Peer Code Review as Formative Assessment: A Case Study from a Database Course Project*

Rune Hjelsvold[0000−0002−5955−1603] and Nipuna Hiranya Weeratunge[0000−0003−4612−0342]

Norwegian University of Science and Technology
{rune.hjelsvold,nipuna.h.weeratunge}@ntnu.no

Abstract. Code review is an important quality assurance activity for software engineers. In addition, both reviewers and developers may improve their professional competencies when participating in code reviews. This paper describes a three-week database course project where the students designed and developed a database and a database application. After a three-week project work period, the students submitted their designs and their code repositories for peer review. 115 students participated in the code review of the 68 repositories submitted for review. The students could make use of the feedback they received and what they learned from studying other student's code when revising the project for the final submission, which was graded by the teaching staff.
In their final reports, the students reported on benefits and barriers for effective peer review that seemed to be more accentuated when the purpose was formative assessment. Benefits included high level thinking and deeper levels of learning and giving the students different points of view when reviewing their own solutions. Some of the barriers were the students' lack of domain knowledge and lack of confidence in being reviewers, the quality of the projects submitted for review, the review workload, and the amount of work needed to modify the code after the review.

Keywords: Computer Science Education · Database Course · Project Work · Peer Code Review · Formative Assessment

1 Introduction

Code review is an important software quality assurance activity for software engineers. In a study of 911 developers at Microsoft, MacLeod et al. [5] found that almost 90% of the respondents had been participating in code review the week before answering the survey. They also found that *improve code* and *find defects* where the two highest ranked motivations for developers to participate in code review. The next two slots in the ranked list of motivations were *transfer knowledge* and *explore alternative solutions*. The latter two can be considered as part of on-the-job learning.

* Supported by Excited - The Centre for Excellent IT Education

Peer assessment is also considered to be important for the development of evaluative judgement [9]. Computer science students could therefore benefit from practicing peer code review in their software development projects at the university. There is, however, a lack of papers addressing the use of peer code review as part of the formative assessment taking place during a course rather than as part of summative assessment.

This paper is a case study about the use of formative peer code review in a database course for bachelor CS students. The research questions for this case study are:

- RQ1: What are the benefits of formative peer code review as seen by students?
- RQ2: What are the barriers faced by students in the formative peer code review process?

1.1 Methodology

As discussed by Harland [2], a case study will always be context-dependent. Still, learning from specific cases is one of the ways that can help us understand complex real-life social situations. The researchers behind this case study were the ones teaching the database course and are main beneficiaries of the analysis presented in this paper. The description of the case, the analysis of the participating students' reflections, and the discussion of the implications and future work should be of interest to educators seeking to enhance their understanding of peer code review in student projects.

2 Related Work

This section will detail the existing literature on formative peer code review to investigate the current state of research. An important aspect in this topic would be the use of formative feedback. Thus, in [7] the authors have investigated the formative use of rubrics and their impact on student learning. They analysed 21 studies through content analysis with focus on sample, subject/task, design, procedure and findings, in relation to student performance and self-regulation. The results of the study indicate that rubrics can mediate improved performance through providing assessment transparency which can in-turn reduce student anxiety. In addition, the authors also suggest that the use of rubrics can aid the feedback process, enhance student self-efficacy and encourage student self-regulation, which in-turn may indirectly facilitate better student performance.

Similarly, the study [4] has focused on encouraging active engagement from students with feedback. The authors start by stating that the quality of students' engagement and use of the feedback they receive is crucial for taking advantage of the formative potential feedback. However, the authors also note that the students' engagement with feedback is usually not very productive and the exact reasons for this lack of engagement cannot be identified. With all this complexity,

the authors suggest several methods to facilitate active engagement, such as providing useful feedback to the students, helping students to develop constructive strategies for using feedback and avoiding grades on individual assignments.

In [8] the authors investigate the current state of knowledge in formative peer review. The authors state that peer feedback can be of the same quality as teachers' feedback since it can include feedback that is written from a student-centric perspective in relatively simple language. In addition, they further state that the quality of the peer feedback can be improved by scaffolding interventions, such as scripting the steps for the assessor. Another important suggestion from the authors is that the domain knowledge should be considered using a developmental skill approach where the students are given multiple opportunities to practice. They further emphasise that this approach can provide better quality peer feedback as the effect of practice can make a difference.

The authors of [1] also report using peer review to motivate and engage students. The authors note that participating in peer review promotes critical thinking while creating scholarly engagement and the development of skills in analysing, correcting and commenting on the deliverables of others aligns well with professional skill requirements in many fields. Thus, the authors applied and explored student peer assessment in a project-based course in web development at a university. During this course the students deliver sub-projects and conduct evaluations of other students' code and working prototypes. The authors present two course designs, the original and one revised based on the experiences made in the original version. One of their findings is that the quality of students' comments and evaluations increase significantly when the reviews are written by the students individually and are included in the final grading.

Another relevant study is [6] where the authors attempts to investigate whether peer feedback instructional designs impact students' learning perceptions. In order to assess this, they conducted a comparative study at an educational institute following students during the first two years of a teacher education program. During this study, students participated in two consecutive peer feedback experiences using distinct instructional designs. According to the authors, the results indicate that the students consider that long-term interventions with prior training and double loop feedback processes being more useful for their performance than a short-term experience without face-to-face training and single-loop feedback processes. Furthermore, the authors state that students usually perceive more benefits when they provide feedback than when they receive it.

The authors of [3] had conducted a systematic review of literature on peer code review in higher education, in order to explore the motivations of instructors when conducting peer code review, how the activities were implemented in practice, primary reported benefits and barriers. The authors initially identified 187 prospective studies and analysed 51 studies in detail in the end. According to the authors, the most commonly reported benefit was the development of programming related skills while the most commonly reported barrier was low student engagement.

The authors of [11] did an in-depth study of whether peer review would promote higher level thinking in the students. They divided the students in three groups - one peer review group, one training review group where the students reviewed code prepared by the teaching staff, and one control group. They then analysed differences in the groups' acquired higher level thinking skills in object-oriented programming. They did not find strong evidence that peer review did promote higher level thinking. As future work, they recommend giving the students more complex designs or multiple solutions, each solving the problem in a different way.

2.1 Implications for the Design of the Peer Review Case

In this subsection we will briefly summarise the design choices we made when planning the peer review activity in the database course project:

- We created a project case that was somewhat complex and that would likely result in different solutions to explore the suggestion made by [11] that complex designs and different solutions might promote higher level thinking.
- We developed a detailed feedback form acting as a rubric [7] and a script [8] for the reviews, knowing that students see more benefits from providing feedback than receiving [6].
- Lack of student engagement is a known barrier for peer review [1,3,4]. We took several measures to address this issue. We introduced the students to the review process as a professional skill, and we developed a review feedback form that was well aligned with the expected learning outcomes. In addition, students worked on their reviews individually, and the students' reflections on their review contributions were graded.
- Prior to the project, the students were given four different assignments that would help the students develop domain knowledge [8] that would be beneficial to their review work.

The design of the peer review activity is described in more detail in Section 3.4

3 The Database Course Project

This paper is a case study about the use of formative peer code review in the NTNU Gjøvik course IDATG2204 Data Modelling and Database Systems. The course is 7.5 ECTS and a mandatory fourth semester course for bachelor students in computer science. The study was done in the spring of 2021, when the course was run for the first time in its current form. 139 students had signed up for the course; 109 students submitted their work for formal assessment.

The course was organised as a flipped classroom course with two in-class sessions per week, one four-hour session and one two-hour session used for exercises. The first part of the course covered topics related to the design and implementation of databases and database applications. There were five mandatory assignments in the course. The students had to pass at least four of these

to pass the course. Four of the assignments were related to the topics covered in the first part of the course.

The second part of the course consisted of project work. There were no in-class exercises or mandatory assignments during the project work weeks.

The third part of the course covered topics related to the operation and management of database systems. The fifth mandatory assignment was related to these topics.

The course was designed to have the in-class sessions and project work taking place in computer labs on campus. The assessment was designed to be based on the project work and a three-hour written school exam. As discussed in Section 3.5, these designs had to be changed due to the COVID-19 pandemic.

3.1 Purpose

The database project was designed to give the students an opportunity to refine their skills in designing and implementing databases and database applications by solving a practical case. The peer review activity was introduced to serve three different purposes:

1. To have the students receive formative feedback on their project work.
2. To give the students an opportunity to deepen their database competence by reviewing other solutions.
3. To give the students an opportunity to develop and refine their skills in doing code review.

3.2 The Project Case

The case given to the students was to develop a database for a product manufacturing company. The database should keep information about product types, products, production plans, orders, shipments, customers, and other types of users of the system. A typical database model for the case would consist of some 15-20 entity types and approximately the same number of relationship types. The students were also asked to develop a REST-type API for the various types of users to retrieve, insert, update, and delete information stored in the database. Furthermore, the students were expected to develop unit and API tests for the code that they developed.

The students were to develop the system using MariaDB Server[1] as the DBMS and PHP[2] and the Apache HTTP Server[3] for the implementation of the API. The students would further use Codeception[4] for developing unit and API tests.

[1] `https://mariadb.org/`
[2] `https://www.php.net/`
[3] `https://httpd.apache.org/`
[4] `https://codeception.com/`

3.3 Organisation

The students could work on the project individually or in teams of up to three students. 48 project repositories were submitted in total at the end of the course.

The project period lasted for three weeks, followed by a one-week peer code review period. The deadline for the final submission was set to three weeks after the last in-class session in the course, giving students the opportunity to refine their project based on what they learned from the peer review activity.

3.4 Peer Code Review

Some students had participated in peer code review in earlier courses, but many students had not. A short introduction to peer code review was given to all students at the beginning of the peer code review period.

The code reviews were to be written by the students individually. Each student then got an independent peer code review experience. In addition, each project received more reviews in this setup than if each review had been written jointly by a group of students.The students were recommended to write at least three reviews each to get to see several different designs and implementations.

The review process was managed by a locally developed system called CSAMS (Computer Science Assignment Submission System). The teaching staff created a review form in CSAMS consisting of 20 specific questions covering these five aspects of the project deliverable:

- The conceptual database model (UML)
- The relational database schema
- The database application design
- The implementation
- The system deployment means

For each of the 20 questions, the students were asked to rate the project. For most of the questions, the students could rate the questions with the alternatives *poor*, *fair*, *good*, and *very good*. Each of the alternatives had a question specific description to facilitate consistency among the answers. The students were also encouraged to write more detailed comments for explaining the reason for the rating or about the project in general. Figure 1 shows the conceptual database modeling part of the review form.

3.5 The COVID-19 Pandemic

The course could not be offered as originally designed because the campus was completely closed for students during the whole semester due to the COVID-19 pandemic. The main consequences for the database project and the peer review activity were:

- Students, teaching staff, and teaching assistants did not meet in the computer lab as planned. Therefore, students struggled more to organise their work and to seek help and were less productive in an on-campus setting.

Fig. 1. Screen dump showing the conceptual database modeling review questions

- The students experienced increased levels of stress and increased workloads due to the consequences of the COVID-19 pandemic. Students were therefore less likely to be ready for working in unfamiliar ways.
- Many groups were not able to implement much of their system before the peer review deadline expired. The project period was therefore extended with one additional project week and one extra review period.
- The school exam was replaced by a combination of a home exam and a written peer review experience report.

4 Results

The results presented here are extracted from the students' project reports and from the students' peer review experience report. The results are structured according to the classification of peer review benefits and barriers used by Indriasari et al. [3].

4.1 Benefits

Knowledge or Skill Development Many students had positive reflections on the knowledge and skill development benefits. Several students mentioned how reviewing their peers helped them deepen their own competence, e.g.,

> *Although I learned a lot from revising peer reviews I think reviewing other groups' work was just as productive if not even more so. I got to see a*

lot of good implementations and approaches to problems that I had not thought of on my own. And when I criticized other group's solutions it forced me to think about why their solution was wrong or suboptimal. This led to me gaining a greater insight into the project case and databases in general.

There is something about giving others help which makes me double check my answer, which results in me critically thinking twice about my response before I give an answer. I feel like it is an excellent way of learning for both parties.

Learning Support The review activity seemed to support several students in their learning, e.g.,

For sure I appreciate the teaching methodology our instructors have made. At the beginning many of us, including me, thought that it has no meaning to do peer reviews, it is just a waste of time, but I realized how it works after I invested my time three full days where I went through all peer reviews from others work. This has given me a good understanding of how the project works. I can openly say that this is a better way to learn databases and it did work for me.

Some students reflected on how the questions in the feedback form acted as guidance in their learning, e.g.,

One of the most useful things I got from writing reviews, was the questions asked for each section of the review. Seeing these and writing what could have been done better or what was done well in other projects, made it easy to identify problems in our own project.

Some students reflected on how writing reviews motivated them to revisit course literature, e.g.,

For me, the peer review acted as a repetition of core topics of the course since I was "forced" to read up on some of the topics before giving feedback to others about them.

Some students also commented on how seeing other solutions helped them understand what they were supposed to learn, e.g.,

What I Learned: Giving Peer Reviews
Grading other peoples projects gave me good insight into what was expected of us.

Product Quality Improvement Several students reflected on how reviewing other projects had a positive effect on the product quality, e.g.,

By reviewing others projects I gained several insights and ideas on different ways to interpret the project assignment. This gave me new ideas for how my group could improve our project as well.

Social Implications Several students mentioned how seeing other students' accomplishments was good for their self-confidence, e.g.,

> *For the projects that were similar to mine, I discovered that my group were at least not the only ones that had not understood as much as we thought we needed to. This was quite calming*

> *I could see that my interpretation of the project domain was similar to what the other group had thought, and this gave me a confidence boost.*

Other Several students also reflected on the value of acquiring competence in peer code review, e.g.,

> *Peer reviewing as a concept is something we will encounter in the development industry a lot, be it getting reviewed or reviewing someone else's work. I will take this into account and improve my reviewing knowledge further.*

> *Peer reviewing was an opportunity to gain insight into the work done by my peers. It was also an opportunity to fine-tune my feedback skills and ability to provide constructive criticism.*

4.2 Barriers

Impractical Review Process Two barriers related to the review process seemed to be especially large for the students in this case. The first of these was that many of the project teams had not yet reached the state where they had a "review ready" version of their systems. Many of the reviewers therefore got to review systems with no or only a rudimentary implementation submitted for review, e.g.,

> *The peer review for me wasn't exactly as I had expected, most of the projects I was given wasn't near a final product, so it was hard to come up with any good feedback.*

The second barrier was related to the CSAMS system used for review. The system did not have a snapshot storage mechanism. Therefore, some students experienced that review work was lost due to technical problems, e.g.,

> *For the peer reviews there were a couple of issues. The first issue is that I had a 900+ word review vanished into the air because of bad design. The website where we wrote our reviews timed out and everything I had written got deleted. That was very demotivating and all those hours of looking and reviewing that project were just gone.*

Lack of Knowledge or Ability Several students mentioned the lack of peer review competence as a barrier, e.g.,

> *Giving other fellow students feedback on their work was a task I am unfamiliar with from before. [...] Although we got an introduction through a lecture on feedback, I am the type who needs practical experience before I feel confident in the role. Therefor it made me feel uncomfortable in some way, since I did not have enough experience to be confident in my role.*

Several students also reflected on the lack of domain knowledge, especially database application implementation skills, e.g.,

> *As per the quick note above – I feel that my database competency at the time of the peer-reviews, impacted the reviews negatively.*

> *As mentioned in the reflection above, I didn't know a lot about the API aspect or the code at the time I was writing the reviews. I therefore focused mostly on the conceptual and logical models.*

Low Learning Engagement Some students reflected on a general lack of motivation for participating in the peer review activity, e.g.,

> *By not actively reflecting on why we did this and what I could have gotten out of it, my learning outcome from this activity was to some degree impaired. My motivation was also reduced as result of this.*

Some students reflected on lack of engagement due to the amount of time needed to participate in the peer review activity, e.g.,

> *However, I felt like I needed more time to write the feedback than I got. I devoted as much time as possible to finishing the reviews, but I felt like it still wasn't enough time.*

> *If we were still at the modeling stage of the project, I would have greatly benefited from the lessons I'd learnt from seeing others solutions. Sadly, we were past that stage, and time constraints did not permit going back and making such changes.*

Some students attributed lack of engagement to the diversity of solutions among the projects to be reviewed, e.g.,

> *In my opinion I didn't really enjoy the peer reviews, people do things very differently, some didn't follow the way of doings things as we've been instructed, so it was rather confusing actually.*

Ineffective Administrative Process Some students mentioned unclear information or lack of information as a major barrier, e.g.,

> *I found the process frustrating because again as with the start of the project some vital details have not been considered from lecturer.*

> *On the other side, the execution was not that great in my opinion. This might be due to a lack of clear flow of information, making it difficult to plan ahead about when to work and not fully understanding the scope of the reviews.*

Other Finally, there were some students mentioning negative social implications of participating in the peer review activity, e.g.,

> *The whole peer review process felt very frustrating, from [...] feeling like your own program/design is a lot worse than what others have come up with.*

5 Discussion

5.1 Course Design Choices

In section 2.1, we summarised our design choices when designing the peer review activity. In this section, we will discuss these design decisions in regards to the results presented in Section 4.

Rubric The students' reflections suggest that the review form was helpful as a guidance for the review process and as a tool for students to improve their own project and for deepening their domain knowledge. The two types of barriers that can be seen mentioned by the students are (1) the time and efforts required to address all the questions in the form and (2) the negative impact on the students' self-confidence if they don't feel competent to address all the issues.

Student Engagement Many students reflected on code review being an important professional skill and reported on the peer review activity being a positive experience both for acquiring deepened domain knowledge and for extending the peer review competence itself. Two of the main engagement barriers in this study were related to the weaknesses of the CSAMS system used for peer reviewing and the unfinished state of many of the projects submitted for review.

Domain Competence Lack of domain competence is known to be another major barrier. Several students did mention lack of domain competence in this study too, even though four assignments had been prepared to help the students develop necessary domain knowledge prior to the project. The four assignments, however, are all related to the topics of database design and querying; none of

the assignments were related to the topic of database application development. Several students commented that they were more confident in reviewing the database design than the database application development. This might be related to the fact that the students had more experience in database design from the assignments.

5.2 COVID-19 Implications

As mentioned in Section 3.5, the course could not be run as originally designed due to the COVID-19 restrictions. It is likely that the COVID-19 implications on the students' general motivation, workload, productivity, and communication with the educators contributed significantly to the barriers that the students experienced.

5.3 Planned Modifications to the Course

Based on the experiences discussed in this paper, we will be making several changes to the course next year.

Prior Domain Knowledge Lack of database application development competence seemed to be a major barrier for the students. Next year, we will therefore have one of the assignments devoted to database application development. We will also introduce the review form early into the course so that the students are more familiar with the questions and what assessments are expected for each of them.

Review State Reviewers suffered from many projects not being in a "ready for review" state. Next year, we will therefore define a set of minimal requirements to be met by projects to be entitled for submission for peer review.

Review System Reviewers suffered from weaknesses in the review system. Next year, we will therefore ensure that there are no such weaknesses in the chosen review system.

5.4 Limitations

As a case study based on student-reported results, this study has limitations regarding generalisability. The study may still help us, as a research community, to gain a better understanding of the benefits and barriers of peer code review in CS education. It should be noted, however, that the results may be somewhat biased because the research was conducted by the educators involved in the course and because the students' reflections were taken from reports that were formally graded by the said educators. It is quite likely that the students were more positive than what they might have been otherwise and that they reported more on the expected experiences than what they really experienced.

6 Conclusion and Future Work

In this paper, we have described and discussed the introduction of peer code review as formative assessment in a database course. Several students reported on the peer review activity being a valuable experience allowing them to develop professional code review skills and encouraging them to deepen their domain knowledge and to review their own solution from additional points of views. We have also learned that peer code review as formative assessment relies on projects being ready for review before being submitted and that lack of domain knowledge is a major barrier.

There is a lot more to learn about peer code review in CS education. More studies should help us better understand how to design effective peer code review activities.

A Peer Review Classification

This appendix includes a detailed classification of the peer review activity, following the typology developed by [10].

Variable	**Classification**
Curriculum area/subject	Databases
Objectives	Developing peer code review skills; formative assessment; improved learning by reviewing other solutions
Focus	Mainly qualitative feedback but also quantitative scores
Product/output	Design and development project work
Relation to staff assessment	Peer code review substituted staff formative feedback; student peer review reflection reports were formally assessed by staff
Official weight	No direct relationship to the formal assessment
Directionality	Unidirectional from reviewer to reviewed
Privacy	Anonymous (although code repository addresses could reveal the identify of individual group members)
Contact	Written feedback
Year	Second year bachelor students
Ability	Random assignment of reviewers
Constellation Assessors	Each student reviewed projects individually; students were encouraged to review three different projects

Variable	Classification
Constellation Assessed	Up to three students could work on one project team
Place	Reviews were exchanged through the locally developed CSAMS system
Time	Four lab hours were reserved for the writing feedback; students did also use home time for the review
Requirement	Participation was not directly required but it would be hard for the students to write their reflection report without participating
Reward	No extrinsic rewards; students could improve their project and their project grades based on what they learned from being a reviewer and from the reviews they received

References

1. Aalberg, T., Lorås, M.: Motivating and engaging students with peer review. MNT Conference **3**(1), 234–238 (2019)
2. Harland, T.: Learning about case study methodology to research higher education. Higher Education Research & Development **33**(6), 1113–1122 (2014)
3. Indriasari, T.D., Luxton-Reilly, A., Denny, P.: A review of peer code review in higher education. ACM Transactions on Computing Education (TOCE) **20**(3), 1–25 (2020)
4. Jönsson, A.: Facilitating students' active engagement with feedback. In: Lipnevich, A.A., Smith, J.K. (eds.) The Cambridge Handbook of Instructional Feedback, chap. 24, pp. 531–553. Cambridge University Press (2018)
5. MacLeod, L., Greiler, M., Storey, M.A., Bird, C., Czerwonka, J.: Code reviewing in the trenches: Challenges and best practices. IEEE Software **35**(4), 34–42 (2018)
6. Mercader, C., Ion, G., Díaz-Vicario, A.: Factors influencing students' peer feedback uptake: instructional design matters. Assessment & Evaluation in Higher Education **45**(8), 1169–1180 (2020)
7. Panadero, E., Jonsson, A.: The use of scoring rubrics for formative assessment purposes revisited: A review. Educational research review **9**, 129–144 (2013)
8. Panadero, E., Jönsson, A., Alqassab, M.: Providing formative peer feedback: What do we know? In: Lipnevich, A.A., Smith, J.K. (eds.) The Cambridge Handbook of Instructional Feedback, chap. 18, pp. 409–431. Cambridge University Press (2018)
9. Tai, J., Ajjawi, R., Boud, D., Dawson, P., Panadero, E.: Developing evaluative judgement: enabling students to make decisions about the quality of work. Higher Education **76**, 467–481 (2018)
10. Topping, K.: Peer assessment between students in colleges and universities. Review of Educational Research **68**(3), 249–276 (1998)
11. Turner, S.A., Pérez-Quinones, M.A., Edwards, S.H.: Peer review in cs2: Conceptual learning and high-level thinking. ACM Transactions on Computing Education (TOCE) **18**(3), 1–37 (2018)