBM Portfolio Manager).

Acknowledgements

This work was supported in part by the Alberta Ingenuity New Faculty Award no. 200600673. The author is grateful to the students and also the teaching assistant of the course for their enthusiasm and collaboration in this experiment.

References