Ethics of scientific peer review: Are we judging or helping the review recipients?

Oluwabunmi Adewoyin
Department of Computer Science
University of Saskatchewan
Saskatoon, Canada
Oluwabunmi.adewoyin@usask.ca

Julita Vassileva
Department of Computer Science
University of Saskatchewan
Saskatoon, Canada
jiv@cs.usask.ca

Abstract— Traditionally, ethics of a profession or organization are laid down by their pioneers, or subtly emerge over time as the organization advance. Getting conversant to these ethics requires teaching new or upcoming professionals, in order to avoid any form of misconduct, either deliberately or unknowingly. Peer review has been used as a quality control measure in the scientific community to ensure that only novel, high-quality and significant research work can be published. Typically, experienced and well respected scientists are selected to review the work of their peers or other upcoming scientists. Ideally, people who ethically qualify as reviewers, should have high reputation in terms of their ability to give objective and well-informed judgement, write constructive and helpful critique in a timely manner and, are honest and open in revealing any conflict of interest that may exist. The key objectives of peer review are two fold: 1) summative - to assess the quality of scholarly work, and 2) formative - to provide constructive feedback and thus, to mentor authors to become both better researchers, and better writers.

We focus on the second objective – the use of peer-review system to mentor authors to become better in their field of expertise and also to become better communicators. Unfortunately, often, reviews fail to accomplish this objective; the symptoms are short, brash, very critical reviews, which can be insulting and de-motivating for the authors. The reason for the increase in this kind of reviews is most often - the high reviewing load, often - conflict of interest, and finally - the lack of a feedback system that would allow poor-quality reviews to be penalized. Traditionally, the peer review procedure does not involve either the summative or formative feedback from authors to reviewers; it is up to the editorial staff (program chairs, journal editors) to determine the quality of the review. Often there is unofficial information kept or circulated about reviewers in terms of their typical review quality and timeliness, but these features refer to the summative component of the review, rather than to the formative, i.e. how helpful the review is for the author. Involving authors in evaluating the reviews received by their papers and considering this feedback in determining the reputation of reviewers would be a natural solution, since authors are in the best position to judge how helpful the review and the review session have been. A public reputation system for reviewers based on both the feedback from authors (regarding the formative aspects of the review) and from the editorial staff (regarding the summative aspects) will create an incentive for reviewers to write thorough, competent and constructive reviews, since they would be invited to review only if they have a high reputation.

We propose to modify the peer-review process by including a step requesting feedback from authors on the helpfulness of the reviews they got and another step, requesting feedback from reviewers about learning from each review session. We used the feedback to compute reviewers' reputation over time using the feedback from the authors. Making the reputation of reviewers visible to everyone without compromising the anonymity of peer-review encourages social comparison among reviewers and provides motivation to write constructive and helpful reviews. We ran a pilot study with an undergraduate class "Ethics in IT" for final year computer science (software engineering) students. As a part of their coursework, students submitted and peer-reviewed 7 writing assignments, with the dual goal to apply ethics principles to problems arising in the practice of IT profession and to improve their argumentation and writing skills. Our findings showed that students felt the feedback they received from their peers was better than the feedback from the teaching assistant and that their argumentation and writing skills have improved significantly. The students as reviewers found the feedback from the authors helpful in improving their skills for writing critical and constructive reviews. The students unanimously suggested that same peer-review procedure should be applied to coursework in other courses. The results of this pilot study suggest that it is promising to experiment with the proposed modified peer-review process not only in educational setting, but also in research reviewing, for conferences, workshops and journals. Conference management software tools augmented with the additional step of author feedback and reviewer reputation keeping can be used to both implement the process and provide research communities with the reputation systems they need to be able to select good reviewers and encourage reviewers to behave cooperatively.

Keywords—peer review; group formation; ethics and IT; ethics education

I. Introduction

Peer review has been part of the culture in the science research community since the 1600s [9, 14, 21]. It is used to evaluate project proposals in order to ensure that research funds are allocated to competent researchers and novel ideas. Also, it is used in critiquing journal and conference papers, in order to ensure that only high-quality research (with novel and

significant ideas and good technical execution) are published [2]. The peer review process comprises the authors on one side, and the reviewers on the other side [17]. When authors submit their papers for review, the papers are assigned to reviewers by the programme committee based on their expertise, considering their level of experience. Once reviewers have completed reviewing the submissions, authors are given access to the reviews of their papers, accompanied with the decision made by the program chairs to either accept or reject their papers. In some situations, authors might be given the opportunity to write a rebuttal, to address some criticisms in the reviews, to strengthen their case for the acceptance of their papers, before the final decisions are made. Peer reviewing can be open, single blind, or double blind. Open peer review is when authors and reviewers' identities are not concealed. Single blind peer review is when the reviewers' identities are concealed from the authors, but authors' identities are known to the reviewers, while double blind peer review requires the concealment of the identities of both the authors and the reviewers [3]. In the past, authors had to send their manuscripts by mail to the conference organizers or journal editors, who, after assigning reviewers also sent the manuscripts to these reviewers for review by mail [15]. The reviews were sent by mail back to the editor and from there – to the authors. Therefore the process took very long time. Currently the process is supported by web-based manuscript management systems, which enable paper submission, review assignment and the subsequent reviews to be done very efficiently. Popular existing manuscript management systems include EasyChair, Precision Conference and OpenConf, Manuscript Central, among others.

Despite the wide acceptance of peer review in the research community, it has been criticised by some researchers as being slow, expensive, and plagued with bias and some inconsistencies [16, 3]. It is believed that reviewers take a long time to turn in their reviews because reviewing takes substantial time that they can spend on their own research, and there is no particular incentive for reviewing (no payment). For the same reasons, some reviewers do very superficial reviews. In addition, reviewers are often inconsistent in their judgement of certain papers and research work. The research interests and knowledge of reviewers influence strongly their opinion about the papers being reviewed. Conflict of interest can motivate some reviewers to suppress the publication of a competing researcher to ensure their own advantage in publishing first the idea.

In other research, authors' perception and the success of the peer review process are determined by whether their papers are accepted or not [4, 22]. This is one-sided, because the decision to accept paper does not necessarily reflect the quality and thoroughness of the reviews. The major objectives of peer reviews are - to access the quality of scholarly work and to give constructive feedback to authors. We believe the feedback can be judged as constructive only by the review recipients through the back-evaluation of their reviews. The traditional peer review process does not include back-evaluation of reviews by the authors. Instead, the program chair, committee, or the editorial team only have the chance to evaluate the quality of the reviews and in some extreme cases, to disqualify the reviews or the reviewer [3, 14]. Yet the author is in the best

position to judge whether the reviews are constructive and helpful. If authors lack the opportunity to evaluate the reviews of their papers, reviewers cannot receive feedback that could motivate them and help them improve the quality of their reviews. Including authors' feedback and evaluation of the reviews they received in the peer-review process will support both of its main objectives, evaluating the quality and providing constructive feedback to scholarly work, since scholarly work involves not only creating new research, but also being able to competently and constructively criticise the research of others.

In this paper, we propose an approach to peer review whereby authors are given the opportunity of evaluating the reviews they have received and giving feedback on the peer review session in a double-blind review process; the reviewers also give their feedback on the peer review session. We applied experimentally the approach in an undergraduate computer science class of ethics and IT; and the approach was meant to provide useful conclusions on both the formative and summative objectives of the peer review process. We proposed the following research questions:

1. Does the modified peer review process serve well the formative objective to produce high-quality, helpful and constructive reviews?

2. Did the modified peer-review process motivate reviewers to give constructive and helpful reviews?

3. Is it essential to use a double blind review process as a basis?

The rest of this paper is organised as follows. Section 2 contains the review of related work on peer review and ethical issues involved. Section 3 comprises the discussion of our methodology, while section 4 contains the discussion of the results of our experiments and section 5 concludes this paper.

II. RELATED WORK

The rules and norms related to scientific peer review have evolved over the years, driven by the desire to improve the quality of the reviewing process. Yet this goal seems to be hard to achieve. First, it is not clear how to define the quality of reviews. Possible desirable features are: competent, thorough, unbiased, and constructive [25]. Rockwell (2006) discusses some of the problems with the rules and ethical issues underlying the scientific peer review [15]. For example, the choice of reviewers by the journal editors of conference chair is limited by their expertise and network of colleagues. The anonymity of authors can also be compromised in these ways in double-blind review. The quality of reviews may suffer from bias, especially when the reviewer's own research is very close to the one under review, which is often the case, since the selection of reviewers is mostly based on their expertise in the area of the reviewed paper. Differences or similarities in opinion, research methodology or philosophy may bias the review in either direction [12]. Peer review can be affected also by unethical behavior on the side of the reviewers, who fail to disclose conflict of interest or bias that they are aware of to the program chairs or editors. . Also,

individual differences of using rating schemes by reviewers can bring noise in the process. Some reviewers give contradicting feedback, which put editors in dilemma on whether to accept or reject the paper and also can result in putting authors in dilemma of what is wrong and right with their writing and research outcome [15, 16].

To encourage reviewers to provide honest feedback, *blind review* process has been used frequently, which allows the reviewers to remain anonymous. Van Rooyen *et* al. (1998) and Justice *et* al. (1998), in their research on the effect of blinding reviewers' identities while leaving authors' identities open, found out that anonymity does not necessarily improve reviews' quality [8, 18]. However, their measure of quality of reviews was based on the judgment of the editors and also validated by the review quality instrument by [13]. Justice *et al*. (1998) goes further to argue that the lack of improvement in the quality of review when authors' identities are not known to reviewers is due to the fact that reviewers' identities are open, which might have a significant impact on reviewers' judgment of their papers most especially for the reputable authors in the field [8]**.** In comparison with blind review, *open review process*, where the identities of both author and reviewer are known to each other, does not affect quality of reviews but increases the likelihood of reviewers declining to review [19]. In the *Double blind review*, both reviewers' and authors' identities are not revealed to each other [3]. However, Hill and Provost (2003) show that authors in a double-blind review can be identified by their citations [6]. Therefore, it is unlikely that double-blinding will improve the quality of reviews because it is easy to identify the research interest and the writing style of well-known authors and reviewers [8].

A possible measure for the review quality is the author's satisfaction with the review. Satisfaction shows how the proposed system has been able to meet the expectation of the users, while helpfulness is the congruence between their expectation and their experience with the system [11].In the domain of conference peer review, the author's satisfaction with the review can be influenced by the decision of acceptance or rejection of their paper, i.e. their satisfaction may not be associated with the review quality, which is a measure of helpfulness and learning from the reviews [22]. In the learning domain there is no acceptance or rejection decision, and the peer review is not used for grading, so we don't expect that there was a bias in the users' satisfaction. Lan *et al*. (2011) proposed a five stage peer review process which comprises online writing, peer feedback, feedback of feedback, rewriting and publishing, with the hope that the inclusion of the feedback of feedback will help to increase the authors' satisfaction with their reviews [10]. Hart-Davidson *et al*. (2010) also included authors' feedback on their reviews as a way of measuring the helpfulness of and their satisfaction with the reviews of their papers [5]. Yet, the author's perception of the review process is usually subjective and, according to [22] and [4], depends on whether their papers were accepted or not, in the case of a conference or journal peer review process. We

believe that these researches only discuss the reviewers' and editors' aspect of the peer review process, which is to accept or reject papers depending on their perception of the quality of these papers. Therefore, they did not show the relationship between author's satisfaction and the review quality. Hart-Davidson *et al*. (2010) designed a conference management called "Eli", which they use for linguistic analysis of reviews and generate a helpfulness score for every reviewer [5]. Xiong and Littman (2011) also used natural language processing (NLP) to automatically predict the helpfulness of peer reviews, by checking how reviewers were able to localize the problems found in the papers that they are given to review [23].

In summary, the existing research on peer review in science mostly discusses the faults in the existing peer review processes, such as blind, double-blind and open review focusing mostly on whether these processes support fair, unbiased and competent reviews. Yet, there is no agreed upon definition of what constitutes a high quality review, and the criterion considered so far has been the author's satisfaction in general (with the process and fairness). Only recently the helpfulness of the review in improving the author's research paper has been considered. The existing literature lacks a discussion of the factors that can motivate reviewers to improve the constructiveness of their reviews, or how incentives for reviewers can be included in the reviewing process. Also, appropriate measures of the success of a peer review session from –point of view of both reviewers and editors on one hand, and authors on the other hand need to be developed.

III. APPROACH AND METHODOLOGY

To address the shortcomings of described of existing approaches, described in the previous section, we propose a modified double-blind peer review process, where the authors are given the opportunity of evaluating the reviews they have received in a review process; in addition both the authors and the reviews provide their evaluation and feedback on success of the peer review session. We propose to evaluate this process in a learning context, where the formative role of peer-review is particularly important. We chose a fourth year undergraduate computer science class on Ethics and Information Technology. The class has a dual purpose of teaching students the principles of ethical reasoning and argumentation, and to help them improve their writing skills. The coursework required from the students in the class involves multiple short writing assignments. This type of assignments naturally yield themselves to peer-review, due to the common question / case assigned for discussion to everyone in the class, and the moderate size of the expected text (300-500 words), so writing an essay and reviewing 2 or 3 essays by peers is not too onerous task and can be done weekly. Since reviewing peer's assignments is time-consuming, it is important to ensure that students have motivation to review the work of their peers carefully and write constructive feedback. Our hypothesis is that:

A peer-review process that allows authors to provide feedback to the reviewers on the competence, constructiveness and helpfulness of their reviews will:

1) *Provide feedback that will help reviewers improve their critique skills*

2) *Provide motivation for reviewers that will lead to higher quality reviews*

3) *Allow authors to develop better writing skills due to the higher quality reviews.*

Paper assignment is an important step in the peer-review process. Typically in scientific reviewing, information about the reviewers' interests and areas of competence, as well as their seniority level, are used to select reviewers for each paper. In a learning domain, however, the peers are students with initially unknown skills. An initial writing assignment is used to elicit information about the students' writing and argumentation capabilities. The students are divided into 3 categories (advanced, medium and weak), and in the next assignment, peer review groups are formed taking care to have members of each of the three categories of expertise in each group. In this way the peers with relatively weak writing skills will learn from reviewing the papers of their advanced peers, and will receive valuable feedback from them as reviewers. The advanced writers will learn from helping their weaker peers to provide constructive criticism and will be able to improve their critique skills. The reviewing groups are changed every week, ensuring the presence of at least one member of each category (if there were sufficient number of students) in a group. In addition to the peer-review, all students' papers and reviews were reviewed by the teaching assistant (TA), who serves as an expert in the ethics domain and in writing. The grades obtained by students for their essays each week were used as a basis to categorise and reassign the students to reviewing group in the next week. Fig.1 shows the procedure followed during the peer-review process.
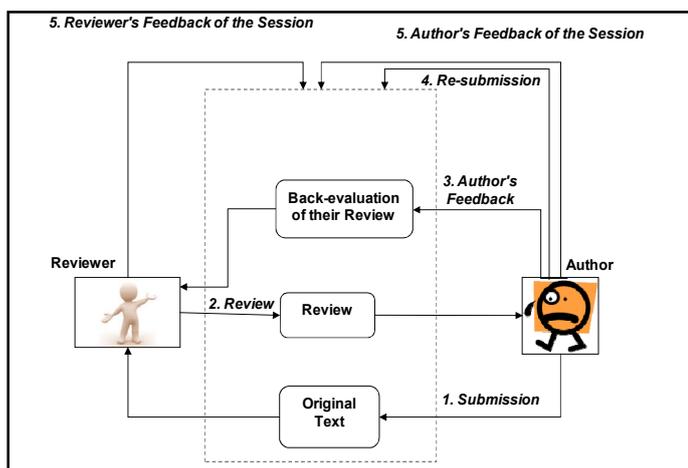

Fig. 1: Modified peer review process

Seven reviewing sessions were organized over 7 consecutive weeks during the winter term (2nd term, 2012/2013). In six of them, double-blind reviewing process was followed. We hypothesized that the double blind process will ensure unbiased reviews and counter possible gaming. Yet, we wanted to give the participants a stable identity to allow them to build a reputation as reviewers based on the evaluation of their reviews by the authors. Also we wanted the students to be able to distinguish the reviewers of their papers, even if they didn't know their real identities. Therefore we asked them to use pseudonyms.

In the middle of the 7 weeks period, in week 4, an open review process was followed, to test the importance of authors and reviewers' anonymity in the process.

We used the EasyChair conference management system, with some added features (e.g. authors' feedback, authors' and reviewers' grand feedback) that were implemented, through a separate website. Every week, students were given a different question or case to discuss in a short essay (500 words), which had to be submitted by the end of Tuesday in the peer-review system. The papers were assigned for review to the two peers from the group of the author and the author was assigned as reviewer to the papers submitted by the other two group members. The reviews were due by the end of Thursday. On Friday, the authors were invited by email to see the reviews of their papers and evaluate them, providing feedback to the reviewer regarding the helpfulness, constructiveness and competence of the review. The authors' feedback was sent out to the reviewers in order to know what authors felt about the reviews they provided. The email contained also an invitation for the students to check the other two reviews of the papers that they were assigned to review, and to compare their own reviews of the paper with the other two reviews. The reviewers could also engage in discussion with the other reviewers. They were also invited to provide more general comments about how much they learned from the peer review session, both as authors, and as reviewers. The final paper was due on Sunday. At the end of the term, the students were asked to evaluate their overall experience with the peer review process, both as authors and as reviewers. All questionnaires that were used are provided in the Appendix.

There were six students enrolled in the Ethics and IT Class (2012/2013, T2); all of them were male Canadians aged between 20 and 30. All of them participated voluntarily in the study and signed a consent form to permit their data to be used for this research. Every week, they were divided in two groups that reviewed each other's papers, based on their writing scores from the previous week (evaluated by the teaching assistant, not by peer-review). The data from the peer-evaluations, the authors' evaluations of reviewers, and the overall evaluations of each reviewing session were collected and analysed. The results of our experiment are discussed in the next section.

IV.   RESULTS AND DISCUSSION

We present the results below as they relate to the research questions defined in section 1.

1. Does the modified peer review process serve well the formative objective to produce high-quality, helpful and constructive reviews?

To answer this question we asked two kinds of questions: first, questions related to the quality of reviews they received, and second, questions related to how much they felt they learned and improved their skills, both as writers and as reviewers.

In the exit questionnaire (see the Appendix), the students provided feedback regarding their satisfaction as authors with the reviews they got. All questions measured satisfaction on a scale from 1-10 (1-lowest, 10-highest, see Table 1). The feedback showed that authors are, on the average, satisfied with their reviews. The highest mean satisfaction (8.50) was with "good points raised" by the reviews, with useful suggestions about style and grammar (8.33) , with the suggestions provided in the reviews on good ways of expressing ideas (8.17). The mean satisfaction with the friendliness of the reviews was average (5.17). The standard deviations for these means are relatively low.

Table 1: Satisfaction levels of authors with the reviews they got (agreement with the criteria on the scale of 1 (lowest) to 10 (highest))

Criteria	Mean	Standard deviation
Friendliness	5.17	2.86
Good points raised	8.50	1.34
Useful suggestions to style and grammar	8.33	1.63
Suggestions on good ways of expressing ideas	8.17	2.23
Lack friendliness	3.67	2.07
Lack good points in the review	4.00	3.03
Lack useful suggestions to style and grammar	4.00	3.69

These positive results are confirmed by the test questions intended to verify the truthfulness of the answers given previous questions. The questions were phrased in reverse, and lower level of agreement should be interpreted as a positive opinion regarding the corresponding quality of the reviews. The results show that the authors have lower levels of dissatisfaction with the friendliness, points and suggestions raised in the review of their papers (see rows 4, 5 & 6 in Table 1). The standard deviation for these answers is relatively high, however, which probably indicates some confusion in the students with the purpose or phrasing of the test questions.

These generally positive results about the authors' satisfaction with the reviews they received show that the reviewing assignments were done well, and that our method of grouping peers with diverse levels of writing capabilities was successful. We believe that this approach ensured that peers as authors are matched with capable peers as reviewers, and that peers were motivated to work hard both as authors, to get good feedback, and as reviewers, to provide helpful reviews. Since the goal of this class is to help students improve their writing and critique skills, the final questionnaire contained a question asking if the participant thought that their writing and critique skills have improved as a result of the peer review process. The results show that 50% of the peers believed that their writing skills have improved, and 86% of them believed that their critique skill has also improved as a result of the peer review.

As a result of this, they all agreed that the course work involving writing assignments using peer review should continue in the Ethics and IT class. The majority of the students, 83% of them, suggested that the peer review process should be used in other classes. We allowed reviewers to be able to view other reviews of the same papers that they review. 71% of the students stated that they improved their reviewing and critiquing skills from being able to view the other reviewers' feedback of the same paper they review. One of the reviewers provided the comment quoted below, which shows that they did learn from seeing other reviewers' comments / feedback.

" *...sometimes another reviewer would make a good point that i had missed or take on an entire other perspective so it was nice to see...*"

An analysis of the students' activities in the conference management system showed that in 4 out of 7 review sessions, reviewers tend to make changes to their reviews after seeing other reviewers' feedback, which is also an evidence of their learning from the reviews of others. We also did an analysis of their review grades from the expert's judgement (TA) and we found out that the average mark recorded in the class was 4.67 (out of 5) in the last peer review assignment. This shows a significant improvement in their reviewing skill when compared with the average performance of 4.17 (out of 5) recorded in their first peer review assignment.

2. Did the modified peer-review process motivate reviewers to give constructive and helpful reviews?

The modified process included explicit feedback from the authors on the quality and helpfulness of the reviews they received. In the exit questionnaire (see Appendix), the reviewers were asked questions about what motivated them to give constructive and helpful reviews. The results show that 50% of the reviewers were motivated by the feedback given by the authors of the papers that they reviewed. The motivation of these reviewers may be explained with seeking reputation as good reviewers (based on the opinion of fellow colleague), or self-efficacy (seeking to hone their skills, overcoming challenges), according to the spectrum of motivations, described in [20]. Since reputation is a powerful motivator which can be used also as a personal achievement measure for self-efficacy driven reviewers, a design recommendation for peer-review systems using the modified process is that their reviewers interface should maintain a reputation system and public visualization of reviewers' reputation. This would motivate reviewers to give thorough and helpful reviews since they would not like to soil their reputation among their peers.

We believe that this motivator will be powerful not just in learning applications of peer review, as the one presented in this study, but would generalize also for the scientific peer-review process, since the reputation of reviewers with their peers, program chairs and journal editors is important to ensure future invitations to program committees and editorial boards.

In addition to the 50% of students motivated by reputation, one-third of the participants were motivated to achieve high marks on their reviews from the teaching assistant (extrinsic motivation), while 16.67% stated that they were motivated by

just being helpful to their peers (intrinsic motivation or seeking reciprocity with their peers). This shows that even in a small user group, there is a wide diversity of motivations to write constructive reviews.

3. Is it essential to use a double blind review process as a basis?

We had seven peer review sessions, in 7 consecutive weeks. Blind review with pseudonyms took place in six of the seven sessions, $RS_1 - RS_3$, and $RS_5 - RS_7$. In RS_4, an open review was carried out. To investigate the effect that anonymity had on the quality of reviews and the quality of feedback given by the authors, we present in the middle column of Table 2 for each session the average satisfaction of all the participants with the reviews that they received immediately after this session. In the right column of Table 2 we show the average satisfaction of the participants with each review session, but according to the final questionnaire. Not surprisingly, the students are more positive in their final, retrospective evaluation, which is probably positively biased due to the overall positive experience and the elapsed time that normally filters out negative experience. Remarkably, none of the authors filled the satisfaction feedback for RS_4 immediately after the session. The average value of student satisfaction for this session shows that the participants were least satisfied with the reviews they got from this session (see third column for RS_4 on Table 2).

Table 2: Authors' satisfaction with their reviews (%)

Review session (RS)	Weekly questionnaire (% Satisfaction)	Exit questionnaire (% Satisfaction)
RS₁	72	85
RS₂	80	85
RS₃	75	85
RS₄	-	70
RS₅	62	85
RS₆	52	86
RS₇	80	90

These results speak clearly of the advantage of double blind review. While peers' identities were pseudonymous, they were objective and thorough in the reviews of their peers' papers. However, in the fourth week when they used their real identities, the students felt obliged to give each other bland and polite reviews, which eventually resulted in the authors feeling that the reviews they received were not thorough and helpful.

V. CONCLUSION

We analysed the current peer-review systems with respect to their formative aspect, that is, how they ensure high quality, constructive reviews that help both authors to improve their work and reviewers to improve their critique skills and we found them lacking. In response, we proposed a modified double-blind (pseudonymous) peer review process that allows authors to provide evaluation and feedback to the reviewers of their papers. A reputation value of reviewers is calculated based on these evaluations and presented to the reviewers. We implemented the proposed process in a undergraduate Ethics and IT class that has as one of its objectives developing argumentation and writing skills and involves numerous writing assignments. The choice of a learning domain rather than science research reviewing (conference or journal) was driven by the fact that the formative aspect is central in learning domains, while the summative aspect (evaluating papers) is secondary. Also it allowed to carry out many reviewing sessions with the same participants and to compare them, which is much harder to do in a conference review domain. Our results, even though with a very small number of participants, show that the modified process worked very well and ensured high satisfaction of the students with their learning both as authors and as reviewers. Our results show that the proposed method of matching reviewers and authors in smaller groups based on their skill levels leads to a general increase in the quality of the reviews by the reviewers and their subsequent helpfulness to the authors.

Our results show that double blind review is essential to ensure honest and critical reviews, but pseudonyms are important to ensure a permanent identity of reviewers and allow them to build reputation. These results, even though demonstrated in the learning domain, we believe will hold also for scientific reviewing, especially considering that popular online peer review systems, such as EasyChair, allow the same set of scientists to review and author papers for many different conferences, to have permanent identities and possibly, to build on their reviewer reputation over years. We believe, this would present a strong motivation for reviewers to invest more time and effort and write more constructive reviews that fulfil the formative objective of peer-review.

Directions for future work include evaluating the modified peer-review process in larger educational settings and in scientific review (e.g. the TRUM (Trust Reputation and User Modeling) workshop series organized by the authors at the UMAP (User Modeling, Adaptation and Personalization conferences).

REFERENCES

[1] Casati, F.; Marchese, M.; Mirylenka, K. and Ragone, A. (2010). Reviewing peer review: a quantitative analysis of peer review. Technical Report # DISI-10-014.

[2] Chubin, D. and Hackett, E. (1990). Peerless Science: Peer Review and US Science Policy. Albany: SUNY Press.

[3] Do, E. Y. (2003). Afterword- Why peer review journals? International journal of architectural computing, issue 02, volume 01, *pp.* 253-264.

[4] Gibson, M.; Spong, C. Y.; Martin, S. And Scott, J. R.: Author perception of peer review. Obstetrics and gynecology, Volume 112(3), pp. 646-652 (2008).

[5] Hart-Davidson, W.; McLeod, M.; Klerkx, C. and Wojcik, M. (2010). A method for measuring helpfulness in online peer review. *SIGDOC 2010*, October27-29, 2010.

[6] Hill, S. and Provost, F. (2003). The Myth of the Double-Blind Review? Author Identification Using Only Citations. SIGKDD Explorations, Volume 5, issue 2, pp. 179-184.

[7] Jefferson, T.; Wager, E. And Davidoff, F. (2002). Measuring the quality of editorial peer review. The journal of American medical association (JAMA), volume 281, number 21, pp. 2786-2790.

[8] Justice, A. C.; Cho, M. K.; Winker, M. A.; Berlin, J. A.; Rennie, D. And Peer Investigators (1998). Does masking author's identity improve the quality of peer review? A randomised controlled trial. The journal of American medical association (JAMA), Volume 280, number 3, pp. 240-242

[9] Kronick, D. A. (1990). Peer review in 18th century scientific journalism. Journal of American medical association (JAMA), volume 263, pp. 1321-1322.

[10] Lan, Y.; Wang, J.; Hsu, S. and Chan, T. (2011). Peer feedback in online writing system. In: M. Chang et al. (Eds.): Edutainment 2011, LNCS 6872, pp. 126–129, 2011.

[11] Latu, T. M. and Everett, A. M. (2000). Review of satisfaction research and measurement approaches. Science and research internal report 183, department of conservation, Wellington, New Zealand.

[12] Mahoney, M. J. (1977). Publication prejudices: an experimental study of confirmatory bias in the peer review system. Cognitive therapy and research, volume 1, number 2, pp. 161-175.

[13] McNutt, R. A.; Evans, A. T.; Fletcher, R. H. and Fletcher, S. W. (1990). The Effects of Blinding on the Quality of Peer Review: A Randomized Trial. Journal of American medical association (JAMA), volume 263(10), *pp.* 1371-1376

[14] Ranalli, B. (2011). A Prehistory of Peer Review: Religious Blueprints from the Hartlib Circle. Spontaneous Generations: A Journal for the History and Philosophy of Science, Vol. 5, No. 1 (2011) pp. 12-18.

[15] Rockwell, S. (2006). Ethics of Peer Review: A Guide for Manuscript Reviewers. URL: http://ori.dhhs.gov/education/products/yale/prethics.pdf. [Accessed on: 01-09-2013].

[16] Smith, R. (2006). Peer review: a flawed process at the heart of science and journals. Journal of the royal society of Medicine.

[17] Spier, R. (2002). The history of the peer-review process. Trends in Biotechnology, Volume 20, Issue 8, Pages 357–358

[18] Van Rooyen, S.; Godlee, F., Evans, S.; Smith, R. And Black, N. (1998). Effect of blinding and unmasking on the quality of peer review. The journal of American medical association (JAMA), volume 280, pp. 234-237.

[19] Van Rooyen, S.; Godlee, F.; Evans, S.; Black, N. And Smith, R. (1999).Effect of open peer review on quality of reviews and on reviewers' recommendations: a randomised trial. BMJ, Volume318, pp. 23-27.

[20] Vassileva, J. (2012) Motivating Participation in Social Computing Applications: A User Modeling Perspective. User Modeling and User-Adapted Interaction, 22 (1-2),177-201

[21] Wagner, W. And Steinzor, R. (2006). Rescuing Science from Politics Regulation and the Distortion of Scientific Research. Cambridge University Press, 1st edition

[22] Weber, E. J.; Katz, P. P.; Waeckerle, J. F. And Callaham, M. L. (2002). Author perception of peer review: Impact of review quality and acceptance on satisfaction. The journal of American medical association (JAMA), volume 28, number 21, pp. 2790-2793.

[23] Xiong , W. and Litman, D. (2011). Automatically predicting peer-review helpfulness, Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies: short papers, June 19-24, 2011, Portland, Oregon

[24] Xiong, W.; Litman, D. and Schunn, C. (2010). Impact of Annotation Difficulty on Automatically Detecting Problem Localization of Peer-Review Feedback. In: Workshop on Computer Supported Peer Review in Education, Intelligent Tutoring System Conference (ITS) 2010.

[25] Evans, A. T.; McNutt, R. A.; Fletcher, S. W. and Fletcher, R. H. (1993). The characteristics of peer reviewers who produce good quality reviews. Journal of general internal medicine, volume 8, pp. 422-428

WEEKLY RATING OF REVIEW - ETHICS AND IT

Please rate the reviews you received in the form below

1. Please enter your name *

2. As an author, please indicate the helpfulness of each review you got on your essay *
   Please provide the option, from 1 to 5, that best describes the helpfulness of the review you got (1 is the least, 5 is the best)

	1	2	3	4	5
Reviewer 1	☐	☐	☐	☐	☐
Reviewer 2	☐	☐	☐	☐	☐

   Comments

3. As a reviewer, did you look at the other reviews of the essays you were assigned? *

   ☐ Yes   ☐ No

4. If applicable, how helpful did you find the other reviews?
   Please rate on a scale of 1 to 5, how helpful you think was the other review of the essay (1 is the least, 5 is the best)

	1	2	3	4	5
Paper 1	☐	☐	☐	☐	☐
Paper 2	☐	☐	☐	☐	☐

5. How did you learn from your experience as the reviewer? *
   Please provide the option, from 1 to 5, the best that describes your learning experience from each paper you reviewed (1 is "Not at all" 5 is "A lot")

	1	2	3	4	5
Paper 1	☐	☐	☐	☐	☐
Paper 2	☐	☐	☐	☐	☐

FINAL SURVEY-ETHICS AND IT

This is the final survey to evaluate your overall experience with the peer-review coursework used in the CMPT 408 class. The survey contains 29 questions and requires about 20 minutes to complete. Thank you for your time and feedback.

1. Please enter your name *

2. As an author, what is your general impression of the reviewers? *
   Please select all relevant options

   ☐ Competent

   ☐ Constructive

   ☐ Provided detailed suggestions

   ☐ Provided useful correction to style and grammar

   ☐ Too polite

   ☐ Not too helpful

   ☐ Not too thorough

   ☐ Too negative

   ☐ Lacked substance

   ☐ Didn't always understand the points I was trying to make

3. Did you notice any difference in the reviews you got over time?

   ◉ Yes

   ◉ Yes

   ◉ Yes

4. Do you think the competence of the reviewers improved over time? *

   ◉ Yes

   ◉ No

   ◉ Hard to say

   Comment

5. Do you think your essays were reviewed by the same set of reviewers over the weeks? *

   ◉ Yes

   ◉ No

⬚ Hard to say

Comment

```

```

6. How do you think the assignment of essays to reviewers was done? *

⬚ Based on some principle

⬚ Randomly

⬚ Hard to say

7. If you chose "Based on some principle", can you guess the principle that the assignment was based on?

```

```

8. Please rate the reviews you got from the following essays on the scale of 1 to 10
(1- very poor, 10 – excellent)

	1	2	3	4	5	6	7	8	9	10
Privacy in Social Networks	⬚	⬚	⬚	⬚	⬚	⬚	⬚	⬚	⬚	⬚
Intellectual Property	⬚	⬚	⬚	⬚	⬚	⬚	⬚	⬚	⬚	⬚
Games and Gamification	⬚	⬚	⬚	⬚	⬚	⬚	⬚	⬚	⬚	⬚
The Future of Education	⬚	⬚	⬚	⬚	⬚	⬚	⬚	⬚	⬚	⬚
Cybercrime and Cyber Security	⬚	⬚	⬚	⬚	⬚	⬚	⬚	⬚	⬚	⬚
Evaluating and Controlling Technology	⬚	⬚	⬚	⬚	⬚	⬚	⬚	⬚	⬚	⬚
Final Paper	⬚	⬚	⬚	⬚	⬚	⬚	⬚	⬚	⬚	⬚

9. Why did you not evaluate the review you got from the 4th essay? *
(4th essay was "The Future of Education" from 11th to 17th February 2013)

```

```

10. Please rank the following, on the scale of 1 to 10, according to their degree of relevance in the choice of the rating you gave to the review of your essay *
(1- very poor, 10 – excellent)

	1	2	3	4	5	6	7	8	9	10
Friendliness	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Good points raised	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Useful corrections to style and grammar	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Suggestions on good ways of expressing ideas	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Lack friendliness	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Lack good points in the review	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Lack useful suggestion on ways of expressing ideas	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

11. As a reviewer, what factor(s) was more important for you to improve your reviews?

☐ Feedback from the authors of the essays you reviewed

☐ Marks given by the marker

☐ Just trying to be helpful to my colleague

☐ Neither

12. If neither of the above was a motivation for you, what motivated you to give good reviews?

13. What do you think of the quality of the reviews you wrote over the weeks? *

☐ Improved

☐ Stayed the same

☐ Got worse

Comment

14. Can you compare the feedback from the authors of the essays you reviewed and the marks given by the marker? *

    ☐ Similar

    ☐ Not similar

    ☐ Hard to say: some were similar, some were not

15. In the case that you answered "Not similar", or "Hard to say:", which feedback did you think was fairer?

    ☐ Feedback from authors

    ☐ Marks given by the marker

    Comment

    ```
    ┌─────────────────────────────┐
    │                             │
    │                             │
    │                             │
    │                             │
    └─────────────────────────────┘
    ```

16. Did you check the reviews provided by the other reviewers of the same essays that you reviewed? *

    ☐ Yes

    ☐ No

    Comment

    ```
    ┌─────────────────────────────┐
    │                             │
    │                             │
    │                             │
    │                             │
    └─────────────────────────────┘
    ```

17. If "Yes", did you feel that you learned something from the reviews given by others?

    ☐ Yes

    ☐ No

    Comment

    ```
    ┌─────────────────────────────┐
    │                             │
    │                             │
    │                             │
    │                             │
    └─────────────────────────────┘
    ```

18. Did you ever engage in a discussion with the other reviewers about the essays you reviewed? *

    ☐ Yes

    ☐ No

Comment

```
┌─────────────────────────┐
│                         │
│                         │
│                         │
│                         │
└─────────────────────────┘
```

19. If "Yes", what did you discuss with other reviewers?
    (Please select all relevant options)

    ☐ Understanding the essay

    ☐ Emphasizing specific points

    ☐ Grammar

    ☐ Style of writing

    ☐ Other   ┌─────────────────────┐
                       │   │
                       └─────────────────────┘

20. If "No", why didn't you discuss?

```
┌─────────────────────────┐
│                         │
│                         │
│                         │
│                         │
└─────────────────────────┘
```

21. How did you decide what rating to give to an essay based on the issues you raised in the review? *

```
┌─────────────────────────┐
│                         │
│                         │
│                         │
│                         │
└─────────────────────────┘
```

22. Do you think your writing skill has improved as a result of this type of coursework? *

    ⬡ Yes

    ⬡ No

    Comment

```
┌─────────────────────────┐
│                         │
│                         │
│                         │
│                         │
└─────────────────────────┘
```

23. Do you think your skills as a critic/reviewer have improved as a result of this type of coursework? *

    ⬡ Yes

    ⬡ No

    ⬡ Other   ┌─────────────────────┐
                       │   │
                       └─────────────────────┘

24. Do you think this kind of coursework (peer-review) should continue in CMPT 408? *

    ⬡ Yes

☐ No

25. Do you think this kind of coursework (peer-review) should be used in other classes? *

☐ Yes

☐ No

26. If "Yes", please suggest other classes

```



```

27. If "No", please give reason(s)

```



```

28. What do you think of EasyChair as a tool to support the peer review process?

```



```

29. Please, provide any additional comments or suggestions, if you have some, below:

```



```