Aropä and PeerWise: Supporting Student Contributed Pedagogy in Large Classes

John Hamer, Paul Denny, Andrew Luxton-Reilly, John Hosking, Beryl Plimmer

Abstract

Aropä and PeerWise are two web-based tools that support collaborative learning in large, undergraduate classes. Aropä manages peer assessment activities, allowing students to take part in double-blind refereeing of their peers' coursework. PeerWise is a data bank of multi-choice questions contributed, explained and discussed entirely by students.

These systems leverage the latent intellectual capacity of a large class to provide new opportunities for learning. Using Aropä, each student might review three or four essays and receive a corresponding amount of feedback about their own essays, all within a few days. The immediacy and diversity of the feedback is substantially greater than can be produced by a tutor. While the quality of the reviewing is typically variable, there are affective benefits in challenging students to distinguish between good and poor feedback. By eliminating the stamp of authority and introducing diverse, possibly conflicting feedback, students are required to exercise their critical judgement in deciding what information to accept and reject. Moreover, tutor marking can still be used, and can even be mixed in with the peer reviewing.

PeerWise leverages the energy of a large class in a different way, building an annotated question bank that can contain thousands of multiple-choice questions. Each question is accompanied by an explanation written by the question author, overall quality and difficult ratings assigned by students who have answered the question, and possibly a forum in which misunderstandings and possible improvements are discussed. The question bank thus serves two complementary purposes: a creative medium in which students engage in deep learning and critical reflection; and a drill-and-test library for developing fluency with the course content.

We have statistical evidence to show that active use of these tools strongly correlates with performance in formal examinations. Further, as a side-effect of channeling all interaction through a central database, a detailed record of student interaction is collected. This record allows instructors to monitor overall class performance and to assess individual students over time in modes that limit opportunities for plagiarism. With routine use, a rich picture of student performance is collected.

We are currently at the point of building additional tools to further exploit the interaction data. These include reputation systems, whereby the quality of an individual's comments and feedback is judged by the recipients, and recommender systems, in which participants are able to highlight instances of high quality work. Both of these ideas are present in popular online auction and shopping sites, but have not been widely adapted for educational use.

This paper describes the Aropä and PeerWise tools, discusses the education theory behind the ideas, presents results from ongoing research study into student learning and attitudes toward the tools, and elaborates some of our ideas for future development.

1. Introduction and Related Work

The "knowledge economy" demands a radical shift in pedagogy to prepare students appropriately. Teachers in higher education must help students to develop the skills necessary to work independently, filter large amounts of information, critically evaluate the quality of information, act as part of a community, and use online tools to communicate effectively. Development of these skills can be facilitated by the use of contribution-based pedagogies (CBPs) which involve students in creating and sharing learning resources [21]. Significant benefits of CBPs include the development of communication, teamwork and peer- and self-assessment skills, which are integral to effective operation in the “knowledge economy” and help build a foundation to support lifelong learning. Higher-order cognitive processes such as evaluation, reflection and critical thinking are emphasized, and students are transformed from passive receptors of information to active and critical members of a community engaged in the process of constructing knowledge.

Peer assessment has been used in many institutions for more than 50 years [26], in a wide range of higher education contexts such as academic writing, science, engineering, business and medicine [16, 19]. The literature on both self-assessment [6, 17] and peer assessment [28] reports numerous benefits for students. Boud [5] suggests that self-assessment is critically important, and courses that do not encourage self-assessment can actually undermine lifelong learning [4]. Stefani [27] states that if we want our students to be autonomous, reflective and independent, then our assessment practices should include these qualities. The use of peer-assessment in a course may also improve student’s self-assessment skills [29]. Kern et al. [23] recently noted that while the use of peer review in the classroom has numerous benefits, there has been only one report of large scale education application of peer review, conducted in 2001, and involving 411 students. They report that peer review should be used in higher education and should become regular educational practice.

A number of approaches and systems that facilitate production and sharing of student generated assessment materials have been described in the literature. Horgen [22] used a lecture management system to share student generated MCQs. Fellenz [18] reported on a course where students generated MCQs which were reviewed by their peers, although technology was not used to support this process. Fellenz reported that the activity increased student ownership of the material and motivated students to participate. Barak [2] reports on a system named QSIA used in a postgraduate MBA course in which students contribute questions to an on-line repository and rank the contributions of their peers. Arthur [1] reports on a large course activity in which students in one lecture stream prepare quiz questions for students in another stream. Yu [31] has students construct MCQ items and submit them to an on-line database where they are peer-assessed. Feedback about quality is used to improve the items before they are transferred to a test bank database to be used for drill-and-practice exercises. All of these reports agree that student-contributed MCQs is a powerful idea.

A community of practice is described by Wenger, McDermott and Snyder [30] as "a group of people who share a concern, a set of problems, or a passion about a topic, and who deepen their knowledge and expertise in this area by interacting on an ongoing basis". The focus that CBPs place on student-driven learning helps to form a community of practice amongst the student population, in which students learn from and with each other. The development of learning communities and the interactions among classmates that they produce are recommended by Cohoon [9]. Brookfield [7] notes that students learn from their peers through advice, information and skill modeling and suggests that peer learning is, in fact, crucial for success. He states: "The learning activities of successful self-directed learners are placed within a social context, and other people are cited as the most important learning resource". CBPs may even have additional significant benefits for gender or ethnic minorities. Barker et al. [3] suggest that retention of female students in Computer Science may be improved by creating a classroom culture in which learning is a social or community practice rather than a solitary one.

CBPs can be successfully employed in small classes easily as the overhead of handling the distribution and management of student contributed resources is light. However, this overhead grows quickly and for large classes the use of technology is required to provide support for these approaches.

In this paper, we describe two scalable web-based tools, Aropä and PeerWise, that facilitate meaningful student contribution and task based collaborative learning. Using these tools, students can share and comment on each other's documents, practice peer assessment and develop valuable learning resources with minimal effort required from academic staff. In addition, the rich data that these tools collect on student performance make them useful in classes of all sizes. These tools are regularly used in over a dozen courses and by at least 1000 students each semester, including students from multiple institutions.

2. Aropä

Aropä electronically facilitates the administration of peer assessment activities. Students upload documents to be reviewed and download documents for reviewing online. Reviews are entered online via an instructor customisable web form. Aropä manages the allocation of reviews, calculates weighted average grades, and an assortment of other administrative tasks. The Aropä tool and an evaluation of it are described in detail in [20]. For the evaluation, qualitative data collected from two large courses was analysed. The evidence suggested that Aropä successfully supported peer review activities in large classes and contributed to student learning on many different levels.

Our motivation for developing Aropä was to make peer assessment a routine activity even in large, undergraduate classes. Since it was developed in 2004, over 1,000 students have used Aropä each semester in classes in disciplines including Academic Practice, Business, Civil Engineering, Commercial Law, Computer Science, English, Electrical Engineering, Environmental Science, Information Management, Medical Science, Pharmacology, and Software Engineering. The class sizes range from 12 to 850.

For a student, using Aropä is straightforward. It involves firstly document upload, e.g. a report, computer program, or essay, via a standard web file upload page, then reading and reviewing allocated submissions and finally reading the feedback provided by anonymous reviewers. A screenshot of the main interface is shown in Figure 1.

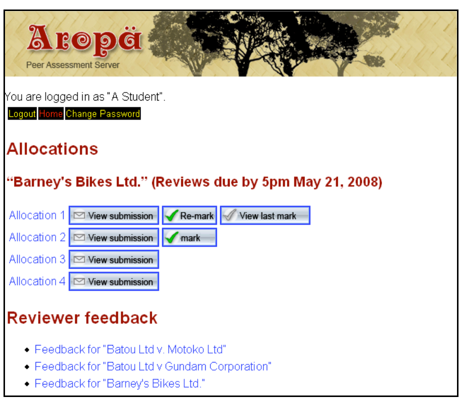

Figure 1: The main page for a student using Aropä. All three student activities (uploading, as author; reviewing; and receiving feedback) are available from this page (uploading is not shown).

Instructors must decide on the date submissions are due, the length of the review period, the number of reviews to be allocated to each student, and the grading rubric to be used. We have found the following guidelines to instructors to be useful:

- Allocate three to five reviews to each student. This seems to give a good balance between the need to provide a variety of feedback and the time required to write the reviews.
- Allow two days to at most a week for reviewing. Long review periods have largely proven unsuccessful.
- Start reviewing immediately after submissions are complete. One of the strengths of peer assessment is in providing rapid feedback; participation rates fall away if the reviewing does not start immediately.

We have found that instructors tend to use Aropä in three distinctive ways, differing in their timing, degree of compulsion and grading rubric:

- formative feedback on a draft;
- critical reflection after an assignment; and
- summative assessment.

Figure 2 is a typical example of a summative assessment activity grading rubric. The criteria are quite precise and detailed as the instructor is mainly interested in giving feedback to the author, rather than having the reviewer reflect closely on the marking process. Marking consistency is a major consideration for summative assessment. Aropä provides support for identifying poor reviewers and to give weight to reviewers who have done a good job. Marks are often awarded for both the authoring and the reviewing. This motivates students to take the reviewing seriously.

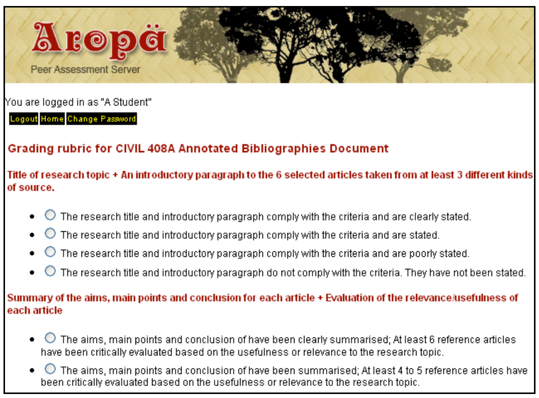

Figure 2: A grading rubric, showing Likert-style selections with detailed commentary provided by the instructor. Rubrics can also include open-ended comment fields and yes/no check boxes, as well as images, tables, multiple levels of headings, and character and paragraph styles.

Figure 3 shows a grading rubric for a more formative feedback style on a draft essay. Here the rubric has no quantitative elements, instead guiding the reviewer to provide more reflective comments via a series of open-ended comment boxes. The author has an opportunity to incorporate feedback provided into a final version of the essay. Significantly, students report that reviewing other essays helps them to identify faults in their own writing providing double value.

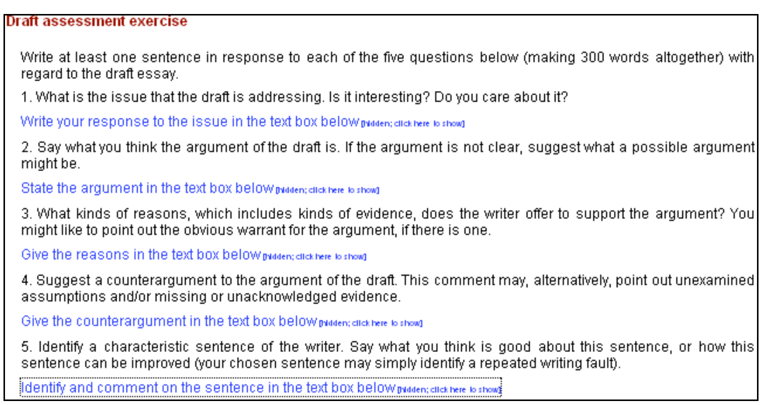

Figure 3: A formative grading rubric. The rubric is comprised of open-ended responses only, and explicitly requests the reviewer state their opinion, with no expectation of there being a single “correct answer”

Combination rubrics, including both formative and summative elements, are often used to add a peer assessment step to courses with regular short weekly or fortnightly assignments. Students report significant value in reading work from their peers, permitting them to judge their own performance relative to the class. Over-confident students are given a reality check, and students lacking in confidence (often women in IT disciplines) are reassured that they are doing better than they perceived..

3. PeerWise

PeerWise allows students to compose multiple choice questions with associated explanations, and share them so that they can be answered and evaluated by other students in the class. Any student can answer any of the available questions in a standard drill and practice fashion, and can critique and rate each question answered for quality and difficulty. These ratings are also shared, and can be used to assist students in deciding which questions to answer.

The design of the PeerWise tool has previously been reported [10]. Common usage patterns over a range of courses have been examined [11], and students of all ability levels appear willing to answer many more questions than they are required. An analysis of the quality of student contributions has been conducted [14] and it was found that not only were most questions clear, correct and with good associated explanations but in the few cases where questions contained errors, those errors were effectively detected and discussed by other students. Even when no guidance is provided to students on the topics for which questions should be created, we have found all major course topics are usually well represented in the repository of questions that the students develop [15]. The efficacy of the tool has been evaluated [12] by measuring the correlation between student engagement and exam performance, with the most actively engaged students performing significantly better in written exams than students of equivalent ability who are less active.

Typically, courses start with empty repositories, and students develop relevant questions as the courses progress. All of the content remains available to students for revision purposes prior to final examinations. All activity on PeerWise, such as contributing new questions, answering existing questions, and rating and providing feedback on questions is confidential.

When a student first logs in to PeerWise, they are shown a list of the courses to which they belong. After selecting the relevant course, the main menu is displayed to the student (Figure 4). The main menu is divided into three sections: the questions that the student has contributed; the questions which have been created by others and which the student has answered; and the questions which have been created by others but which the student has not yet answered. Each of these sections is described in detail next.

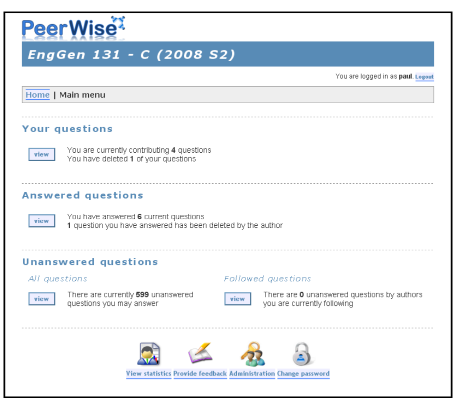

Figure 4: The main menu for PeerWise

3.1 Your questions

All of the questions that a student has contributed are displayed in this section (Figure 5). The items are displayed in a table with columns listing the date the question was created, how many times the question has been answered and the current rating of the question. The table can be sorted with respect to any of these columns. There is also a column that displays the difficulty of the question, as perceived by students who have answered it.

Students can elect to "follow" question authors who have contributed questions that they particularly like. Following an author provides a simple way of discovering good questions, as all of the other questions that author has contributed become easily accessible in a separate section. For question authors, attracting followers is an endorsement of the quality of the questions being contributed, and the number of followers a student has is displayed in the "Your questions" section.

The details of a question can be viewed by selecting it from the table. These details include the question text, as well as a histogram showing how often each alternative has been selected, and any feedback that has been provided by students who have answered it. Question authors are able to respond to any feedback written about their questions.

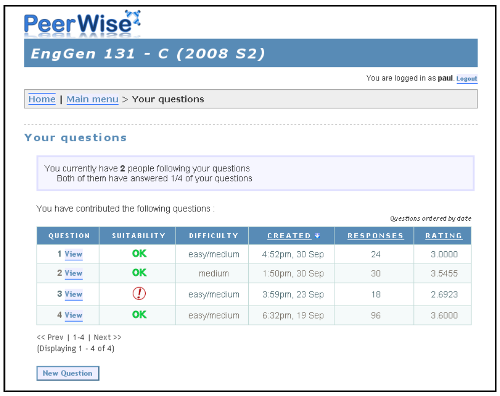

Figure 5: Page showing the questions written by the student

Contributing a new question involves providing a question stem and between two and five alternatives. The question author indicates which of the alternatives is correct, and provides an explanation for the answer. The explanation is displayed to all students upon answering the question, and is intended to assist students who have answered the question incorrectly to understand their mistake. Question contributors are able to "tag" their questions to indicate relevant topics. The tags are presented in a cloud that enables students to quickly locate questions on topics of interest. As soon as a new question is contributed, it immediately becomes available in the "Unanswered questions" section (Figure 6) for all other students in the course.

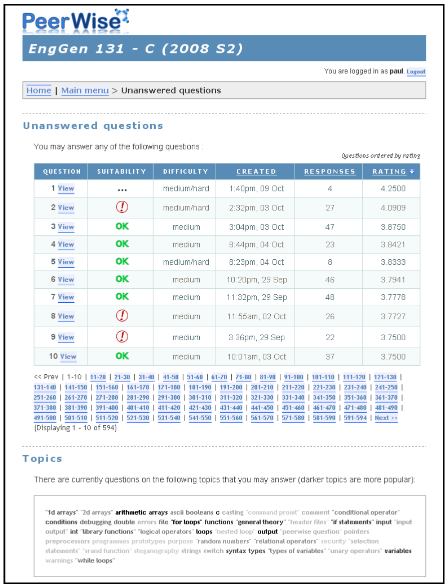

Figure 6: Page showing unanswered questions and topic cloud

3.2 Unanswered questions

Every student has access to all of the questions in the system. The unanswered questions are presented in a table from which the student can select individual questions to answer. The columns of this table include the perceived difficulty, number of responses, and current rating of each question. As the questions are also tagged by topic, students using PeerWise for drill-and-practice revision can spend their time answering highly rated questions on topics of interest to them, at a difficulty level they feel comfortable with.

Once a student selects an answer to a question, they are immediately shown the correct answer suggested by the author of the question, and the number of times each alternative was selected by other students in the course. The explanation provided by the question author is also displayed, as are all student comments written about the question. A simple metric is used to assess whether the selected answer is correct. If the answer selected by the student matches the answer suggested by the question author, and in turn this matches the most popular answer selected by other students, then the answer is deemed to be correct. In other scenarios, different icons are displayed depending on whether the student agreed with the author, or with the most popular answer selected by other students.

At this point, a rating form is displayed, which allows the student to rate the quality and difficulty of the question, and provide their own feedback. The quality rating is on a scale of 0 to 5, and the difficulty can be specified as either "easy", "medium" or "hard". As questions are answered, they move from the “Unanswered questions” section to the "Answered questions" section (Figure 7) where they remain available for review at any time.

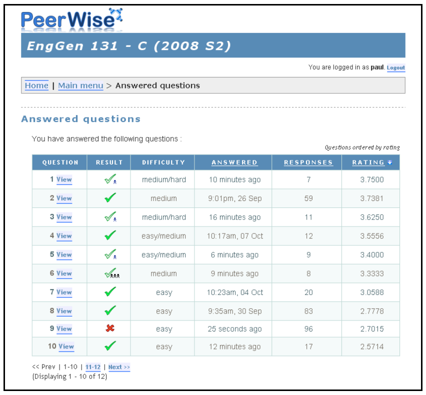

Figure 7: Page showing answered questions

3.3 Answered questions

The questions that a student has currently answered are displayed in a table in the "Answered questions" section. If a student has provided feedback on a question, they can check to see if the author has responded to their feedback in this section. In addition, as more students submit answers to these questions, the accuracy of the correctness metric improves.

3.4 Leaderboard

A simple leaderboard is available, which anonymously ranks certain kinds of contributions (Figure 8). This provides each student with an opportunity to compare their performance against that of the most active students. Specific tables display the top rated questions, and rank students on the number of questions that have been answered, on the popularity of question authors and on the popularity of students who have written feedback on questions.

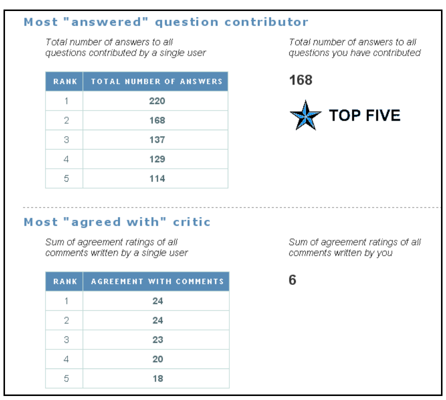

Figure 8: A section of the set of leaderboard tables

4. Discussion

4.1. Student attitudes

In general, students have responded positively to these tools. We have noticed that workload has a big impact on student attitudes. Students have responded negatively to their use in semesters where the required participation for PeerWise has been very high, or when students are allocated a large number of Aropa reviews.

In 2006, we conducted an anonymous survey of students' attitudes towards using Aropa for reviewing programming assignments in an introductory programming course [20]. We received 155 responses to the following two open-ended questions:

- What did you like most about the peer marking system?
- In what ways could the peer marking system be improved?

We coded the student responses to these questions, and Figure 10 summarises the five most common themes in the responses.

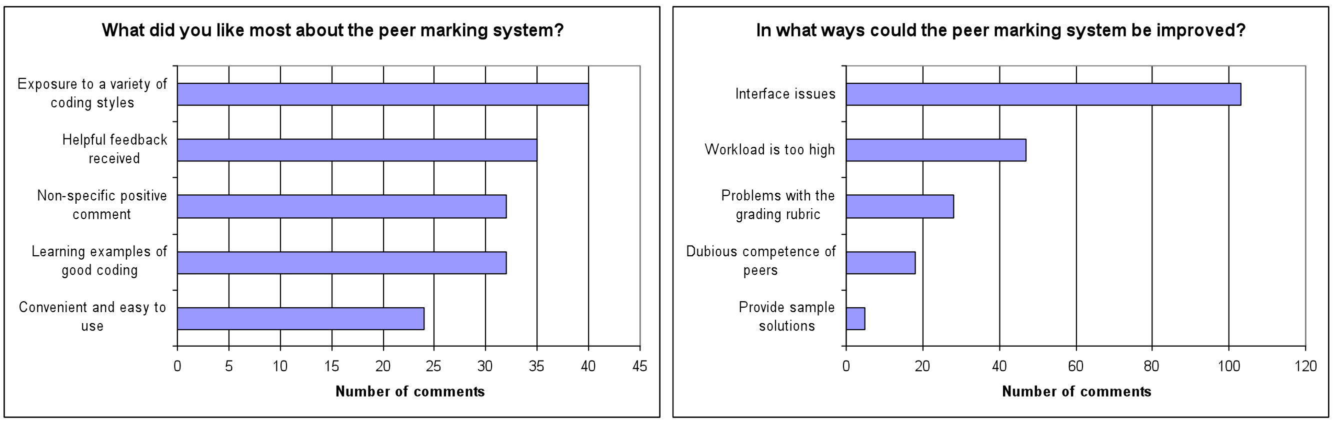

Figure 10: Most common responses to open-ended survey on student perceptions of Aropa

Nearly 70% of respondents found the Aropa interface difficult to use, and this has been improved since the survey was conducted. Around a quarter of the students claimed they liked being exposed to a variety of coding styles and felt the feedback they received on their own work was helpful.

In 2007 we conducted a similar survey regarding students' attitudes towards using PeerWise [13]. We had noticed throughout the semester that students tended to contribute more than the minimum requirement for assessment, and we included a question on the survey to investigate this further. The survey also included questions similar to those on the Aropa survey, specifically on the features that students found most useful / interesting / enjoyable, and on perceived problems with PeerWise:

- Q1: Which features of PeerWise did you find most useful/interesting/enjoyable?
- Q2: If you contributed more than the minimum requirement, why?
- Q3: What do you believe are the biggest problems with PeerWise?
- Q4: What do you believe are the biggest benefits of using PeerWise?

The PeerWise survey had 439 respondents, and the common responses are summarised in Figure 11.

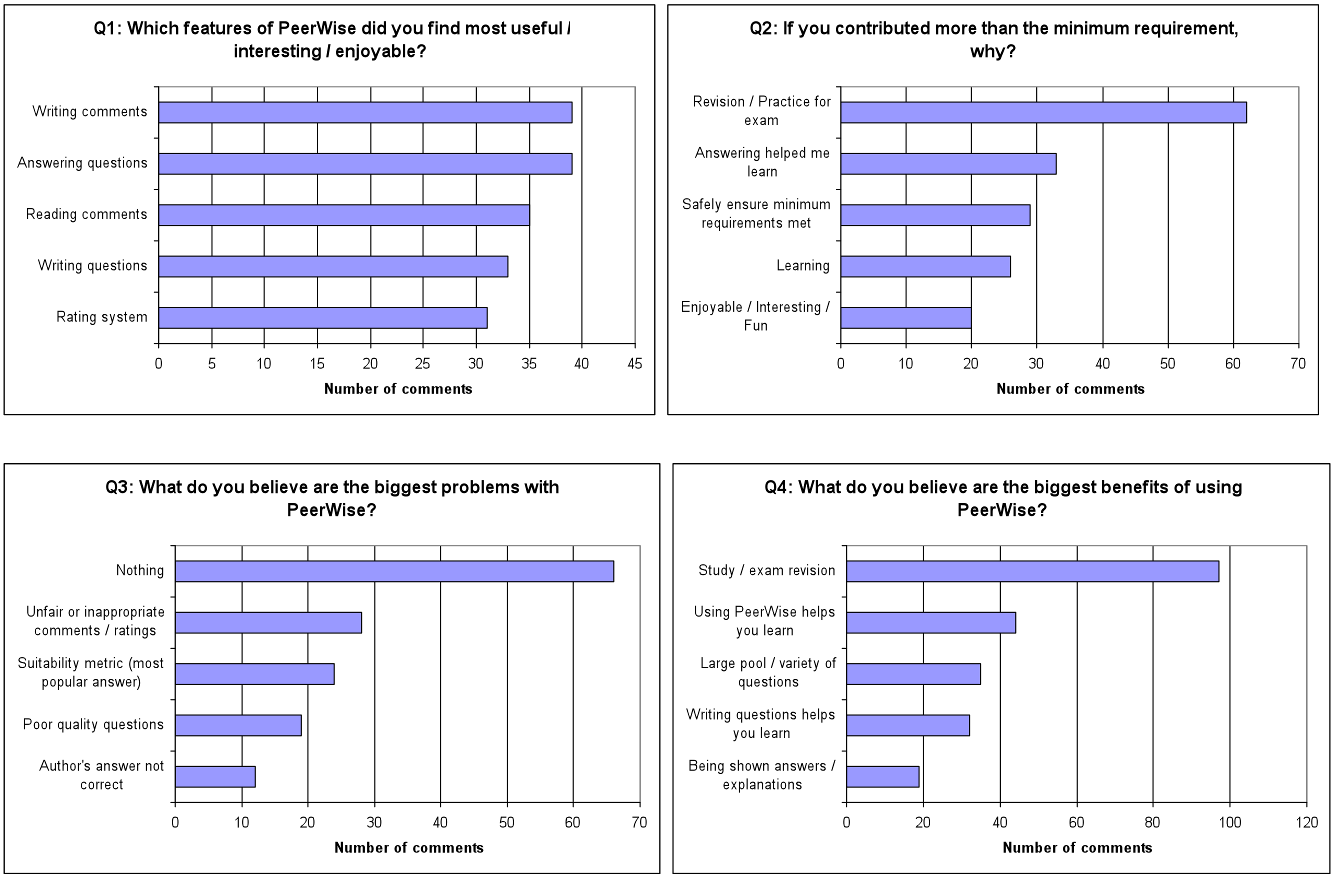

Figure 11: Most common responses to open-ended survey on student perceptions of PeerWise

Both the Aropa and PeerWise surveys also included a series of Likert scale questions. Table 1 displays the results of these questions.

Developing new questions helped me learn		I feel comfortable with other students reading my work	
Answering other student's questions helped me learn		Marking other students' work helps me spot mistakes in my own work	
The ability to read, write and endorse comments was useful		The comments from other students were helpful to me	
I would like to use PeerWise again next year		I would like more assignments to be peer marked	

Table 1: Likert scale survey questions and results on PeerWise (left) and Aropa (right). The graphics show the distribution of responses, Agree on the right and Disagree on the left, darker for Strongly

Student responses were considerably more positive to the PeerWise survey. The different nature of the tasks involved, and the fact that the workload for Aropa was perceived to be high in the semester the survey was conducted may explain much of these differences in attitude.

4.2. Student learning

As a first step in examining the benefits of our tools, we conducted a correlation study which indicated that student engagement with PeerWise is correlated with results in formal examinations [12]. We plan to carry out a similar study with Aropa.

In this study, students were split into quartiles based on their performance in an examination held prior to the introduction of PeerWise. Within each quartile, students were divided into the most active and least active groups, where activity was measured as a combination of the core activities – contributing questions, answering questions and providing feedback – as well as a measure of time on task. The final examination scores of these groups were then compared, as illustrated in Figure 9. In each quartile, the mean exam mark of the most active group was significantly higher ($p < 0.05$) than the mean exam mark of the least active group.

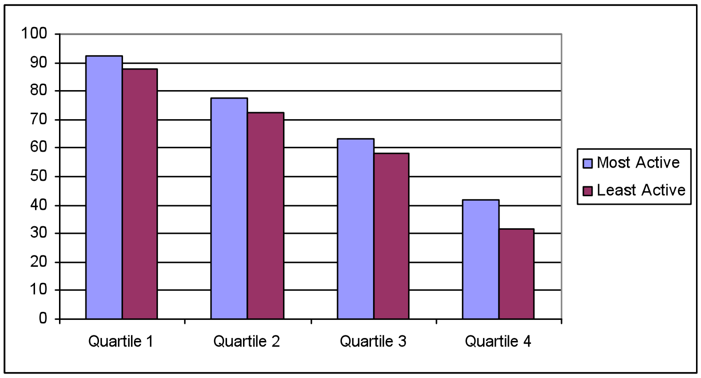

Figure 9: The mean examination marks of the most and least active groups within each quartile

4.3. Student usage

One advantage of using tools for supporting a CBP is that student activity can be logged and so a detailed record of student contributions is collected. These records allow instructors to monitor overall class performance and to assess individual student contributions to an extent that is not possible without the support of technology. A good example is measuring students' time on task, which is widely considered to be associated with learning gains [8, 23, 24].

In the context of Aropa, time on task might be a measure of how long students spend on their reviews. Measuring time spent reviewing accurately is difficult because many students download their items for review well before they start the reviewing process. Although we attempted to survey students on their time spent reviewing [20], the responses varied wildly from "not much" to 1.5 days. However it is possible to easily identify students who submit their completed reviews very soon, in a few cases nearly immediately, after downloading their allocations.

In the context of PeerWise, time on task might be a measure of how much time students spend answering questions. Students tend to view questions to answer only when they are ready to answer them, and so the time between viewing a question and submitting an answer is generally a fairly accurate measure of time on task. Outliers, which may occur if a student is interrupted during the process of answering a question and then returns to it later, can be ignored by calculating the median time on task. Figure 11 shows the median time taken by students in an introductory programming course to answer questions on PeerWise. There were 446 students in the analysis, which considers all students who answered a minimum of 10 questions.

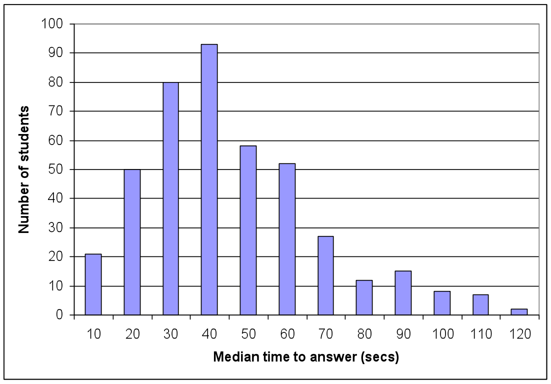

Figure 11: The median time taken to answer a question

From Figure 11, we can see that most students spend around 40 seconds answering each question, the median value being 37 seconds. Of concern to an instructor may be the group of students who spend very little time on task. Using the logged activity data, it is simple to identify such students and examine their individual activity in greater detail.

For example, Figure 12 targets an individual student who was identified as having a median time of 1 second to answer a set of 31 questions. The questions in this Figure are presented in the order in which the student answered them. Most questions were answered in 1 second, with the most time spent on any question being 6 seconds.

In contrast, the student in Figure 13 chose to answer fewer questions, a total of 20, but spent far longer on each one. Only 5 of the questions were answered in less than two minutes.

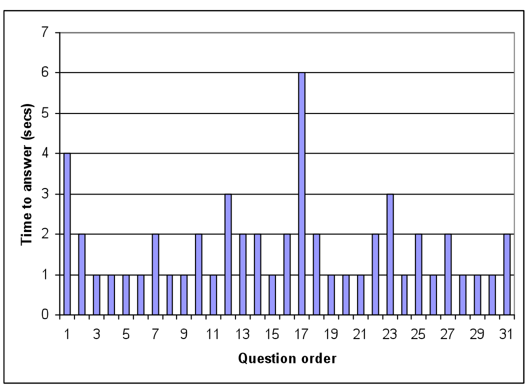

Figure 12: Time spent answering questions by an individual student (median time = 1 second)

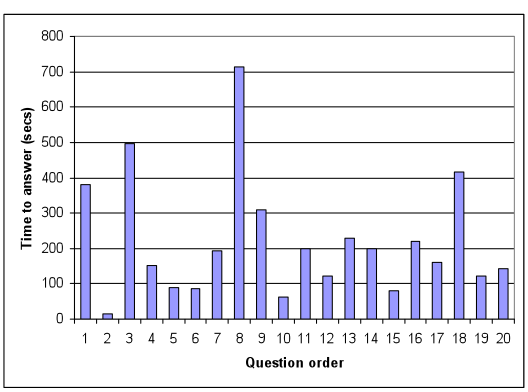

Figure 13: Time spent answering questions by an individual student (median time = 178 seconds)

The ability to monitor when students choose to interact with PeerWise provides insight into how valuable they perceive the contributions of their fellow students. Any voluntary use provides a good indication that the peer-created resource is valued. Consider Figures 14 and 15 which show the number of questions contributed and the number of questions answered in an introductory programming class over a 5 week period. The deadline for the assessment of student contributions was 2nd June, and the final examination for this course was the 14th June.

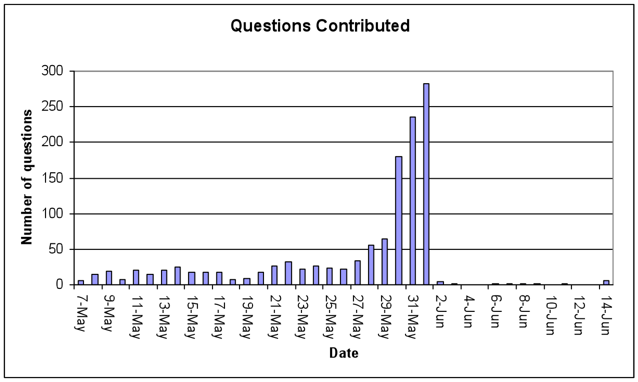

Figure 14: The number of new questions contributed each day (assessment deadline 2nd June)

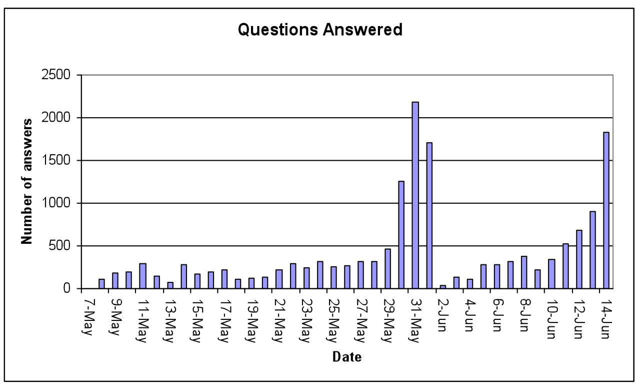

Figure 15: The number of questions answered each day (assessment deadline 2nd June, examination 14th June)

In Figure 14, it is clear that very few new questions were contributed to the repository after the assessment deadline indicating that this activity is not perceived to be of value to students given the effort that is required to develop original questions. However, as illustrated in Figure 15, the system was voluntarily used heavily prior to the final examination. Students see real value in spending their limited revision time using the resources that have been constructed by their peers.

4.4. Content quality

One concern that instructors may have regarding the validity of CBPs is the quality of resources created by students. If student contributions are often erroneous or of low quality, then a resource based on them is likely to be of little value. While it is inevitable that there will be a variety of quality amongst all student contributions, we have found that students generally produce resources of good quality [14].

We have found students to be good at evaluating the quality of resources and, when shared with other students, these evaluations can be used to locate high quality content. In PeerWise, students can rate questions on a 6 point scale, from 0 to 5. We found a strong correlation (0.54) between the ratings assigned by students to a sample of questions, and the ratings assigned by academic staff to the same set of questions. In addition, students make use of the ratings assigned by their peers when choosing which questions to answer. Figure 16 shows the average rating and number of responses for a set of 329 questions developed by students in an introductory programming course. This clearly illustrates that poorly rated questions do not get answered very frequently, whereas only highly rated questions become popular.

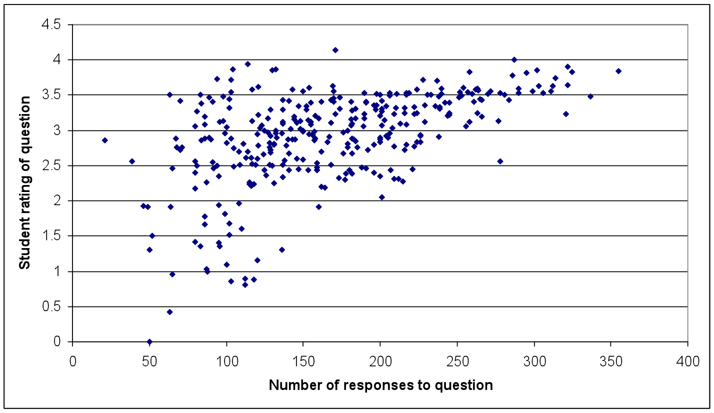

Figure 16: The average rating and number of responses for a complete set of 329 questions

With Aropa, the quality of the items that are available for review can be considered separately to the quality of the reviews themselves. Having a variety of items available for review is, in fact, the feature that students have highlighted as being most beneficial. Therefore, it is not necessary for all of the items to be high quality, as low quality items provide an opportunity for thoughtful and constructive criticism. The quality of the reviews generally seems to be satisfactory, with around 10% of survey respondents noting that they felt the competence of their peers was doubtful.

6. Future work

The tools described in this paper provide complementary environments that engage students in different forms of collaborative learning. As web-based systems, they are available for students to use wherever and whenever they like, and scale to large class sizes with no additional load on instructors. In using these tools, students are encouraged to become both receivers and generators of knowledge within a learning community.

We are beginning to understand how students perceive the activities supported by these tools, and have enough experience now to make recommendations on how they can best be used in a range of classes and subject areas. We still need to gather more evidence on the efficacy of these tools, and try to understand more about the way students use them and how they can be improved and applied more effectively.

REFERENCES

[1] Arthur N. Using student-generated assessment items to enhance teamwork, feedback and the learning process. Synergy, 24:21–23, 2006. www.itl.usyd.edu.au/synergy.

[2] Barak M, and Rafaeli S. On-line question-posing and peer-assessment as means for web-based knowledge sharing in learning. International Journal of Human-Computer Studies, 61:84–103, 2004.

[3] Barker LJ, Garvin-Doxas K, Roberts E. What can computer science learn from a fine arts approach to teaching? In: SIGCSE '05: Proceedings of the 36th SIGCSE technical symposium on Computer science; 2005. p. 421–5.

[4] Boud D. Assessment and learning: contradictory or complementary? In P. Knight, editor, Assessment for Learning in Higher Education, pages 35–48. Kogan Page, London, 1995.

[5] Boud D, Cohen R, and Sampson J. Peer learning and assessment. In David Boud, Ruth Cohen, and Jane Sampson, editors, Peer learning in higher education, pages 67–81. Kogan Page, 2001.

[6] Boud D and Falchikov N. Quantitative studies of student self-assessment in higher education: a critical analysis of findings. Higher Education, 18(5):529–549, 1989.

[7] Brookfield S. (1985): Self-directed learning: a critical review of research. New Directions for Adult and Continuing Education 1985;25:5-16.

[8] Chickering AW and Gamson ZF. (Eds.). "Applying the Seven Principles for Good Practice in Undergraduate Education." New Directions for Teaching and Learning, Vol. 47, 1991.

[9] Cohoon JM. Recruiting and retaining women in undergraduate computing majors. ACM SIGCSE Bulletin 2002;34(2): 48–52.

[10] Denny P, Luxton-Reilly A, Hamer J. The PeerWise system of student contributed assessment questions. In Simon, Hamilton M, editors, ACE '08: Proceedings of the tenth Australasian conference on computing education; Wollongong, NSW, Australia; 2008. p. 69–74. (CRPIT; vol 78).

[11] Denny P, Luxton-Reilly A, Hamer J. Student use of the PeerWise system. ITiCSE '08, June 30–July 2, 2008, Madrid, Spain; 2008. p. 73-77.

[12] Denny P, Hamer J, Luxton-Reilly A, Purchase H. The effect of PeerWise activity on exam performance. ICER'08, 6–7 September, 2008, Sydney, Australia; 2008. (to appear)

[13] Denny P, Hamer J, Luxton-Reilly A. Tools that support contribution-based pedagogies. Proceedings of the ACM-IFIP IEEIII 2008, Informatics Education Europe III Conference, Venice, Italy, December 4-5, 2008 (to appear).

[14] Denny P, Luxton-Reilly A, Simon B. Quality of student contributed questions using PeerWise. In ACE2009: Proceedings of the 11th Australasian Computing Education Conference, Wellington, New Zealand, January 2009 (submitted for publication).

[15] Denny P, Luxton-Reilly A, Hamer J, Purchase H. Coverage of course topics with student generated assessment. In SIGCSE'09: Proceedings of the 40th ACM Technical Symposium on Computer Science Education, Chattanooga, TN USA, Mar. 2009. ACM. (submitted for publication).

[16] Falchikov N. (1995). Peer Feedback Marking: developing peer assessment. Innovations in Education and Training International, 32(2). 175-187

[17] Falchikov N., and Boud, D. Student self-assessment in higher education: a meta-analysis. Review of Educational Research, 59(4): 395–430, 1989.

[18] Fellenz M. Using assessment to support higher level learning: the multiple choice item development assignment. Assessment and Evaluation in Higher Education, 29(6):703–719, 2004.

[19] Freeman MA and McKenzie J, (2002) "Implementing and evaluating SPARK, a confidential web-based template for self and peer assessment of student teamwork: benefits of evaluating across different subjects", British Journal of Educational Technology, 33 (5), pp. 553-572.

[20] Hamer J, Kell C, Spence F. Peer assessment using Aropa. In: ACE ’07: Proceedings of the ninth Australasian conference on computing education; Darlinghurst, Australia; 2007. p. 43–54.

[21] Hamer J, Cutts Q, Jackova J, Luxton-Reilly A, Purchase H, McCartney R, Riedesel C, Saeli M, Sanders K, and Sheard J. Contributing student pedagogy. To be published in SIGCSE Bull., 2008.

[22] Horgen SA. Pedagogical use of multiple choice tests – students create their own tests. In P. Kefalas, A. Sotiriadou, G. Davies, and A. McGettrick, editors, Proceedings of the Informatics Education Europe II Conference. SEERC, 2007.

[23] Kern, VM, dos Santos Pacheco, RC, Saraiva, LM, and Pernigotti, JM. Peer review in computer sciences: Toward a regular, large scale educational approach. In Francisco Milton Mendes Neto and Francisco Vilar Brasileiro, editors, Advances in Computer Supported Learning, pages 45–65. Information Science Publishing, 2007.

[24] McKeachie, WJ. 1999. Teaching Tips: Strategies, Research, and Theory for College and University Teachers. 10th ed. Boston: Houghton Mifflin.

[25] Scorcinelli, MD. “Research Findings on the Seven Principles,” New Directions for Teaching and Learning, v47, 1991, 13-25.

[26] Sluijsmans D., Moerkerke G., Dochy F., & Van Merriënboer J. (2001). Peer assessment in problem based learning. Studies in Educational Evaluation, 27, 2, 153-173.

[27] Stefani LAJ. Assessment in partnership with learners. Assessment and Evaluation in Higher Education 1998;23:339-50.

[28] Topping K. Peer assessment between students in colleges and universities. Review of Educational Research, 68(3):249–276, 1998.

[29] Towler L., and Broadfoot P. (1992) Self-assessment in the primary school. Educational Review, 44(2), 137–151.

[30] Wenger E, McDermott R, Snyder WM. Cultivating communities of practice: a guide to managing knowledge, Harvard Business School Press; 2002.

[31] Yu F, Liu Y, and Chan T. A web-based learning system for question posing and peer assessment. Innovations in Education and Teaching International, 42(4):337–348, 2005.