DIFFERENCES OF ONLINE AND FACE-TO-FACE PEER REVIEWS REGARDING TYPE AND QUALITY

Christine Bauer, Kathrin Figl
Research Lab for Educational Technologies, University of Vienna
Rathausstrasse 19/9, 1010 Vienna, Austria

ABSTRACT

Peer reviews are used in a wide variety of disciplines. Nevertheless research investigating the impact of technology on peer-reviewing mainly derives from the field of writing classes. This paper presents an experimental study exploring the quality and kind of feedback given in a peer-reviewing task in an IT Project Management course. The study analyzes differences between the face-to-face and the online setting. The results show that students commented on fewer topics in the online version but described them in more detail than in the paper version. The online version was experienced as time-efficient and easy to fill out, while students found it significantly easier to express feedback in the paper version.

KEYWORDS

Peer reviews, online collaboration, online assessment

1. INTRODUCTION

Project-based teamwork and peer-reviewing are gaining importance as didactic techniques in computer science courses. Peer reviews have been used in a wide variety of disciplines, among them, accounting, engineering, mathematics, and social sciences (e.g., (Derntl, 2006, Dochy et al., 1999, Trahasch, 2004)). Interestingly, the most common use has so far been in writing classes. Main reason behind was the purpose to eliminate the need for instructors to read and grade hundreds of student essays (Gehringer, 2000).

However, peer-reviewing has benefits far beyond relieving the instructor's workload on grading. Through reviewing their peers' work, students develop evaluation skills, increase their reflection ability, and develop awareness of their own work's quality (Dochy et al., 1999, Trahasch, 2004). Of central concern in this research is determining how the face-to-face and the online setting impact the peer review outcome.

2. THEORETICAL BACKGROUND

2.1 Use of Peer Reviews

Peer Reviews in Writing and Language Classes. Dating back to the 1970s, writing instruction moved from a "product view" to a "process view" with multiple drafting. Feedback across the various drafts gained importance and peer reviews have become a significant activity in writing classes. Students take on the role of editors or tutors in providing information on each other's writing. There is a growing body of literature investigating the effects of peer- and instructor-reviews on revision (e.g., (Gehringer, 2000, Hanson, 2005)).

Peer Reviews in Software Development. Quality assurance plays a significant role in the software development cycle. Usually the term quality assurance evokes an image of executing software to see whether it functions as intended. An alternative form of quality control is to invite colleagues to examine your work in order to detect bugs and faults in software documents and code, and to find improvement opportunities: a peer review. While traditional testing is limited to executable code, peer reviews can be applied to any

software deliverable, design, or document at any stage of the development process. Hence, peer reviews have long been recognized as a powerful way to improve quality (Mashayekhi et al., 1993, Wiegers, 2002).

2.2 Positive Effects of Peer Reviewing

Studies reveal that peer reviews have many positive effects on students' learnings (cf. (Dochy et al., 1999, Trahasch, 2004)). Reviewing peers promotes the reflection on a student's own performance and may, hence, increase the awareness of the quality of the student's own work. In traditional course settings, students are often not interested in other students' contributions, as long as they get their own work right. When reviewing their peers, students are given the opportunity to learn to know work contributed by other teams (Derntl, 2006). Peer reviews help students to improve their own performances based on insights gained form their peers' comments (Gehringer, 2000).

2.3 Online versus Face-to-Face Peer Reviews

The few studies comparing traditional pen-and-paper and online peer reviews come mainly from the area of language and writing courses and typically concentrate on students' attitudes towards different modes of peer response, the impact of online peer reviews on the amount of participation by students, and on the effects of online peer reviews on the number of comments generated during the peer-reviewing task (Hanson, 2005).

In a study of (Liu and Sadler, 2003), the online peer review could count revisions although the face-to-face interaction seemed to be more effective than online communication. Concerning participation and interaction among peers, studies show inconsistent results. While (Sullivan and Pratt, 1996) report that participation in the online version was much higher than in the face-to-face version, (Huang, 1998) revealed that in face-to-face peer reviews more speech was produced and the full group participated more often compared to a synchronous online version. (Liu and Sadler, 2003) found that comments in online peer reviews were more revision-oriented than those in a paper version.

3. STUDY

3.1 Research Questions

Most previous studies analyzing differences of online and face-to-face reviews regarding the type and quality of reviews were undertaken in writing classes. This study will explore peer reviews in a completely different context – in an IT project management course. Existing studies notwithstanding, clear results on how review comments differ in face-to-face and online settings are still sparse. Therefore following research questions will be addressed: Do the online and the face-to-face version show differences in the amount of reviews written? How do reviews differ as regards content? Are there differences in the perceived review quality?

3.2 Design

The study was conducted in a blended learning course on "Project Management" for computer science students (Master program) in the winter term 2005/2006. Students were assigned to create and plan software development projects, which they had to elaborate in small teams of three persons. They had to elaborate documents like a mission statement of the project, an analysis of status quo and a to-be analysis, use cases and a project proposal, the product and project structure, an IT process model and structure, as well as MS Project charts. Since quality assurance and review processes were a significant part of this course's subject matter, respectively two (partner) teams were assigned to review one another.

In an early stage of the assigned projects, students reviewed their peers' project documents using a face-to-face peer-review form. The documents, which had to be reviewed, were provided printed out in class. Teams wrote comments about their respective partner team's project documents. The partner teams discussed the review results and, the reviewed team revised the documents accordingly. In a later project phase, this

peer-reviewing process was carried out online by a new combination of partner teams. Students used a web-based review tool that was designed in accordance with the paper version (Figl et al., 2006). For this review process, each team made its documents available on the course platform, such that the reviewer teams had online access to their partner teams' documents. In both versions about six documents had to be reviewed.

For investigating students' perceptions, students were asked to individually fill out a closed and open ended online questionnaire after they had completed both versions of peer reviews. The questionnaire, which was evaluated mainly statistically, included questions on review activities during the review task as well as questions on work efficiency and review quality. For comparing the online and paper reviews, text and qualitative content analysis were applied. In section 4.2, the formation of categories is described in detail.

3.2.1 Online-System

The web-based review tool that we used for the online peer review (Figl et al., 2006) was especially designed in accordance with the face-to-face version. It allows registering and managing deficiencies in a cooperative way and makes the progress of each team's project visible to all participants. Figure 1 gives a screenshot of the user interface for registering deficiencies.

The form contains following input fields: *Date, Time – Regarding* (both, date and time, can be set as well as under the heading "Regarding" the student project where the error has been found), *Author* (will be set automatically, as only logged-in users can add deficiencies), a *Short Description* of the problem that will be shown in the list of shortcomings (Figure 2), name of the *Document*, *Chapter* and *Page* number, a detailed *Description* of the deficiency, and *Additional Comments* that do not fit into any of the above fields.

For categorizing the deficiencies, we offered following categories in the tool: C3 – Severe (Rethink whole structure), C2 – Major – (Problem that has to be corrected), and C1 – Minor (Minor deficiency).

During the process of revision, users can set the problem-state of a reported deficiency. States in our configuration include: *Problem still exists*, *Problem has been solved*, *Duplicated* (the same problem already exists in the database), and *Interesting but subjective*.

Reported deficiencies are listed as can be seen in Figure 2. Every user can edit the deficiencies he/she reported; the administrator (instructor) can change any deficiency. The list also provides a link to the according comments section of a problem and a short description, which is clickable and links to the detailed description of the problem. For editing comments the simple WIKI syntax can be used. The comment system provides means to communicate with the partner team, and eventually close the problem after it is solved.

Figure 1. The interface for registering deficiencies

Figure 2. List of reported deficiencies

3.2.2 Face-to-Face Review Forms

In the face-to-face version, students reviewed their peers' documents in class. There was the possibility to discuss the reviews directly with the partner teams. We used a paper review form, the structure of which was in accordance with the online review form. Students had to fill in following information: reviewer team's name, project name that is reviewed, documents that are reviewed, detailed description of shortcomings found, classification of the shortcomings' severeness (C3 – Severe (Rethink whole structure), C2 – Major (Problem that has to be corrected), C1 – Minor (Minor deficiency)), overall judgment if a further review is necessary, and date as well as the reviewer team's signature.

4. RESULTS

4.1 Differences in the Number of Words

The number of words written by the students was counted for the face-to-face and the online version of the peer reviews. Results show that in the online version students wrote about twice as much as in the face-to-face version. Interestingly, students commented on much more different topics in the face-to-face version compared to the online one. Therefore the words per statement were much higher in the online version than in the face-to-face version. Results are summarized in Table 1. The number of words per statement was rounded to the next full number.

Discussion. The higher number of words and higher amount of words per statement in the online version could arise from the fact that the space for writing was limited in the paper review form, whereas in the online version it was unlimited. In the online version, teams could easily see how many comments other teams provided; probably therefore teams submitted about the same number of statements.

Table 1. Number of words in peer reviews

	Face-to-Face Version	Online Version	Total
Number of Words	194	386	580
Number of Statements	35	19	54
Average Number of Words per Statement	5	20	11

4.2 Differences in Review Categories

Positive and Negative Feedback. Table 2 gives an overview of the frequency of positive and negative feedback in the reviews. In general, students gave much more negative (74%) than positive (26%) feedback, meaning that they rather noted shortcomings instead of praising their partner team's work. Moreover, hardly any positive feedback was given in the online compared to the face-to-face version (2 versus 12 statements). Concerning the rated severeness of shortcomings, the online and the face-to-face setting do not differ.

Discussion. It does not surprise that there was much more negative than positive feedback. Actually it goes in accord with the online and paper forms. As the forms were called "list of shortcomings", it was probably clear to students that they should point out aspects their colleagues should improve. On the other hand, the instructor communicated to her students that they could give positive feedback as well.

Table 2. Positive and negative feedback

	Face-to-Face Version		Online Version		Total	
Positive Feedback	12	34%	2	11%	14	26%
Negative Feedback	23	66%	17	89%	40	74%
• Minor	14		13		27	
• Major	6		4		10	
• Severe	2		–		2	
• Missing	1		–		1	
Total	35	100%	19	100%	40	100%

Classification of Negative Feedback. For a content categorization of the negative feedback a classification scheme was developed inductively from raw data. Categorization was undertaken by two researchers in order to provide inter-subjectivity. In a first step, four different types of statements were classified. Shortcomings and suggestions were distinguished from each other, as well as categories for statements referring to a wrong level of detail and missing data. In a second step, all shortcomings and suggestions were classified into formal and technical ones. For "missing data" and "wrong level of detail" no further classification seamed reasonable. The main categories of review statements are described below:

- Level of Detail. The classification "level of detail" refers to students' judgments claiming that the reviewed documents were either too long or too short (miss important details). For example: "The technical specifications concerning the project implementation and the goals are quite a bit too short."

426

- Missing. The classification "missing" is used for all statements indicating that documents or parts of documents were missing in order to review them. For example: "We could not find any document or annotation, where the choice of the process model is mentioned or reasoned."
- Shortcomings (formal / technical). The classification "shortcoming" is used for the description of specific shortcomings. Basically we can distinguish formal and technical shortcomings.
  - "Formal" shortcomings refer to the structure, grammar and use of language in documents as well as aspects of providing documents in the right form (versioning, suitable document format). These are circumstantial aspects compared to "technical" shortcomings that are closely related to the project management and documentation skills, which students had to learn in the course.
  - The category "technical" shortcomings includes shortcomings regarding the real content of documents (e.g. correct and plausible description of the team's project) and also good planning of project plans or right modeling of unified-modeling diagrams. A reviewer team stated, for instance, "Unified Modeling standard was not followed (arrows, names of relationships)."
- Suggestions (formal / technical). Suggestions are not mere descriptions of shortcomings but include constructive hints for improvement. They are likely more revision-oriented than "shortcomings".
  - Suggestion (formal): "Make presentation more colorful: more color would give relief to it."
  - Suggestion (technical): "'Logging in' should be connected with 'extends' instead of 'uses'."

In general, the major part of negative feedback consisted of shortcomings (38%), followed by suggestions (25%), comments on the level of detail (20%), and on missing documents (18%). It appears that the percentage of noted shortcomings was lower in the online version (18% versus 52%) and for other types of feedback, the frequency was higher in the online version (Table 3).

Discussion. The fact that in the face-to-face version a stronger focus laid on shortcomings rather than on other types of negative feedback, could be due to several reasons. Since students evaluated projects in a later project phase in the online version, documents were probably already elaborated in more detail. This would explain that it was easier to give further suggestions. However, the categories "level of detail" and "missing" should also have declined. Another explanation could be that in the face-to-face version students had the documents printed out and could, possibly, therefore concentrate on noting shortcomings more easily. The finding that there were more suggestions than shortcomings in the online version could approve the findings of (Liu and Sadler, 2003) that online comments are more revision-oriented.

Table 3. Types of negative feedback

	Face-to-Face Version		Online Version		Total	
Level of Detail	4	17%	4	24%	8	20%
Missing	3	13%	4	24%	7	18%
Shortcomings	13	52%	3	18%	15	38%
• formal	5		1		6	
• technical	7		2		9	
Suggestions	4	17%	6	35%	10	25%
• formal	0		3		3	
• technical	4		3		7	
Total	23	100%	17	100%	40	100%

4.3 Differences in Work Efficiency and Review Quality (cf. (Figl et al., 2006))

For determining which setting for the peer reviews students rated better at specific dimensions, one-tailed Student's t-tests (n=16) with the reference value of 3.5 were calculated. The value "1" stands for the paper version, the value "6" for the online version.

Students found it significantly easier to express hints (M=2.63, SD=1.63, $t_{df=15}$=-2.15, α= 0.05) and to give helpful feedback (M=2.81, SD=1.33, $t_{df=15}$=-2.07, α= 0.06) in the paper version. Those students who are favoring the face-to-face peer review stated the possibility for direct communication with the partner team as the main reason. The results indicate that the paper version takes more time (M=2.75, SD=1.69, $t_{df=15}$=-1.77, α= 0.10). On the other hand, the online version can be filled out more easily (M=4.19, SD=1.47, $t_{df=15}$=-1.87, α= 0.08), which is also due to the fact that the documents can be better analyzed in digital form. Both versions were judged equally regarding the quality of feedback (Figure 3).
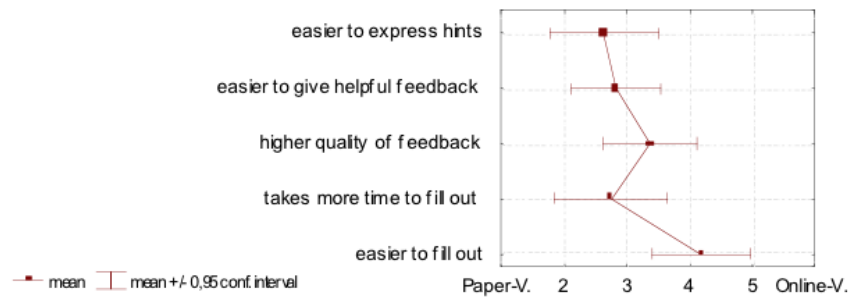
Figure 3. The process of the peer-reviewing activities in our course

5. CONCLUSION

This experimental study explored the quality and kind of feedback given in a peer-reviewing task. In particular, differences between the face-to-face and online setting were analyzed. Students commented on fewer topics in the online peer review but described them more detailed than in the paper version. There was hardly any positive feedback given in the online version, though in the paper version the ratio of positive comments was only slightly higher. The generally high frequency of negative feedback was due to the assignment since deficiencies had to be detected. Students did not perceive the shortcomings' severeness differently in the two different settings. Regarding the quality and kind of reviews, the versions did not differ either. Finally, the online version was experienced as time-efficient and easier to fill out.

The results of this study indicate that online reviews are as valuable as face-to-face reviews. Although data of our study suggest that there are hardly any differences between these settings, it has to be considered that only written statements were analyzed and interactive elements like teamwork in producing reviews and discussion about review results were not evaluated. However, these are additional factors, which impact the appropriate choice between the settings in a pedagogical situation. Therefore, further research needs to analyze the effects of possible choices on the number and quality of revisions and whether other elements of the peer-reviewing task (e.g., a follow-up discussion) should better take place face-to-face or online.

REFERENCES

Derntl, M. 2006. A Generic View on Web-Supported Peer-Evaluation, *Proceedings of ED-MEDIA*, Orlando.

Dochy, F., Segers, M. and Sluijsmans, D. 1999. The Use of Self-, Peer and Co-assessment in Higher Education: A review, In *Studies in Higher Education*, Vol. 24, No. 3, pp. 331-350.

Figl, K., Bauer, C. and Mangler, J. 2006. Online versus Face-to-Face Peer Team Reviews, *Proceedings of ASEE/IEEE Frontiers in Education Conference*, San Diego.

Gehringer, E. F. 2000. Strategies and Mechanisms for Electronic Peer Review, *Proceedings of ASEE/IEEE Frontiers in Education Conference*.

Hanson, J. G. 2005. Cooperative Learning Methods and the Teaching of English Writing: Peer Response, In *STETS Language & Communication Review*, Vol. 4.

Huang, S. Y. 1998. *A comparison between Chinese EFL students' peer response sessions held on networked computers and those held in a FF setting*, (Ed, Lansing, E.) National Center for Research on Teacher Learning.

Liu, J. and Sadler, R. W. 2003. The effects and affect of peer review in electronic versus traditional modes on L2 writing, In *English for Academic Purposes*, Vol. 2, pp. 193-227.

Mashayekhi, V., Drake, J. M., Tsai, W. and Riedl, J. 1993. Distributed, Collaborative Software Inspection, In *IEEE Software*, Vol. 10, No. 5, pp. 66-75.

Sullivan, N. and Pratt, E. 1996. A comparative study of two ESL writing environments: A computer-assisted classroom and a traditional oral classroom, In *System*, Vol. 29, No. 4, pp. 491-501.

Trahasch, S. 2004. From Peer Assessment Towards Collaborative Learning., *Proceedings of 34th ASEE/IEEE Frontiers in Education Conference*.

Wiegers, K. E. 2002. Seven Truths About Peer Reviews, In *Cutter IT Journal*, Vol. 15, No. 7.