Peer Review in Practice: eight years of experiences with Aropä

Helen C. Purchase and John Hamer

School of Computing Science, University of Glasgow
31st January 2017

Abstract

Aropä is an online peer-review system provided free, worldwide, since 2009. Voluntarily developed and maintained by two Computing Science academics, it has significant global reach: from Australia and New Zealand to the US and Canada, including Malaysia and Spain in between. Its use covers a vast range of subjects: Anthropology and Accounting, Classics and Commercial Law, Human Biology and Health Policy, Palaeontology and Pharmacology. To date, the system has supported the development of critical thinking and reflective skills in over forty-two thousand students worldwide.

1. Introduction

The Aropä project has actively supported online peer-review activities since 2009, permitting multi-institutional access from 2010. Aropä is offered free, worldwide, and is wholly designed, developed, and maintained by two academic Computing Science staff members at Scottish university. At the time of writing[1], 1,160 successful peer-review assignments devised by 106 instructors have been supported at 20 institutions across the world, in nine countries. Large classes are easily supported; for example, in August 2015 948 students each submitted a report for a Commercial Law course at The University of Auckland, and together wrote a total of 2,716 reviews of their peers' work within a week.

In the last two academic years alone, nearly 13,000 unique students used the system to write reviews on their peers' work. Since, as Nichol (2014) argues, the most significant learning outcome from peer-review activity is as a result of students writing reviews, the effect of this system on the skill development of thousands of students across the globe is considerable.

This paper first describes the peer-review process, and discusses existing peer-review systems. It describes the initial features offered by Aropa in its version of 2009 (section 3), before introducing and discussing the enhancements made in response to user requests (section 4). Looking at these user requests means that we can identify trends in perspectives on peer-review, and these are summarised in eight categories. Section 5 presents usage data, and the paper concludes with observations on the peer-review process, and guidance for instructors.

---

[1] All data is correct as at 31st December 2016.

2. Peer review and Peer review systems

2.1 Peer review

In its basic form, the process of peer-review entails an author submitting an artefact (for example, an essay, a research article, a computer program, a presentation), and their peers stating their opinion of it. It is common practise both in academia (especially with respect to research publication) and in industry. In the case of student peer-review, students are typically first provided with an assignment specification. They then produce the artefact as required by the specification, and submit it (as would be the case with typical assessment activities). The set of student submissions are distributed amongst all students, and each student provides feedback on one or more of their peers' submissions, with reference to a set of criteria. At the end of this process, students reflect on the range of feedback that has been provided on their own submission.

Students therefore take on a variety of roles:
- As author – in submitting an assignment
- As reader – in examining submissions written by peers
- As critic – in judging submissions against a set of criteria
- As the criticised – in reflecting on feedback from peers

Benefits of student peer-review activities include the development of evaluation, critical thinking, written expression and other transferrable skills (Nulty 2011), and individual experience reports over a range of subjects describe successful use of peer-review (e.g. Ramm et al, 2015, Walker & Simpson, 2013).

2.2 Peer review tools

The experience of using an online tool for peer reviewing assignments is qualitatively different to both face-to-face reviewing and using pen-and-paper to mark. Several online tools for peer review have been developed in the past couple of decades; in many cases, these have been developed in-house by computing science academics for use in their own classes. Useful reviews of these tools can be found in Hains-Wesson (2014) and Luxton-Reilly (2009).

For some tools, while their use has been reported in academic publications, there is no evidence of their further use outside that of the specific peer review activities conducted by the paper authors or system developers; examples in this category are PeerGrader (Gehringer, 2000), Web-SPA (Sung, Chang, Chiou, & Hou, 2005), OPAS (Trahasch, 2004), CeLS (Ronen, Kohen-Vacs, & Raz-Fogel, 2006) and PRAISE (de Raadt, Toleman, & Watson, 2005).

However, five systems initially developed and used in-house have over the years been distributed and used more widely: SWoRD (Cho & Schunn, 2007) and PeerScholar (Pare´ & Joordens, 2008) are now licenced pay-for-use products (Panther Learning and Pearson respectively). PRAZE (Mulder & Pearce, 2007), OASYS (Bhalerao & Ward, 2001) and Aropä (Hamer, Purchase et al, 2014) are offered through their university connections (Melbourne, Warwick, Auckland/Glasgow respectively), with facilities to extend their use to other universities worldwide.

PRAZE is administered and managed by the University of Melbourne Learning Environments Unit, and offers extensive online documentation. There is no indication as to how widely it is used at other universities. OASYS is open source, and can be downloaded from sourceforge: it does not

appear to offer support to users. Aropä is hosted by the Universities of Glasgow and Auckland, and administered, managed and supported by the authors; although it is open source, it has proved more effective if hosted locally.

Based on our analysis of these systems, important features of an effective peer review system are:

- Anonymity. Authors should not know who their reviewers are, but the flexibility for reviewers to know who the authors are is occasionally useful.
- Allocation methods. Allocations can be done by individual, or by topic, or by group, and being able to specify the number of reviews that every student should complete is useful. There needs to be flexibility with regards students who need miss specified deadlines.
- Submission methods. It should be possible for any type of file to be required for submission.
- Marking criteria. Instructors should be able to specify a marking rubric for reviewers to follow.
- Mark calculation. A mark for every student can be calculated based on marks given by student reviewers (with discrepancies highlighted).
- Backwards feedback. Authors can respond to their reviews, and reviews can be marked. There should be a facility for dialogue and discussion.

The five tools listed above that have been extended for use after their presentation and publication as research articles all mostly satisfy these core criteria.

The two authors of this paper are the sole designers, developers and maintainers of the Aropä system; we have complete control over it, can observe all activity, and have direct communication with the instructors who use it. We are therefore in a unique position to report on the range and scope of the peer-review activities, including trends, emerging issues and individualisation. During the past eight years we have adapted the system in response to several instructors' individual requests, each request revealing the instructor's perspective on the nature of a successful peer-review activity.

3. The Aropä System

In its basic form, Aropä supports the principal peer-review activity: anonymous, randomly allocated peer-reviewing, based on a rubric devised by the instructor, with an interface that allows students to upload their submissions before the submission deadline, write their reviews before the review deadline, and then view the feedback given to their own submission by other students.

Aropä, first created in 2007, was originally intended for use by Computing Science academics. In 2009, its use was extended to other subjects, and in 2010, provision was made for multi-institutional access. Since 2010, it has stabilised to an extent that effective use typically requires no input or advice from the Aropä team.

The steps comprising an Aropä assignment are:
1. The instructor specifies the
   - submission and review deadlines
   - a rubric to guide students in their reviewing
   - author/reviewer allocation method (random, groups, tags)
2. Students work on their submissions, and submit them to Aropä
3. At the submission deadline, reviewers are automatically allocated to submissions

4. Students then use the system to read their peers' submissions, and write reviews based on the rubric
5. After the submission deadline, students read their peers' feedback

3.1 Infrastructure

Aropä is hosted on two servers, at the University of Glasgow and at the University of Auckland. Documentation pages are hosted at Glasgow.

The system is entirely web-based, and there is nothing for either instructors or students to install. There is no requirement to integrate with existing institutional authentication systems or course management systems. Aropä can be run as a self-contained system. However, support for several common institutional authentication protocols (e.g., LTI and LDAP) is provided. Class lists are entered by the instructor, using a simple cut-and-paste from an existing class roll. Experience has taught us that classes often include a number of informal enrolments, or students who for some reason or other do not appear on the class list. Such students can be entered manually by the instructor, and gain immediate access to the Aropä system. Independence of centralised systems also makes Aropä largely immune from outages in systems over which we have no control.

The system itself is written in PHP, and uses a MySQL (or MariaDB) database. We have recently rolled out a refreshed user interface, using the Bootstrap CSS framework. We are particularly fortunate to be a pair of Computing Science academics who both have a research interest in education as well as expert skills in software development and user interface design. This has allowed us to maintain close control over the ongoing development of the Aropä software. The system itself has become "hardened" over many years of use, and provides high availability to institutions throughout the world.

When a new institution is on-boarded, we create a self-contained "institution" context, which specifies the institution name, logo, time zone and authentication methods. All student identifiers and course material are local to that context.

Instructors who communicate directly with us are provided with a "getting started" guide, and are given a "blessed" status in the system. Blessed users are able to create new classes and add new instructors to classes they create. Users who are introduced to the system indirectly are not blessed, and are unable to create new classes. This simple mechanism serves to keep us abreast of new adopters while minimising the need to be involved in routine day-to-day management. We recently received an email from a lecturer from the University of Auckland to say she has been using the system for the past four years and asking if there is way for her to create a new class each year without having to ask her colleague? This is precisely how we like to see the system used – with new adopters able to get started with help from a colleague, while still encouraging them to (eventually) make contact and tell us how they are getting on.

3.2 Standard Features

Aside from the basic peer-review functionality, Aropä provides flexibility in many of its features:

- The instructor can specify the required form of the authors' submission (for example, pdf, Word, PowerPoint). The system will warn students who attempt to submit in the wrong format, or the wrong number of files.
- The number of reviews to be written by each student can be specified, with an option for self-review. Aropä will randomly allocate submissions to ensure that all students are asked to review the correct number of submissions. A pairing between a submission and a reviewer is known as an 'allocation'.
- If students are working in groups, they can both submit their work as a group, and/or review as a group.
- Submissions can be tagged by topic, so that students only review on the topic relating to their own submission.
- The rubric can include radio-button 'closed' responses, text-box 'open' responses, and file upload.
- Students can see both the comments and marks in their reviews, or only the comments – this is useful if an instructor wishes the authors to focus on the qualitative responses rather than the numeric ones.
- A range of different author-reviewer pairs can be specified: authors can be everyone in the class or one member from each group; reviewers can be everyone in the class, or only those who uploaded a submission.

The practicalities of running peer-review activity are also supported:

- The allocations automatically created by Aropä can be adjusted manually if desired.
- Instructors can monitor the progress of the assignment and, at any time, can see which students have uploaded submissions and which reviewers have written reviews. These tasks are particularly important for first-time instructors (who are typically anxious as to whether the students will engage in the activity at all) and if the instructor wishes to send reminders to the class if the progress is slow.
- Quantitative results from the responses to closed questions can be downloaded in a spreadsheet. A facility for eliminating marks from reviewers with outlying marks is provided.
- The instructor can download all the submissions and all the reviews.

3.3 Advanced features

Aropä also offers some advanced features that are not commonly found in other peer-review systems.

- Reviews themselves can be marked. A secondary assignment can be created that takes as input the reviews from the primary assignment, and allows students to mark these reviews. This is useful as part of training students to become constructive reviewers, and allows students to receive feedback on the quality of the reviews they provided.
- Students can see the reviews from other students on the submissions that they themselves reviewed.
- Instructors can 'impersonate' students so as to see what they see at any time in the process.
- Access to the assignment can be given to guests who have read-only access – this is particularly useful in showcasing the peer-review activity, and in assisting instructors new to the system.

4. Enhancements and Trends

4.1 The iterative development process

The most interesting and significant aspect of running this project has been the development of additional system features at the request of individual instructors over the course of several years. Many requests for new features are suggestive of the particular manner in which instructors view the purpose and risks of peer-review, and of their anxieties (and those of their students). The several requests for improved infrastructure requests (for example, ease of uploading class lists) are indicative of increased use of the system, and the need for the process to be made quick and simple for instructors who use it on a regular basis. In addition, we receive many requests that are pedagogically motivated, and which suggest trends in the range of opinions regarding the effectiveness and usefulness of peer-review.

In providing system support to our users and discussing their requirements with them, we have also advised on appropriate peer-review processes – particularly important for first-time users. In doing so, we have gathered a wide range of examples and precedents that we can discuss and refer to. Importantly, we have tried not to bias instructors in the way that they conduct their activities, while at the same time sharing our experiences.

4.2 Analysis of Trends

We analysed nearly ten years of email correspondences with Aropä users (from the start of the project until 31st December 2016), as well as system change logs, with the aim of identifying:

- All system enhancements requested by instructors, and the justification given for them;
- Discussions about the pedagogical motivation for the enhancement;
- Whether the enhancement was implemented or not.

These enhancements were first placed into four categories: as being *pedagogically* motived, *administratively* motivated, *not implemented*, and on our *'wish list'*. A more detailed thematic analysis categorised the enhancements according to eight trends, each discussed in turn below (sections 4.2.1 – 4.2.8, and summarised in table 1).

- The 'right' feedback
- Concerns about plagiarism
- Student concerns
- The graduate attributes agenda
- Pedagogical innovation
- Increased use of quantitative feedback
- Technical solutions to external procedures
- Increasing expectations of high usability

The trends are discussed in order of perceived importance to the instructors.

4.2.1 The 'right' feedback

Instructors and students alike are increasingly sceptical that student reviewers will not give the 'right' feedback. While it is expected that one of the learning outcomes of a peer-review exercise is

that students are able assess critical feedback on their work, filter it, and decide which (potentially contradictory) feedback to take on board, both students and instructors are anxious that students might get feedback that is in some respects 'wrong'.

To address this, we implemented a feature whereby official 'markers' could review students' submissions in addition to the peer-review activity. Official ('right') reviews are written by tutors or instructors who are given a 'marker' role in the system, and can be allocated either all or a proportion of the submissions to review. These markers are directed to the student interface when they log in to complete their reviews.

In the case of a first year Classics assignment, when the review-marking feature was used for tutors to assess the quality of students' reviews, students were so concerned that they might take on board reviews of their work that tutors later considered 'wrong' that they requested a feature whereby they could see how the reviews of their work had been assessed – thus enabling them to see whether each review was consider 'good' or not (and therefore whether it should be taken seriously). Our own view is that an important part of the peer-review process is that students need to sift through potentially contradictory reviews and decide themselves what to take on board and what to discard. We discussed our perspective with the instructor, but after three consecutive years of her getting this request from students, we implemented this feature for her. She subsequently decided not to use it, having, through her experience with the system, eventually come around to our own view. Interestingly, the more recent cohorts of students have not requested this feature.

One way that instructors are hoping the students might get 'correct' feedback is by including correct answers (or a model solution) within the rubric itself, thus ensuring that all students know what the right answer to the question is before they provide feedback. Effectively, in this case, the peer-review process is used as marking process (with respect to a 'correct' answer), thus focussing on quantitative marks rather than qualitative feedback. We had not originally anticipated that a model solution would be used in rubrics; indeed, most of our prior peer-review examples comprised submissions of essays or other artefacts for which a single 'correct' model solution would be inappropriate. To facilitate this process for a Molecular Biology assignment which included a model solution at the top of the rubric, we implemented a feature whereby images could be included in the rubrics.

4.2.2 Concerns about plagiarism

Instructors are increasingly concerned about the fact that peer-review might encourage plagiarism, especially if the peer-review process relates to a draft that is revised in response to feedback, and then submitted for summative assessment. Indeed, anecdotally, this is one of the most common reasons instructors give us for not wishing to engage in peer-review activity at all. We also hear plagiarism concerns expressed by students, especially those in the later years of their degree study.

In response to this concern from a Classics lecturer, we extended the original 'tagged' allocations feature. When students 'tag' their submission according to one of a set of tags specified by the instructor, the allocation method can ensure that students only review other submissions with the same tag as their own: this ensures that students are reviewing on a topic of which they themselves have some knowledge and understanding.

The new 'between-tag' extension ensures that, as far as possible, students will only review essays on a different topic from their own. The between-tag algorithm does the best allocation that is possible:

depending on the distribution of topics amongst the tags, some students may necessarily review on their own topic - but the number of students who do this is minimised by the allocation algorithm.

While this feature was introduced in response to plagiarism concerns, another way of looking at these between-tag allocations is that it gives students a chance to learn about topics that they have not yet engaged with themselves, as part of a Contributing Student Pedagogy approach (Hamer et al., 2008); however, we have yet not heard anyone express this view.

4.2.3 Student concerns

Students are always ready to complain about a pedagogical activity that is new or unfamiliar, and peer-review is no exception. Some of these complaints may be a result of the decisions made by the instructor (for example, the number of submissions a student should review, the length and nature of the rubric, the length of time allowed for reviews); in many cases, we have had requests from instructors to adapt the system so as to address some complaints. One such example has been given above – the provision for students to see how reviews of their own submissions were marked.

One common complaint from students is that they wrote extensive reviews, yet received fewer reviews back than they wrote, since some of their peers did not bother to write reviews at all. While instructors can use external motivation to encourage participation in the peer review process (for example, marks for reviews completed, making completion of reviews a gateway to other assessment items), several instructors requested an internal system-based solution.

Instructors can now indicate that students can access the reviews of their own submission only if they themselves have completed all or some of their reviews. In addition, we updated the process of saving a review, so that a warning is given if a student attempts to save a review that does not include all mark and comment items. The missing items are highlighted with a red border, and if a student insists on saving an empty or incomplete review, the instructor can see this information when monitoring the reviewing process.

Part of dealing with students' complaints is managing their students' expectations as to what to expect from the reviewing process in advance. The ideal set of 'balanced' author-reviewer pairings is one where number of reviews a student is required to write is the same as the number of reviews to be written about their own assignment. In very rare cases, this is impossible to achieve – for example, when all students in the class (even non-submitting students) are allocated to review submission from a subset of the class. Some students therefore complained that they were doing more reviews than others; as a result the Aropä allocation algorithm now prioritises student effort, ensuring that all students review exactly the number of submissions that the instructor has specified (even if this means that students may get a varying number of reviews back)

Providing additional information to the instructor about unbalanced allocations can also be useful, so that the instructor can warn the students in advance. On our wish list is a feature that will inform the instructors if there are any problems with the allocation balance. This sometimes happens when several submission extensions are given after the submission deadline. Each time a submission extension is given, a new allocation is created – this allocation cannot affect any of the previously created allocations where the reviewer has already either read the submission, or completed the review. The more time that passes between the submission deadline and the giving of a submission extension, the more difficult it will be to create a balanced set of allocations.

Information on inappropriate 'between-tag' allocations will also be useful for the instructor. The between-tag algorithm does the best allocation that is possible. However, unless there are an equal number of submissions for each tag, it is impossible to satisfy the between-tag criterion exactly. The number of students who need review same-tag submissions is minimised by the allocation algorithm. Instructors would benefit from knowing how many students will necessarily have to review submissions with their own tag, so that they can warn students of this possibility in advance.

Students sometimes say that they think that their reviews are 'unfair'. The instructor of an Anatomy class requested a feature which would give the students the opportunity to (anonymously) respond to their critics – this was not mandatory, and had no effect on the students' marks, but she felt that this opportunity would both ensure that the students engage with their feedback, and also allow them to provide a response to criticism. Students engaged with this process willingly, demonstrating that, when given a chance, they can show that they take the comments provided by their peers seriously and actively wish to engage in the review process.

We have no expectation that these enhancements will remove student complaints entirely, but our experiences is that once the process is embedded in a course, and repeated for a couple of years, complaints from the incoming cohort decrease.

4.2.4 The graduate attributes agenda

There is increasing emphasis on facilitating the development of 'Graduate Attributes' or 'Transferrable Skills' as part of an undergraduate degree programme (Hager and Holland, 2007). Peer-review is an excellent vehicle for encouraging such skill development, and many of the feature requests from instructors demonstrate their commitment to developing critical analysis skills. These include, as mentioned above, facilitating a means for students to provide a response to their reviewers.

Several instructors wish to emphasise to their students that the quality of their reviews was important. While a review-marking feature was available in the early version of the system (allowing students to comment on the quality of their peers' reviews), this did not extend to allowing tutors to mark the reviews. The instructor for a Classics class wanted to award course marks the students' critical analysis skills, thus not only embedding transferrable skills in the curriculum, but also formally assessing them.

In some cases, it is simply the engagement in the review process, rather than the quality of the reviews, that is of interest to the instructors, and for which they give marks. The facility for monitoring the reviewing process was enhanced so that the instructor could not only see which allocations have a completed review, but also can see which reviewers have read their allocated submissions, and, after the review deadline, which students have read the reviews of their own work. This is useful to instructors if reading and responding to reviews is considered an important part of the assignment. It reveals how the peer review activity is progressing, and which students are engaged enough that they care about seeing their feedback. This review completed/read information can be downloaded to a spreadsheet if marks are to be allocated for engagement in the activity.

Students are sometimes expected to use their review feedback to enhance a draft, thus focussing on the skills of analysis (interpreting the reviews), critical judgement (selecting which comments should be addressed), reflection (reflecting on the quality of the draft) and revision and integration (making appropriate changes without complete re-writing). An instructor in Biomedical sciences wished to

assess the extent to which students had used their reviews in producing their final document (submitted for formal assessment).We implemented a feature whereby submissions could be downloaded together with their reviews so as to facilitate easy access to all the information that she needed when assessing each student.

Subsequent to the 'response to reviewers' feature implemented for the Anatomy assignments (described above), an Accounting lecturer used this feature as a means for students to describe what specific changes they made to their initial draft in response to the reviews.

Only one instructor has highlighted the issue that students might benefit from some training in critical analysis before reviewing peers' work. We implemented a feature whereby an instructor in English as Second Language can submit a document written by himself, and then ask the students to review it. His own critique of this document is made available to the students as soon as they have written their own review (using a new 'locking facility' which allows students to lock-in their own review, removing the requirement to wait until the review deadline) so that they can compare it with their own review.

4.2.5 Pedagogical innovation

It is clear that many of the instructors who introduce peer-review into their classes are keen on promoting educational innovation. Indeed, many go on to publish their experiences with Aropä in the educational scholarship literature (e.g. Patterson, 2009), often without our knowledge. We have included a facility for instructors to define and administer a survey to their students, as a way of eliciting student feedback on the experience, and, on request, have defined a standard survey.

We have made provision for effective re-use of the system; that is, instructors using it for the same assignments in subsequent years. Much of the work of creating a peer-review activity lies in the creation of a suitable rubric. Set-up effort is reduced by the inclusion of a drop-down list of previously created rubrics that can be loaded and re-used. Having discovered that some instructors who wished to repeat their peer-review activities the following year were simply changing the list of students associated with the course, and the submission dates (a process that does not remove the allocations and reviews from the previous year, and so compromises the subsequent allocation algorithm), a recent addition is a feature that copies an entire class and its assignments for reuse, while removing the student list, all submissions and reviews, and all specified deadlines.

While not exactly pedagogical innovation, some instructors use Aropä simply for online marking – as a convenient and efficient way of collecting submissions, marking them, and distributing the feedback. Anecdotal feedback suggests that this is because Aropä is much easier to use than other online marking tools – an instructor in English Language has told us that he as abandoned Moodle for online marking, and uses Aropä exclusively instead. Although we do not actively advertise Aropä as an online marking tool, we have no objection to instructors using it in a way that suits them (and indeed, use it this way ourselves).

4.2.6 Increased use of quantitative feedback

Qualitative comments are often more useful in feedback that quantitative responses, and all rubrics use a mixture of both, with most rubrics containing more comment box items than radio button items. However, there is increasing use by instructors of the quantitative data. In particular, we have introduced a features whereby instructors can associate marks with each of the radio button options

in the rubric; thus, for example, in the case of a set of three radio buttons, 10 marks could be given for the 'excellent' option, 8 marks for 'very good' option, and 1 mark for 'fair'. This increases the flexibility of the rubric, especially if the quantitative marks given by reviewers are to be subsequently used in any formative or summative manner.

We have resisted, however, allowing radio-button rubric items that solely represent marks. An Engineering instructor wished to have a rubric feature where the student would specify a mark (eg, a value from 1 to 10) by using a drop-down list that included all available marks. We resisted this enhancement on the basis that giving marks out of ten (for example) suggests that there are ten distinct levels of achievement, and a drop-down list of numbers does not give any space for describing what these achievement levels are. Thus, there is more chance of unnecessary variability in the marks given by students, since it is not clear what each mark (from one to ten) actually means. The existing radio-button rubric elements require that a descriptive label be associated with each option.

4.2.7 Technical solutions to external procedures

We get several requests for Aropä to provide additional support to instructors so as to facilitate their conduct of a peer-review exercise: this is support over and above the basic features required for successful peer-review, and is often requested so as to make administrative procedures easier – collecting student feedback through an Aropä-facilitated survey (as described above), is one such example.

For example, some instructors give the full specification for an assignment in the text which is shown to students at the time that they submit their file to Aropä; we had expected that the assignment specification would be given externally to the system. On request, we extended the system so that this text area could include images to facilitate this use case.

Motivating students to produce appropriate reviews is an area in which instructors commonly request support. One such feature, whereby motivation for completing reviews is facilitated by the system by limiting access to reviews to only those students who produce reviews has been mentioned above. A Biology instructor asked whether instructors could specify a minimum comment length for individual sections of the rubric, since students had complained that the comments that they have received from their peers were too brief. While it would be possible to implement this, we have resisted on the basis that if student is prevented from submitting a review because a comment is not long enough, we would not want to encourage them to simply fill up their word quota with irrelevant text. We would like students to think carefully about what they write, and setting a minimum comment length suggests that we do not care what they write (simply how much they write).

As a way of encouraging students to submit and review, we are often asked if Aropä could send emails to students to remind them of deadlines. We have resisted this so far. Providing this facility would require that either the student or the instructor enter into Aropä an email address for each student in the class, since Aropä is not integrated with existing Learning Management systems. This means that students for whom the system does not have the correct email address will not receive any notification, thus making the system open to error. And we would not wish instructors and students to rely on a system that is potentially unreliable.

An early addition to the system was a facility for giving extensions to students, so that they could submit their file or write their reviews after the official deadline. Aropä automatically adjusts the

allocations to take extensions into account, and the status of extensions can be monitored – including showing when they have expired without submission. Most instructors are very grateful for this feature and the ability to award extensions to individual students. However, Aropä is non-forgiving for students who submit just after a deadline – a couple of seconds will mean the difference between being able to submit a file, and not being able to. Many instructors are quite happy to allow students to submit 'just a little bit' late, and would rather be able to specify an implicit extension period than have to deal with giving individual extensions to students who have only just missed the deadline. Allowing instructors to specify a brief 'implicit' extension period is one of our planned features.

There are some aspects of the system use that it would be impossible to support. Aropä facilitates anonymous peer-review; however, if a student (deliberately or otherwise) puts their own name at the top of their submission document or names their submitted file with their own name or student number, Aropä cannot remove this information so as to protect their identity. In addition (although less obvious) meta-data associated with submitted files cannot be removed – although only those students who deliberately wish to subvert the process will seek this information. Some instructors (and indeed students) have expressed concern about this; we encourage use of the Aropä online editor for submissions where possible (rather than submission by file upload), since this will guarantee anonymity.

4.2.8 Increasing expectations of high usability

Request for interface improvements are always welcome as part of an iterative development model, and the interface design, interaction model and user manuals have developed extensively over the years.

The most significant improvement was the inclusion of a WYSIWYG editor for rubrics, to replace the original XML-based format. Since the system was originally developed for use in a Computing Science Department, a formalised syntax was adequate. The system was very quickly adopted in other departments, and so a more user-friendly method of rubric specification was necessary. The previous approach of instructors having to explicitly launch the allocation algorithm after the submission deadline was abandoned in favour of a process whereby the allocation are automatically created – we found that many instructors were forgetting this crucial step.

Other usability enhancements include: cut-and-paste entry of student, marker and group identifiers, distinguishing different sets of radio buttons in the rubric by colour, a facility for instructors to reset their students' passwords, listing the students who have not yet submitted and those who have missed extension deadlines, and a clear and simple interaction flow for common tasks. A simple editor is provided for submissions that do not require diagrams or tables or advanced formatting. We increased the maximum file upload limit at the request of a Biology instructor, and, importantly, made instructions regarding expected file type (and file size limit) more explicit.

An interesting usability enhancement was requested by an Engineering Mathematics instructor, so as to allow for submissions to be annotated. A feature whereby a document can be uploaded into a review was implemented. This allowed students to print their allocated submissions, hand-write comments, and then scan and upload these into the Aropä review form. We understand that this feature is no longer used by this instructor, who had difficulty persuading his class of the benefits of peer-review. The additional administrative load for the students in printing and scanning documents was seen by the students to be unreasonable. Tablet technology may allow for easier annotation of submissions in future developments.

Not all suggested interface enhancements are desirable (e.g. drop-down lists of marks, as above), or, indeed, possible (e.g. automatic integration with various universities' learning management systems so that class lists can be automatically imported).

System enhancement	Description	Category
Pedagogically motivated enhancements		
Tutor marking of submissions	Some users are given special 'tutor' status. This means that they can login and review student submissions, using the student reviewing interface.	The 'Right' Feedback Pedagogical innovation
Tutor marking of reviews	A separate second-level process can be defined that enables tutors to comment on the quality of the reviews.	The Graduate Attributes agenda
Tutor marking of group submissions	Tutors are able to mark any type of submission.	Pedagogical innovation The 'Right' Feedback
Inserting images into rubrics	Rich text and images can be used in the rubric – usually to specify a model solution	The 'Right' Feedback
'Between-tag' allocations	Students can tag their submissions by topic, and will only review submissions on other topics.	Concerns about plagiarism
Response to reviews	Students can write a response to their peer reviewers.	Student concerns The Graduate Attributes agenda
Access to review marking	Students can see how tutors have marked the reviews that they have received, and so see which ones are considered 'valid'.	The 'Right' Feedback Student concerns
Monitoring the reviewing process	At any time, instructors can see who has completed their reviews, and, after the review deadlines, who has read the reviews on their own submissions.	The Graduate Attributes agenda
Non-student authors	Students can review artefacts submitted by an instructor, so as to first gain experience and training in the review process.	The Graduate Attributes agenda
Save reviews with submissions	Instructors can easily see a submission together with its reviews.	The Graduate Attributes agenda
Upload file feature in rubric	Reviewers can upload a file into the rubric as part of their review	Increasing expectations of high usability
Limiting access to reviews	Only those students who have done some or all of their reviews are able to see the feedback on their own submission.	Student concerns Technical solutions to external procedures
Marks allocated to radio buttons	Mark values can be allocated to each option in a radio button list.	Increased use of quantitative feedback
Empty reviews	Reviews that are incomplete or empty are highlighted in the instructor view.	Student concerns
Standard Survey	The Examples Class that all new instructors are given guest access to includes a standard survey that elicits feedback on the experiences of peer-review.	Pedagogical innovation Technical solutions to external procedures
Administratively motivated enhancements		
Re-using rubrics	Instructors can reload rubrics they have previously defined, or load and adapt a rubric created by someone else.	Pedagogical innovation
Extensions	The deadline for submission and review can be extended for specified students.	Technical solutions to external procedures
WYSIWYG editor for rubric creation	The rubric editor shows exactly what the reviewers will see when writing their reviews.	Increasing expectations of high usability
Automatic allocations	Allocations between authors and reviewers are created automatically after the submission deadline: no intervention from	Increasing expectations of high usability

	the instructor is necessary.	
Password control	Instructors can reset student's passwords, and can set passwords for people with marker role.	Increasing expectations of high usability
Identifying errant students	Information about students who have not yet submitted or who have missed extension deadlines is shown to the instructor	Increasing expectations of high usability
Interface enhancements	The interaction model has been carefully defined and refined so that the flow of tasks is intuitive.	Increasing expectations of high usability
Roll-over	Classes can be copied in their entirety for use in the following year.	Pedagogical innovation
Documentation	The documentation is updated regularly, with a growing list of FAQs. There is student documentation as well as instructor documentation online.	Increasing expectations of high usability
Allocation priorities	Students are always allocated the same number of submissions to review as their peers.	Student concerns
Inserting images into the 'instructions for students' text	Instructors can put their whole assignment specification in the system, to be read before the student submits.	Technical solutions to external procedures
Unsatisfied requests		
Drop-down lists for giving marks (rather than radio buttons)	The rubric allows for drop-down lists that enable a numeric mark to be specified.	Increased use of quantitative feedback
		Increasing expectations of high usability
Specifying a minimum length of comment	Students are unable to complete a review unless their comment boxes contain a minimum number of words.	Technical solutions to external procedures
Identifying information	Students' identity can be determined by information given in the filename, the document itself, or by meta data. Ideally, the system would remove this identifying information.	Technical solutions to external procedures
Emailing reminders to students	Students would be reminded about submission and review dates by email.	Technical solutions to external procedures
On our wish-list		
Implicit extensions	Instructors can specify a brief 'grace period' after submission and review deadlines; this period would not be advertised to students.	Technical solutions to external procedures
Feedback on unbalanced allocations	Instructors would be informed if, for whatever reason, the allocations are created such that some students are required to review more submissions than others.	Student concerns
Feedback on inappropriate between-tag allocations	Instructors would be informed if some students will be required to review submissions on their own topic, and how many students fall into this category.	Student concerns

Table 1. Summary and classification of system enhancements requested by instructors.

4.2.9 Summary

It is clear then, that the priorities and concerns of instructors interested in peer-review are multi-faceted, and are always changing. The current increasing use of technology in higher education means that expectations are likely to increase even further: however, we have no plans to fall in the

trap of providing unnecessary additional functionality. For example, we do not intend to link Aropä users with Twitter handles, to include a discussion forum, or to integrate with in-class response systems. Our focus is on providing a tool that supports peer-review (and does so efficiently and effectively). More functionality means more complexity – and we have seen too many all-singing-all-dancing systems that are clumsy to navigate and to use to wish to make the system more complex than it needs to be.

We would, however, like to be able to integrate Aropä with Learning Management Systems (e.g. Moodle, BrightSpace, Blackboard) so as to facilitate easy class list upload and import of email addresses, but our investigations so far have indicated that simple transfer of information from these systems into the Aropä database is not possible.

5. Aropä Use

5.1 Key Performance Indicators

Nicol (2014) claims that the greatest benefit comes from writing reviews, and so one of our main KPIs is the number of unique students who have written reviews in an academic year. We are particularly interested in the number of instructors who have returned for repeated use (an indication of system usability), and, since our intention is to support large classes in particular, class size. Table 2 summarises the key performance indicators; some additional noteworthy statistics are shown in Table 3, which focusses on a one year period.

Number of successful peer-review assignments	Since 2009	1,160
	in the last two years	513
Number of unique students (worldwide) who have used the system to write at least one review	Since 2009	42,698
	in the last two years	12,880
Number of repeat instructors: that is, instructors who used the system for more than one peer-review activity	Since 2009	66
Number of higher education institutions with at least one successful assignment	Since 2009	20
Largest class size associated with a successful assignment	Since 2009 (August 2015)	948 submissions 2,716 reviews

Table 2: Aropä KPI usage data (at 31st December 2016)

Number of unique students writing at least one review at the university of Glasgow (15/16 academic year)	2,075
Unique students writing at least one review at the university of Glasgow (15/16 academic year), as approximate % of the student body	7.6%
Number of unique students writing at least one review at the university of Auckland (2015 academic year)	3,907
Unique students writing at least one review at the university of Auckland (2015 academic year) as approximate % of the student body	7.4%
Total number of completed reviews in the calendar year 2015	44,396
Total number of completed reviews since 2009	269,781

Table 3: Additional usage statistics

5.2. The nature of the assignments

5.2.1 Assignment configuration

Configuration data from successful assignments is easily obtained from the system. Information about 1,160 assignments was extracted on 31st December 2016, and is shown in table 4. This table demonstrates not only the variety of the types of peer-review activities conducted, but also the flexibility of the system.

	minimum	maximum	median
The students			
Class size	2	1,057	71
Number of submissions in an assignment	1	948	18
Number of reviews in an assignment	1	3284	83
The rubric			
Total number of input items in the rubric	2	60	6
Number of radio button items in the rubric	1	44	2
Number of comment boxes in the rubric	1	27	4
The allocations			
Number of submissions for each student to review	1	8	2
Allocation type: random	812		
Allocation type: groups	199		
Allocation type: same tags	8		
Allocation type: different tags	10		
Allocation type: manual	69		
Deadlines			
Number of days between submission deadline and review deadline	0	83	7

Table 4: Assignment configuration data

5.2.2 Case Studies

The many different ways in which instructors use Aropä is extensive. Several particular assignments have come to our attention over the years.

- Philip Cook from the Auckland University Business School has used Aropä consistently over 12 semesters from Semester 2 2010 to Semester 2 2016 (excepting Semester 2 2013). He has arranged for 20 different Commercial Law classes (both at first and second year levels of study) to use the system, with two or three peer-review assignments per class. Students submit a report giving legal advice on a problem scenario involving various legal issues , and are asked to review three of their peers' submissions. There is a standard rubric for all the assignments: it includes questions like "Are the legal principles correctly identified?" and "Are the arguments developed logically?", with a set of four radio buttons as well as a comment box for students to use in their review. Philip discusses the marking process at length with the students in a lecture session. Not engaging in the review process will lead to students losing marks allocated to tutorial participation; students are also advised that Aropä participation is taken into account when considering particular circumstances regarding final grades (e.g. marginal fails).

- Lisa Hau (who teaches Classics at Glasgow University) has used Aropä every year since 2011, for seven second year classes. Students are given three essay topics to choose from (e.g. Tragedy, FuneralSpeeches, Democracy) and they 'tag' their draft essay according to their topic on submission; as much as is possible, Aropä allocates reviewers to essays with a different tag from their own. Students review two of their peers' essay drafts under six categories (e.g. Knowledge and analysis, Argument, Use of primary literature/primary evidence); for each, a set of radio buttons and a comment box is provided. Each review is then marked by a tutor with respect to fairness, comprehensiveness and helpfulness (using the Aropä 'review marking' feature). Marks are given to the students for the quality of their reviews.

- Abdoul Aziz Fall and his colleagues in the Faculty of Medicine at Cyberjaya University College in Malaysia have used Aropä extensively since 2014, conducting 138 successful peer-review assignments with class sizes totalling 3,154 students. Abdoul's particular use of the system covers medical areas such as Personality Disorders, Emotion, and Social Psychology and Religion. Students are put in groups, and there is one submission per group: the submissions are the slides from the groups' presentations that all students attend. All students belonging to groups that have submitted their slides are invited to review the group submissions, and most students do so. Other use at Cyberjaya includes the submission and review of short statistical reports, and situations where only one report is submitted and all students review that same report. Students are rewarded with marks for participation.

- Todd Whittaker teaches Computing Science at Franklin University, and has used Aropä for three classes per year since the Fall Semester of 2014. These are small classes (typically less than 20 students). There are five assignments for each class: these are draft documents for consecutive stages of a team project; in some cases the submission is a team effort, in others, submissions are done by individuals. Student then submit revised documents for summative assessment. Todd uses Aropä's 'review marking' feature for each of the assignments: every student who reviewed the submitted drafts is randomly allocated four reviews to mark, using a rubric that considers the tone of the review, the extent of constructive feedback, and the accuracy of the review. The marks allocated to the drafts by peers form a small part of the final assessment mark, as do the marks allocated to review quality. Not participating in any step of the review process will result in zero marks for that step.

- Peter Bier teaches the annual "Introduction to Engineering Computation and Software Development" course to all first year engineering students at the University of Auckland. Since 2010, he has used Aropä for classes ranging in size from 575 to 876. Students are required to submit a set of Matlab files – this can be as many as 16 files. Students review the work submitted by three of their peers, with a rubric that covers both style and functionality. With such large class sizes, the standard Aropä feature that allows extensions to be given for late submissions is unwieldly (since each student needs to be explicitly given an extension); Peter gets around this by creating a duplicate 'extension' version of the assignment with later deadlines. In 2016, a total of 2,393 reviews were written by students for this assignment. While students receive 2% of their final grade for completing their, the feedback they receive is for formative purposes only.

5.2.3 Incentives

Aropä cannot, of course, collect information about aspects of the activity that are external to the system and which are communicated to students by the instructor. In particular, we have no way of knowing whether the students have been given any incentive to complete their reviews. The feature that prevents students from seeing their reviews unless they have completed some or all of their own reviews was implemented in 14th August 2012. Since then 161 of 901 successful assignments (17.9%) have used this feature.

A survey of 41 active Aropä instructors revealed that 10 did not offer any incentive to students with respect to engagement in the peer-review activity, but that others used the following methods to encourage students to complete their reviews:

- Marks for participation. This was the most popular category (12) – students are given marks if they complete their reviews. This is clearly a very simple method, since the assessment is simply computational: did the student complete their reviews or not?
- Marks for review quality. The second most popular category (7) – students are given marks based on the quality of the feedback that they produce. This is often done using the 'review marking' feature in Aropä, with tutors or instructors marking the reviews according to a specified rubric.
- Summative marks. Some instructors use the grades provided by the peer-reviewers as summative marks (after excluding outliers), contributing to the students' final grade. This is rare, but appears to be successful, since this method has been repeated in future years.
- Implicit incentive. Students are encouraged to take part because the instructor makes explicit links between this activity and other summative activities. For example, doing reviews gives students access to the rubric used by markers in summative assessment (which they would not get otherwise), or students are reminded that engaging in the peer-review activity will help their performance in later summative exercises. One instructor invites only those students who have completed reviews to an additional workshop where the marking rubric is discussed in detail, thus giving additional guidance to these students for later assessments.
- Mandatory. Some instructors make engaging the peer-review activity mandatory for the award of credit; others simply tell students it is mandatory (with no suggestion that it could ever be optional).

5.2.4 Subjects

One of the remarkable findings of our analysis of Aropä use is the extensive range of subjects for which it is used. Originally developed for peer-review of programming code in Computing Science assignments, it is now used in most subjects offered in Higher Education institutions (see Fig 1). There are a surprising number of medical and life sciences assignments, as well as physiotherapy health sciences.
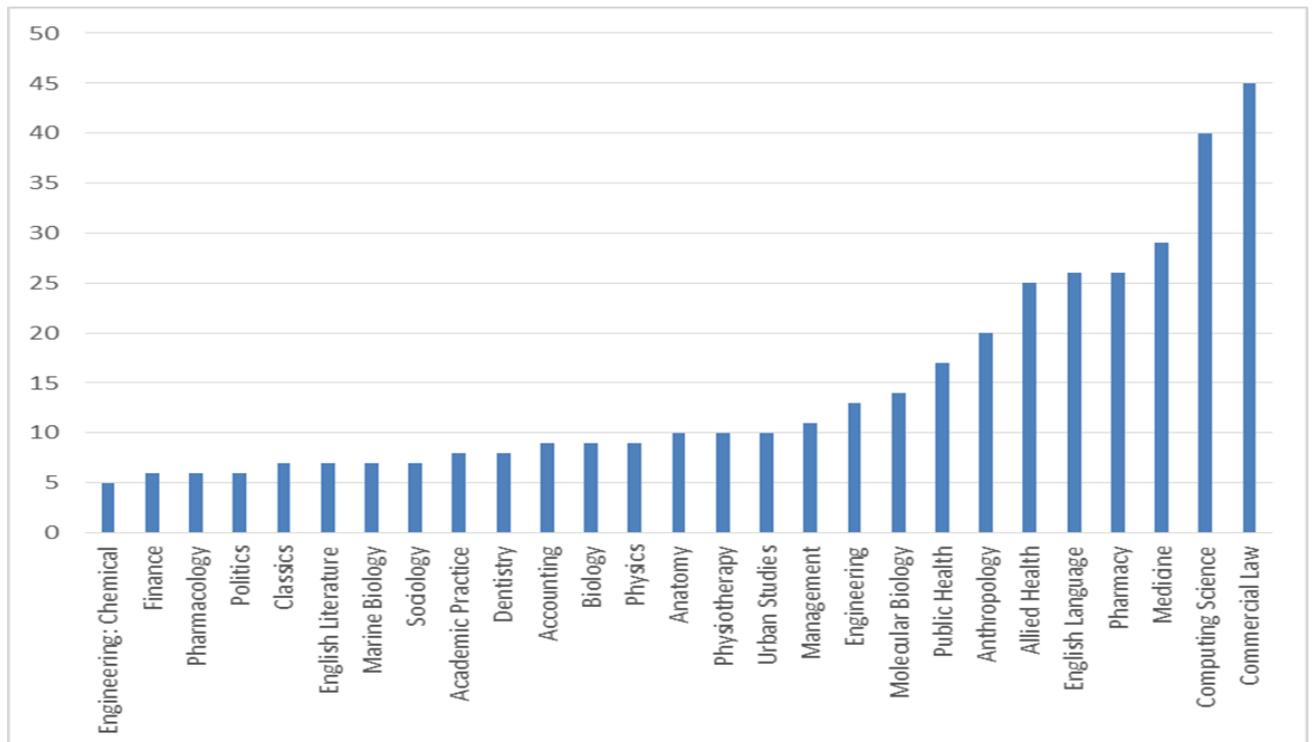
Figure 1: Disciplines which have used Aropä for more than five courses.

5.4 User opinion

5.4.1 Student responses

Many students are apprehensive about peer review at first, but we have found the majority of them recognise the benefits after having taken part. We are able to extract student responses from surveys that are conducted using Aropä: verbatim student comments include:

- Very useful process, went into it reluctantly but was actually good. It didn't take up too much time, and i would recommend to other students.
- It allowed me to learn how to critically examine someone else's work.
- I feel greatful (sic) to all the reviewers of my assignments everytime I read the comments. They are amazing "Little Tutors"!
- I found that by offering suggestions and explaining principles etc.. that students had misunderstood in their essays to be both useful for the student being marked and me personally as you revise your own knowledge at the same time.
- As for the actual system itself I think it's very user friendly and easily the best method of conducting peer reviews that I've came across.
- I felt this was a very useful exercise and should definitely be used in more subjects/ for more reports or practice papers etc.
- Seeing others' thought processes how to go about solving problems enabled me to approach and understand problems that I wouldn't necessarily be able to solve on my own.
- Really beneficial experience to see whether I am on track or not. Overall a worthwhile part of the course!

Students are sometimes sceptical of their peers' analytical and critical skills, and sometimes complain that what they really want is 'expert' feedback. Todd Whittaker (2016) puts it this way: "many feel that the other students aren't qualified to review their project or perceive that minimal

effort is put in to quality reviews. The irony is that the students don't perceive that they aren't qualified to judge others, but just that other's aren't qualified to judge them."

A common complaint from students is that, while they have spent a great deal of time proving extensive feedback, the reviews that they themselves received were sparse. Students often suggest that the quality of the reviews should be assessed so that they would be more useful. (Review-marking is available in Aropä, but not all instructors use it for tutor-marking of reviews).

5.4.2 Instructor responses

The results of recent qualitative survey of 41 instructors were immensely positive, and included comments like:
- incredibly easy to use
- As a first-time user, teaching a very large class, I was initially nervous about what problems might arise, but it proved highly reliable
- I especially appreciate all the personalised help and support I have got from the Aropä creators
- I love it! Thank you for creating it--the only system out there that works for my purposes.
- Easy to learn and use, flexible and focused on academic peer review
- It's great to use! So intuitive and quite straightforward, love using it!
- Personally I think it's a great system. I like the small improvements over time.
- The turnaround times for help and guidance have always been remarkably quick! I think it does a great job of letting staff see how far along the process each student has gone, and the impersonate user function is very useful for troubleshooting.
- It's excellent. I am considering using it for straightforward assignment marking over Moodle right now, as it's so much easier. Easy excel export of marks is essential for us and the Aropä report style is really good.
- Aropä has been great and we are glad it's still going!
- Very dependable and helpful system. The revamp a few years back improved the look greatly and really appreciated the adaptations John built in especially for me. Would hate to lose it.
- It's great - and still better than Peermark.
- Yes thank you for the free use of the system. I think it is a great system and easy to use. The first few years that I used it I needed help from Helen or John and it was always forthcoming, for which I am truly grateful. Now it seems easier to use (or I have got better at using it) and I haven't had to ask for help in recent years. Thank you again.

The only negative comments from instructors included the lack of integration with their own particular university's student enrolment system (so as to facilitate class-list download), and initial start-up learning costs.

Several instructors have published papers reporting their use for Aropä – in most cases unbeknownst to us – in the areas of Music (November, 2011, 2012), Anatomy (Welsh, 2014), Engineering (Patterson, 2009), Physiotherapy (MacDonald, 2013), Scientific writing (Finlay, 2012), and over several disciplines (Huston, 2015).

5.3 Additional observations

Over the many years that Aropä has been actively used in non-Computing Science departments, we have noticed the following:

- The most successful assignments are those where the review period is short (typically less than a week), so that students are reviewing while their own submission is still fresh in their mind.
- Problems with the success of the peer-review activity are often caused 'outside' the system: students forget to submit or write their reviews, some students take the review activity more seriously than others, students violate anonymity by including their own name in their submission, the rubric only includes closed questions so does not permit qualitative feedback, students forget their passwords, the class list provided by administrators at the start of the semester is incomplete etc. While we can give advice to both instructors and students, we cannot dictate how they use the system.
- While many instructors are not shy in contacting us if they have problems or questions, or if they have feature requests, there are many instructors who use the system extensively who we never hear from. For example, we advised the instructor of English Writing 1 at the University of Auckland on using the system in 2009, and he has used it for three courses a year since then. We have never heard from him personally since then – he has reported no problem with his use of the system.
- Some instructors simply use the system as an online submission tool (making no use of the review facilities available), simply because its interface is easier to use that that of other online submission tools. While this is not an intended use of the system, it does indicate high usability.

5.4 Discussion

The usage and survey data all support the view that Aropä is extensively and successfully used, and is clearly appreciated by its users for its ease of use and extent of functionality. The range of approaches to peer-review is extremely diverse (as shown by the variety of configurations used), demonstrating the extensive flexibility that the system offers. While this variety is impressive from an implementation perspective, it is most important from an academic point of view. If an instructor wishes to introduce a peer-review activity for the first time, then they are likely to have an idea as to how they would want it to work: forcing a particular fixed approach would therefore be unreasonable and likely to discourage them.

6. Future Work

There are, of course, many avenues of future work with respect to system enhancements (Table 1), and we expect several more to come our way on a regular basis.[2]

We are, of course, aware that many LMS systems are now including their own Peer Review facilities (e.g. Canvas, recently adopted at the University of Auckland, https://www.canvasvle.co.uk/), although we are often assured that the usability and range of features of Aropä are significantly better – and indeed, instructors at the University of Auckland continue to use Aropä for this stated reason.

---

[2] Two enhancements were suggested and completed during the period of preparation of this article.

However, we are aware that as the LMS technology improves, if it proves impossible to integrate Aropä with learning management systems, instructors will prefer to use a peer-review system that automatically uploads class lists, and downloads marks. We have investigated both Moodle and BrightSpace, but neither have sufficient LTI integration features to permit appropriate integration with Aropä.

We have often been asked why we offer this service for free and have no plans to commercialise it, especially since there are other similar systems which charge for use (e.g. PeerGrade (https://www.peergrade.io/landing). We have a stable system that is highly valued, and we work in a very responsive service mode (with most issues resolved in a day or two). We are passionate about peer-review as a means of developing key cognitive skills, and we want to encourage instructors (even sceptics) to give it a try. Pricing would introduce an unnecessary barrier that may discourage use, and commercialising the system as a business would place us behind a screen of anonymity that would reduce the personal interaction we have with our users.

One instructor recently emailed: "[Aropa being] Free is a phenomenal feature! You probably don't hear it often enough, but I'll say it as often as I remember when I write to you: thank you so very much for the service you're providing to the community through Aropa." (Whittaker, 2016).

7. Conclusions

The benefits of an in-house peer-review system created, developed, and nurtured over time by practising academics is clear. Our approach to the inclusion and availability of system features and to the service that we provide to our instructors is led first and foremost by a firm understanding of the needs of the academic process – from both the instructor and student perspectives.

But the Aropä project is more than simply an academic service – if it were, we would probably commercialise the system. As this article demonstrates, our close collaboration with the instructors and our knowledge of their contexts of use means that we consider the project to be an academic endeavour in itself – allowing us to gather data and comment on approaches to peer review activities worldwide and over several years.

Based on our extensive experiences and analysis of data, our main conclusions are:

- "Peer review" activities do not follow the same one simple model, and different instructors do very different things – different ways of allocating submissions to reviewers, different ways of allowing tutors to mark submissions, different means of restricting access to reviews, different timings, different anonymity parameters, and different length rubrics. Some instructors use the flexibility of Aropä to conduct activities that would not typically come under the banner of 'peer review' – for example, writing reviews on behalf of errant students who have failed to participate, or requiring the whole class to review a single submission.
- An increasing numbers of instructors are willing to try peer-review. We speculate that this is due to increasing student numbers (and corresponding decrease in staff:student ratios), and the increased emphasis on 'assessment and feedback' in response to poor student evaluation of this facet of their education (Pitt and Norton, 2016). More positively, successful peer-review experiences are increasingly being shared amongst instructors, encouraging others to give it a try, and the introduction of teaching-only contracts in

universities means that more instructors are being encouraged to engage in academic innovation and scholarship activities.

- Instructors either like using peer-review in their classes (and continue using it) or don't. We think that this is more to do with the nature of the students in the class or the relationship between instructor and students than with the peer review activity itself. Students are often resistant at first (especially in the first cohort to use peer review), and the instructor's perseverance in the face of dissatisfaction is an important factor affecting continuance.

- Students mostly dislike it, for a variety of reasons. They want 'expert' feedback, they consider writing reviews extra unnecessary work, and they consider that their own efforts are not adequately rewarded. They are, however, very glad to have the opportunity to see other students' work. Key to this dissatisfaction is, we believe, a tendency for students to underestimate the importance of generic transferrable skills: they seldom recognise that there is personal benefit to themselves (in terms of skill development) in critiquing someone else's work, and in analysing a range of (possibly contradictory) opinions of their own work.

- It is very hard to persuade students to take part in a peer-review activity if there is no extrinsic incentive. Those assignments where no incentive has been offered are easy for us to identify in the system – typically, less than a third of the class have provided reviews. This lack of engagement is, of course, self-perpetuating: students who have voluntarily provided reviews in an assignment where participation has been low are unlikely to have received comments on their own work (and so are less likely to voluntarily provide reviews the next time they are asked to do so). There needs to be a critical mass of participation for students to think that the reviewing process has been worth their while: few voluntary peer review activities reach this critical mass.

- Technology can only take you so far, and cannot make the important academic decisions. Aropä will support the mechanics of peer review, but this is only a very small part of the activity – the rubric, the allocation method, incentives, and the instructions given to students are all key to success. Instructors who have believed that conducting a peer-review activity simply means using Aropä have quickly discovered the problems in their approach.

Aropä is a significant educational technology endeavour that has endured for nearly a decade, supporting peer-review activities worldwide. We continue to offer the system and support, free, to anyone, anywhere who wishes to use it, creating new user accounts at the rate of around three or four a month. We advise instructors as to its use and appropriate peer review activities, and solicit feedback and make enhancements in response to requests. As an unfunded, long-term and on-going initiative, we are inspired by its impact on students' learning worldwide and by the enthusiasm, commitment and innovative ideas of the instructors who use it.

Acknowledgements

References

1. Finlay, C., Tierney, A., McVey, M., and Gray, J. (2012) Online peer review of scientific writing – Aropä. In 5th Annual University of Glasgow Learning and Teaching Conference. Available at: http://eprints.gla.ac.uk/91940/ [Accessed 01/09/16].
2. Hager, P., & Holland, S. (Eds.). (2007). Graduate attributes, learning and employability, Introduction (Vol. 6). Springer Science & Business Media, pp1-15.
3. Hamer, J., Kell C., and Spence, F. (2007) Peer assessment using Aropä. Proceedings of the Ninth Australasian Conference on Computing Education. Australian Computer Society, vol 66, pp43-54.
4. Hamer J., et al (2008) Contributing student pedagogy. ACM SIGCSE Bulletin, 40:4, 194-212.
5. Hamer, J., Purchase, H.C., Denny, P., and Luxton-Reilly, A. (2009) Quality of peer assessment in CS1. 5th International Workshop on Computing Education Research. ACM, pp. 27-36.
6. Hamer, J., Purchase, H.C., Luxton-Reilly, A. and Sheard, J. (2010) Tools for "contributing student learning" Proceedings of the 2010 ITiCSE Working Group Reports. ACM, pp1-14.
7. Huston, E., McVey, M., McCallum, S., Deeley, S., Hau, L. and Finlay, C. (2015) Diverse use of the web-based peer assessment tool, Aropä. Scottish QAA Enhancement Themes Conference. Available at: http://www.enhancementthemes.ac.uk/pages/searchresultdetails/docs/case-studies/diverse-use-of-the-web-based-peer-assessment-tool-aropa
8. Lunt, T. and Curran, J. (2010) Are you listening please? The advantages of electronic audio feedback compared to written feedback. Assessment & Evaluation in Higher Education, 35(7), pp759-769.
9. MacDonald, D. (2013) Online peer assessment in undergraduate physiotherapy education. Scottish QAA Enhancement Themes conference. Available at: http://www.enhancementthemes.ac.uk/pages/docdetail/docs/paper/online-peer-assessment-in-undergraduate-physiotherapy-education [Accessed 01/09/16].
10. Nicol, D. et al. (2014) Rethinking feedback practices in higher education. Assessment & Evaluation in Higher Education, 39(1), pp102-122.
11. November, N. (2011) Literacy loops and online groups: promoting writing skills in large undergraduate music classes, Journal of Music History Pedagogy, 2(1), pp5-23.
12. November, N. (2012) Using undergraduates' digital literacy skills to improve their discipline-specific writing: a dialogue. International Journal for the Scholarship of Teaching & Learning, (6)2, article 5.
13. Nulty, D. (2011) Peer and self-assessment in the first year of university. Assessment & Evaluation in Higher Education, (36)5, pp493–507.
14. Patterson, D. (2009) Two peer review systems for summative assessment in a Green Processing engineering course. Conference on Engineering Our Future: Are We up to the Challenge? Engineers Australia, pp2281-2295.
15. Pitt, E. and Norton, L. (2016). Now that's the feedback I want!' Students' reactions to feedback on graded work and what they do with it. Assessment & Evaluation in Higher Education, Published online, 8 Feb 2016.
16. Ramm, D., Thomson, A., and Jackson, A. (2015) Learning clinical skills in the simulation suite. Nurse Education Today, (36)6, pp823–827.

17. Walker, J. and Sampson, V. (2013) Argument-driven inquiry. Journal of Chemical Education, (90)10, pp1269-1274.
18. Welsh, M. (2014) The impact of a new interactive anatomy curriculum for nursing students. Higher Education Academy STEM Conference. 2014.
19. Whittaker, T. (2016) Personal communication.

The Authors

Dr Helen Purchase started her interest in e-learning with a PhD in Intelligent Tutoring Systems, and has several publications relating to the use of technology in education. As well as co-creator of Aropä, she leads the GUSTTO project – an online system for practising academics to share 'Teaching Tips' with their colleagues.

Dr John Hamer received his PhD in Computer Science from the University of Auckland in 1990, where he served as Senior Lecturer until 2010. He initially created Aropä to support his own teaching. He currently works as a commercial software developer, and has an honorary affiliation with the University of Glasgow.