Incorporating a code review process into the assessment

Dr Xiaosong Li

School of Computing and IT
Unitec New Zealand

`xli@unitec.ac.nz`

Abstract

This paper describes an action research process, which is used to implement the Code Review Process (CRP). The CRP was used as one of the assessment methods in a third year undergraduate Web Application Development course. There are two cycles in this study, 2005 Semester 2 and 2006 Semester 2.

Trend study surveys were used to collect the data. This data together with the author's observation and the students' comments were analysed. The issues of the CRP were identified. A set of teaching strategies were proposed to enhance the CRP.

The initial objectives of the CRP were achieved. The improvements introduced in the second cycle were successful in helping to achieve the initial objectives. For further improvement, the students need to gain more peer review related skills.

Keywords: Computing education, code review, coding standards, assessment, peer review.

1   Introduction

This paper describes a Code Review Process (CRP) employed by the author as one of the assessment methods in a third year undergraduate Web Application Development (WAD) course. The purpose of WAD is to provide students with knowledge and skills for developing client-server and web-based applications in the Intranet/Internet environment. The CRP was conducted for the first assignment of WAD. The assignment required the students to design and implement a token web application by using ASP.NET/VB.NE. A large amount of code was involved.

The initiative of the CRP was to motivate students to apply coding standards and to facilitate the communication among the students. It was proposed by Li and Prasad (2005) as a teaching strategy of coding standards. The rational of the CRP can be found in a number of relevant reports. It is reported that it is common practice to implement coding standards by using

CRP in industry, in which a team of developers sit down, check and make suggestions on the code produced by one or more members in the team according to formal style standards, the standards are enforced during code reviews (Parasoft 2005, 2007, and Pfeiffer 2005). Code reviews have the following purposes or benefits (Bogue 2006):

- To make sure that the code that is being produced has sufficient quality to be released.

- To find errors of all types, including those caused by poor structure, those that don't match business process, and also those simple omissions.

- To help developers learn when and how to apply techniques to improve code quality, consistency, and maintainability.

- To exchange ideas about how the source code is written and to establish a standard group.

- To help developers identify problems and envision new solutions.

- To make the code more concise, readable, efficient, flexible, and effective.

Using the CRP in the classroom should equip the students with the code review skills needed in their future workplace. The CRP as a type of peer review, which has been used as a collaborative learning technique in education (Carlson1, Berry and Voltmer 2005 and McGourty, Dominick and Reilly 1998), also provides an opportunity for the students to learn peer review skills.

The CRP has been used in the WAD course as one of the routine assessment instruments since 2005 Semester 2. In 2006 Semester 2, three improvements were introduced. This paper compares and analyses the results of 2005 Semester 2 and 2006 Semester 2 and proposes a set of strategies for future improvements. It is aiming to answer the following questions:

- Were the initial objectives set in 2005 Semester 2 achieved?

- Were the improvements introduced in 2006 Semester 2 successful?

- Was the CRP helpful in gaining peer review related skills?

In the rest of this paper, the overall research methodology is described, which is then followed by the method used for the first cycle and the second cycle . After that, the results are reported and analysed, a set of teaching strategies are proposed, and finally a summary is given.

2   The research methodology

The study of the CRP is an action research process. In action research, initially, a problem is identified and data is collected for a more detailed analysis. This is followed by a collective postulation of several possible solutions, from which a single plan of action is selected for implemention. Data on the results of the intervention are collected and analysed, and the findings are interpreted in light of how successful the action has been. At this point, the problem is re-assessed and the process begins another cycle. This process continues until the problem is resolved (O'Brien, 2001).

The problem was identified and analyzed by Li and Prasad (2005), which was that the students were reluctant to apply the coding standards. As the result of the analysis, the idea of the CRP was proposed. Thus the study of the CRP process was initialized. The CRP was designed and implemented in 2005 Semester 2. At the beginning of the 2006, the initial cycle was analyzed and reported (Li, 2006). With the input of the initial cycle, the process was carried on as the second cycle. This paper analyses, compares and reports the results of the second cycle and proposes   strategies for improvement in the next cycle.

The trend study survey was used as an instrument in this study. A trend study samples different groups of people at different points in time from the same population. The trend studies are uniquely appropriate for assessing change over time (O'Connor 2007 and Web Centre for Social Research Method 2007). In Li (2006), a set of questionnaires was used as one of the approaches to obtain students opinions. To check the results of the improvements introduced in 2006 Semester 2, at the beginning of 2007, the same set of the questionnaires with one extra question were sent to the students of 2006 Semester 2. In the both of the questionnaires, seven identical rating scale questions were used. However, this quantitative data was not for the purpose of carrying out rigorous statistical analysis, but to provide sufficient qualitative data to investigate the best suitable strategies for future teaching practice.

The data from the two surveys were interpreted. An "interpretive" approach would be phenomenological in nature or based on social interactionism. Researchers using this approach would seek to present a holistic view of data rather than a condensed view (Savenye and Robinson 2003). The author was expecting to get suggestions or clues from the students' opinions on how the CRP was done, whether the improvements introduced in 2006 Semester 2 worked. For further improvements, which direction should we go? The findings from the data analysis were combined with the author's own observation, literature review and informal conversations with the students to work out a set of strategies for next, 3rd,cycle.

3   The CRP in 2005 Semester 2

The CRP was introduced in 2005 Semester 2. The main objectives of the CRP were (Li 2006):

- Motivate students learning on Coding Standards
- Facilitate the communication among the students

So the CRP was not intended to be either a major assessment instrument or a collaboration learning tool. Therefore, the students were not trained on the assessment skills and peer review skills. The students were expected to reinforce their coding standards knowledge and learn peer review skills from the process.

The following are the main features of the CRP (Li 2006):

- Face to face review
- The students selected their own reviewers
- Contributed small weight to the course assessment (10% of Assignment 1 which in turn weighted 40% of the whole course)
- The reviews were used as a facilitating tool for coding standards
- The reviews were a part of the quality assurance of the Assignment 1

In 2005 Semester 2, 26 students enrolled in the WAD course. These students were from seven different nationalities, including New Zealand, China, Japan, Cambodia, India, Sri Lanka and Norway. The students' age distribution is shown in Table 1.

Table 1: Students' age distribution, 05s2 = 2005 Semester 2, 06s2 = 2006 Semester 2

	Group 1 21-25	Group 2 26-30	Group 3 31-36	Group 4 37-50
05s2	13	7	5	1
06s2	10	4	3	1

Among the 26 05s2 students, 22 (85%) were male and 4 (15%) were female. In summary, the students had the following features:

- Young majority
- Male majority
- Multiple cultural backgrounds

The CRP was carried out in the following sequences:

1. The students were given code guidelines which specified the coding standards required in the course.

2. The students were given a document which described the steps in the CRP for the students to follow.

3. The students were given a marking schema to be used during the CRP.

4. Each student then chose two peers to conduct the CRP.

5. Each student submitted the reports from the two peers with his or her Assignment 1.

6. The CRP marks for each student was the average of the marks given by his/her two peers.

At the beginning of 2006, an electronic survey was sent out by email. Out of 23 students who participated in the CRP, 12 responded. The response rate was 52%. The details of this information will be described in Section 5.

4   The CRP in 2006 Semester 2

The analysis of the results of 2005 Semester 2 suggested that some of the students felt that the CRP was unfair, because their peers were not well equipped with the necessary skills to assess their work. It also suggested that the reviewers should be assigned randomly to avoid the negative impact of "mutual admiration societies" (Li 2006).

The following improvements were introduced in the CRP for 2006 Semester 2:

- The two reviewers were assigned randomly by the lecturer.

- The lecturer reviewed each student's work and gave a mark as well.

- The CRP marks for each student was the average of the marks given by his/her two peers and which was averaged again with the lecturer's mark.

In 2006 Semester 2, 18 students enrolled in the WAD course. These students were from six different nationalities, including China, Switzerland, Sri Lanka, India, Malaysia and New Zealand. The students' age distribution is also shown in Table 1, which is pretty much the same as 2005 Semester 2. Among the 18 students, 14 (78%) were male and 4 (22%) were female. From this information we can see that the class in 2006 Semester 2 retained the similar features as the class in 2005 Semester 2.

At the beginning of 2007, an electronic survey was sent out by email. Out of 12 students who participated in the CRP, 7 responded. The response rate was 58%. The details of this information will be described in Section 5.

5   The data and analysis

5.1   The questionnaires

In the survey of 2005 Semester 2, there were seven rating scale questions and one selection question. In the survey of 2006 Semester, ninth question was added to check if the students liked the randomly assigned reviewers. The possible answers for each rating scale question are shown in table 2.

Questions and their numbers are shown in Table 3. These questions were discussed by Li (2006).

5.2   The results from the surveys

A summary of the survey results from the two different semesters, 2005 Semester 2 and 2006 Semester 2, is shown in Table 4. For Question 1- 7, averages of 2006

Table 2: Possible answers for the rating scale questions

Answers	Values
Strongly Disagree	0
Disagree	1
Neutral	2
Agree	3
Strongly Agree	4

Table 3: Survey questions, where CRP=Code Review Process

No.	Question
1	The CRP motivates you think more than Coding Standards
2	The CRP is helpful for your learning on Coding Standards.
3	The CRP motivates you to follow Coding Standards.
4	The CRP is helpful for you to improve your communication with your peers.
5	The CRP is helpful for you to make more friends.
6	The CRP is helpful for you to know more people with different culture background.
7	The CRP is helpful for you to learn and practice your critique skills.
8	How many of your code reviewers who you were not familiar with before this course?
9	Do you prefer the reviewers assigned by the lecturer or the reviewers chosen by yourself? (Assigned = 0; Neutral = 1 Chosen = 2, this question is only for 2006 Semester 2)

Semester 2 are higher in general than 2005 Semester 2. The standard deviations of 2006 Semester 2 are lower in general than 2005 Semester 2. These suggest that the students of 2006 Semester 2 like the CRP more than the students of 2005 Semester 2. This is reasonable as 2005 Semester 2 was the first time the CRP was introduced in WAD course. This suggests that the improvements listed in Section 4 were helpful, and thus the improvements could be considered successful. Question 8 was used to check that if the CRP had facilitated the communication among the students. The average of 2006 Semester 2 is much higher that the average of 2005 Semester 2. This suggests that either the students of 2006 Semester 2 had got more chance to communicate with their peers or we have more transferred students (from other institutions) in this class. From the survey data, it's not possible to determine which explanation is closer to the truth. The Question 9 has got high standard deviation as well. This is due to that about half students wanted assigned reviewers and another half wanted self-chosen reviewers.

Table 4: Summary of the survey results, 05s2 = 2005 Semester 2, 06s2 = 2006 Semester 2

No.	Average (05s2)	STD (05s2)	Average (06s2)	STD (06s2)
1	2.83	0.83	3.00	1.00
2	2.67	1.15	2.86	1.07
3	2.75	1.06	3.14	1.07
4	2.67	1.37	3.43	0.53
5	2.50	1.17	2.57	0.53
6	2.42	1.08	3.14	0.69
7	2.67	1.23	3.00	1.15
8	0.83	0.72	1.43	0.79
9			1.14	1.07

The standard deviations of Questions 1, 2, 3 and 7 are still pretty high despite the higher averages. This suggests that there is a minority of the students who still really don't like the CRP. To investigate the opinions of different student groups, the samples were divided into four groups: male students, female students, local students and overseas students. As shown in Table 1, the majority of the students fell between the ages 20 – 35. There was no sign showing the diversity of their opinions in ages. The answers from each of the groups were examined and compared. To list the average values for the answers from each group, a code is given to each group. Table 5 shows the code name of each sample group.

Table 5: The code name of each sample group

Name	Meaning
MAV	the average of male students
FAV	the average of female students
LAV	the average of local students
OAV	the average of oversea students

Table 6: Average of 2005 Semester 2 in groups

No.	Average	MAV	FAV	LAV	OAV
1	2.83	2.9	2.5	2.33	3
2	2.67	2.8	2	2.33	2.78
3	2.75	2.8	2.5	2.67	2.78
4	2.67	2.8	2	2.33	2.78
5	2.5	2.7	1.5	2	2.67
6	2.42	2.7	1	2	2.56
7	2.67	2.8	2	2.33	2.78
8	0.83	0.7	1.5	1.33	0.67
9	2.83	2.9	2.5	2.33	3

Table 6 shows the average values of the answers from 2005 Semester 2. Table 7 shows the average values of the answers from 2006 Semester 2.

Table 7: Average of 2006 Semester 2 in groups

No.	Average	MAV	FAV	LAV	OAV
1	3	3.5	2.33	2	3.4
2	2.86	3	2.67	2	3.2
3	3.14	3.5	2.67	2	3.6
4	3.43	3.75	3	3	3.6
5	2.57	2.75	2.33	2	2.8
6	3.14	3	3.33	3	3.2
7	3	3.5	2.33	2	3.4
8	1.43	1.5	1.33	2	1.2
9	1.14	2	0	0	1.6

Table 6 demonstrated that, for 2005 Semester 2, the average values from the male students are in general higher than the female students. This suggests that in general the male students are happier with the CRP than the female students. The reasons for this could be many. Due to the small size of the samples, it's hard to identify the particular reasons. Table 6 also demonstrated that the average values from the overseas students are in general higher than the local students. This suggests that in general the overseas students are happier with the CRP than the female students. This is consistent with the author's observation reported in (Li 2006). One possible interpretation is that the overseas students missed their families in their home country, so they needed more social life. On the other hand, the local students had enough social life in their homes and the local communities.

Table 7 demonstrated that, for 2006 Semester 2, the average values from the male students are in general higher than the female students as well. Table 7 also demonstrated that the average values from the overseas students are in general higher than the local students. These results are consists with their correspondences in Table 6.

Another point worth to notice is that the averages for Question 2 and 5 are consistently low for both 2005 Semester 2 and 2006 Semester 2, this is shown in Table 4. A close look at these two questions shows that Question 2 is about whether the CRP is helpful for coding standards learning; Question 5 is about whether the CRP is helpful to make friends. On the other hand, Question 3 has got a reasonable high average in 2006 Semester 2, which is about whether the CRP is helpful to motivate the students to follow the coding standards. Question 6 has also got a reasonable high average in 2006 Semester 2, which is about whether the CRP is helpful to know more people. This suggests that the CRP is helpful to know more people; however, they are not close enough to become friends.

In comparing the averages of Question 7, it was 2.67 for 2005 Semester 2 and 3 for 2006 Semester 2. The improvement is not as big as other questions. This question is about critique skills. The data suggests that the CRP does help in gaining critique skills, however, it's not a great help. In particularly, the improvements introduced in 2006 Semester 2 did not make too much difference in this area. This is reasonable as the CRP was not designed as a collaboration learning tool. As the critiques skills should an important learning outcomes of the peer review (Kern, Saraiva and dos Santos Pacheco 2003), there is a big room for future improvement. The lack of the critique skills would have a negative impact on the quality of the CRP.

In summary of the survey data, the CRP did well in achieving the two initial objectives indicated in Section 4. The improvements introduced in 2006 Semester 2 did improve the CRP in terms of achieving the initial objectives. However, this is not enough. In particularly, the CRP should be more helpful in learning coding standards and peer review related skills.

5.3 The data from other sources

In this section, we report and discuss the data from the author's observation and the student's comments.

The students' in general had given higher marks than they should be. In 2005 Semester 2, it was observed that the grade of CRP for a student was in general higher than his/her final grade. Most of the students had got A grade in their CRP, but only 4 students had got A grade in their final grade for that Semester. In 2006 Semester 2, the marks given by the students were 6.41 higher than the marks given by lecturer on average. One of the reasons for this could be due to the students' lack of peer review skills and coding standards knowledge. Another reason of this could be due to the students' superficial engagement; they just quickly gave their peers a satisfied mark and returned back to their own busy work. These are consistent with some of the student's comments:

- *It was very difficult for students to arrange a time outside of class as all students are very busy with other papers etc.*

- *Most students were so worried about trying to finish their own code that they weren't interested in reviewing anyone else's code and saw it as an interruption and inconvenience.*

- *How could a student give another student an accurate grade? They would give them all As.*

The participation rate for the CRP is dropping down. In 2005 Semester 2, among 25 students who attempted the Assignment 1, 23 had participated the CRP, the participation rate was 92%. On the other hand, in 2006 Semester 2, among 16 students who attempted the Assignment 1, 12 had participated the CRP, the participation rate was 75%. The following are possible reasons for this:

- The students had got more work load, so they didn't have time for the CRP.

- The lecturer had made the contribution to the CRP in 2006 Semester 2, so a student wouldn't get zero mark if he/she didn't participate in the CRP.

- The students were not motivated enough to participate the CRP. The students didn't realise the importance of the CRP.

- The students needed to be equipped with more skills to actively participate in the CRP.

The above issues should be addressed by introducing more effective teaching strategies.

6 A set of proposed teaching strategies

To address the issues discussed in Section 5, more effective teaching strategies are required. The main objective of the strategies should be to help the students gain peer review related skills. This should include the following techniques (Dimon 2006):

- Ask questions rather than make statements. Be solicited rather than imposed. Involve sharing information rather than giving advice.

- Avoid the "Why" questions.

- Remember to praise.

- Make sure the discussion stays focused on the code and not the coder. Focus on observed behaviour rather than on the person.

- Remember that there is often more than one way to approach a solution.

- Put a summary comment and be positive.

- Use an electronic mechanism to record the comments.

- Make an upfront agreement that not every question needs to be responded to

- Promote reflection as part of a dialog between the giver and receiver of feedback.

- Be descriptive rather than judgmental.

- Be specific rather than general.

- Consider the needs of both the receiver and giver of feedback.

- Consider the amount of information the receiver can use rather than the amount the observer would like to give.

- Build a supportive, confidential relationship based on trust, honesty, and genuine concern.

An important method to achieve the above is to train the students, give them more guidelines. In those teaching practices where peer review plays an important role in assessment, training is necessary. The examples of these are Carlson1, Berry and Voltmer (2005) and McGourty, Dominick and Reilly (1998). In (Carlson1, Berry and Voltmer 2005), the students were given two practices before the peer review. One was reading three benchmark

samples. Another was self-assessment. These approaches can be adopted for our CRP. We could develop four benchmark CRP samples for poor, good, fine and excellent work respectively. We then let the students do a self code review accordingly before their peer code review. The students will get marks just for participation. Their work will be commented for future improvements.

It should be very helpful if we provide a training document for the CRP. The document should cover the rationale of the CRP, the learning outcomes of the CRP, the process of the CRP, the correct attitude for the CRP and the techniques of the CRP.

To solve the time problem so the students will have enough time to focus on the code review, we will allow the students to submit their code review one week later than their other documents.

To make the CRP more efficient and confidential, we could make use existing online tools for the CRP. For example, Turnitin (Turnitin 2007) can be used for anonymous document CRP. Blackboard chat room (Blackboard 2007) can be used for anonymous interactive CRP.

7 Summary and future plans

This study showed that the current CRP had done well in achieving the following initial objectives:

- Motivate students learning on Coding Standards

- Facilitate the communication among the students

The improvements introduced in 2006 Semester 2 were successful in terms of achieving the initial objectives. However, this is not enough. The CRP should be more helpful in learning coding standards and peer review related skills.

To help the students to learn peer review related skills. We need to introduce a training process. This includes

- Four benchmark CRP samples

- Self code review

- A training document

We also can make use online tools to get anonymous code review.

The proposed strategies will increase the weight of the CRP in the course assessment, which will increase the workload for both teacher and the students. To address this issue, we need to re-arrange assessment framework of the course, for example, introducing group work.

The data samples used in this research was too small to get a complete understanding of the students' opinions. This might a negative impact on the proposed teaching strategies.

8 References

Blackboard. (2007): The Blackboard Content System. http://www.blackboard.com/products/Academic_Suite/Content_System/index. Accessed 11 March 2007.

Bogue, R. (2006). Effective Code Reviews Without the Pain.http://www.developer.com/java/other/print.php/3579756 Accessed 11 March 2007.

Carlson1, P. Berry, F. and Voltmer, D. (2005): Incorporating Student Peer-Review into an Introduction to Engineering Design Course. *Proceedings of 35th ASEE/IEEE Frontiers in Education Conference*, pF2C20-pF2C25.

Dimon,G. (2006): Code Reviews: Write Better Code Overnight. http://www.digital-web.com/articles/code_reviews_write_better_code_overnight/. Accessed 11 March 2007.

Kern V.M, Saraiva L.M. and dos Santos Pacheco R.C (2003): Peer Review in Education: Promoting Collaboration, Written Expression, Critical Thinking, and Professional Responsibility. Education and Information Technologies, 37-46.

Li, X. and Parasad, C., (2005): Effectively Teaching Coding Standards in Programming. *Proceedings of SIGITE*, pp239-244.

Li, X. (2006): Using Peer Review to Assess Coding Standards – A Case Study. *Proceedings of 36th ASEE/IEEE Frontiers in Education Conference*, pT2D9-pT2D14.

McGourty, J., Dominick, P., and Reilly, R., (1998): Incorporating Student Peer Review and Feedback into the Assessment Process. *Proceedings, Frontiers in Education, Session T1A*.

O'Brien, R. (2001): An Overview of the Methodological Approach of Action Research. http://www.web.ca/~robrien/papers/arfinal.html. Accessed 4 May 2007.

O'Connor, T. (2007): Lecture note, Survey Research Design. http://faculty.ncwc.edu/toconnor/308/308lect07.htm. Accessed 11 March 2007.

Parasoft. (2005): Understanding the Workflow in a Coding Standards Implementation, http://www.parasoft.com/jsp/pr/runtimes.jsp?runtimeId=1169 Accessed 4 March 2005.

Parasoft. (2007): Source Code Review Process, http://www.parasoft.com/jsp/aep/aep_practices.jsp?practice=CodeReview Accessed 11 March 2007.

Pfeiffer, J. (2005): Coding Standards in 400- and 500-level Classes. http://www.cs.nmsu.edu/~pfeiffer/classes/general/s05/coding.html. Accessed 1 July 2005.

Savenye, WC. and Robinson, RS. (2003): Qualitative research issues and methods: An introduction for educational technologists. In D. H. Jonassen (Ed.), Handbook of research on educational communications and technology (2nd ed., pp. 1045-1071). http://www.aect.org/edtech/39.pdf. Accessed 11 March 2007.

Turnitin. (2007): Peer Review Turnitin's Collaborative Learning System.

https://turnitin.com/static/pdf/datasheet_preview.pdf
Accessed 11 March 2007.

Web Center for Social Research Method. (2007): Trend studies.
http://www.socialresearchmethods.net/tutorial/Cho2/trend.html . Accessed 11 March 2007.

Weinreich, N. (1996): Integrating Quantitative and Qualitative Methods in Social Marketing Research.
http://www.social-marketing.com/research.html.
Accessed 11 March 2007.

Web Center for Social Research Method. (2007): Trend studies.
http://www.socialresearchmethods.net/tutorial/Cho2/trend.html . Accessed 11 March 2007.

9   Acknowledgement