Online self and peer assessment in large, multi-campus, multi-cohort contexts

Jan Fermelis
Faculty of Business and Law, Deakin University

Richard Tucker
Faculty of Science and Technology, Deakin University

Stuart Palmer
Institute of Teaching and Learning, Deakin University

The increasing use of team assignments within higher education is well documented. The driving forces behind this include desires to facilitate reflective and collaborative learning, to develop generic teamwork skills for graduate employment and to reduce the grading workloads of faculty staff. Students however consistently report dissatisfaction when the assessment of team assignments produces a common grade for all team members. Self-and-peer-assessment (SAPA) is presented as a fair, valid and reliable method of producing information about ongoing team processes. This information can provide ongoing feedback to team members and rich formative data to instructors attempting to assess the team process and students' teamwork skills. This data can also enable individualised summative assessment in dysfunctional teams or situations of uneven team member contributions. Whilst manual SAPA protocols can work effectively for smaller classes, computer-assisted SAPA offers a solution to the problems of large classes. This paper reports on the early stages of an online SAPA tool, originally developed for small classes of architecture students, adapted for use by very large business communication classes comprising up to 1000 students in a semester. This large unit is delivered on four Australian campuses as well as off-campus and in off-shore mode, by up to fourteen instructors at any one time. The paper documents how three researchers from very different backgrounds worked to create their own research team, implement a pilot study, and adapt the online tool, whilst adhering to comparability of assessment constraints and maintaining integrity of research design.

Keywords: team work, group work, self and peer assessment, SAPA, assessment tools

Introduction

While there is a range of online SAPA systems in existence or under development in Australia and internationally, many embody a specific SAPA method or philosophy and would require additional coding to be used in alternative applications. The 2006 Australasian Society for Computers in Learning in Tertiary Education conference included presentations on two such on-line SAPA systems being developed in Australia. Team Contribution Tracking (TeCTra) had been trialled in the Faculty of IT at University of Technology Sydney and allowed peer assessment and review comprising quantitative and qualitative data (Litchfield & Raban 2006). Self and Peer Assessment Resource Kit (SPARK) had also been trialled at the University of Sydney (Freeman et al. 2006). At this point in time it is unclear whether either of these systems or any other system developed in an academic context is freely available for academic use elsewhere, and most appear to be based on open-source database platforms. These are good for minimising software costs but are currently incompatible with the centrally supported Oracle-based on-line environments commonly used within many Australian universities. Whilst it is believed that Blackboard is currently working on development of an online SAPA tool, there is no clear and obvious immediate direction to Australian educators regarding online SAPA systems.

This paper discusses the early stages of a project designed to develop and evaluate an alternative prototype on-line SAPA model in cross-disciplinary and cross-faculty contexts. The original software was developed during a three-year period within Deakin University's School of Architecture and Building, as part of a group-learning research program which focused on the pedagogy of design collaboration in small groups. The current project not only continues the pilot in three different courses within the School of Architecture and Building, Faculty of Science and Technology, but extends it to a business

communication unit within the Bowater School of Management and Marketing, Faculty of Business and Law. The pilot will extend for three semesters and will involve a total of approximately 2,500 students, all of whom are participating in collaborative team assignments.

The project has been labelled the 'Fair Assessment and Effective Reflective Learning', and has the following four specific purposes:

1. To determine the accuracy of the SAPA model by triangulating the quantitative analysis of student marks with qualitative student feedback.
2. To compare the benefits of different reflective components of SAPA across cohorts and disciplines, by exposing some students to qualitative peer feedback, some to quantitative feedback, some to both and some to no peer feedback.
3. To determine the extent of premeditated collusive assessment by students possible in the present model and through this to resolve unfair online peer assessment.
4. To evaluate the practicality and effectiveness of an on-line SAPA model for courses with large, multi-modal, multi-campus, multi-instructor enrolments

Background

The increased use of team work in higher education has been driven by a range factors. These include: i) using peer learning to improve the overall quality of student learning; ii) helping develop specific generic skills sought by employers; and iii) reducing the workload involved in assessing, grading and providing feedback (James, McInnis, & Devlin, 2002). Peer-assessment in team assignments has been shown to enhance in students the motivation for participation (Michaelsen, 1992), to promote independent, reflective, critical learning (Somervell, 1993), and to encourage students to assume increased responsibility for their own learning (Rafiq & Fullerton, 1996). Moreover, online SAPA systems have been be found to solve problems of confidentiality (Freeman & McKenzie, 2002) and improve assessment efficiency (Lin, Liu, & Yuan, 2001) (Freeman & McKenzie, 2002). However, students express several concerns about team assignments within higher education. One is a lack of perceived relevance and overuse of group work; another is the concern that group work may not fairly assess individual contributions (James et al., 2002). In response, Self- and peer-assessment (SAPA) has frequently been advanced as a valid and reliable alternative and/or supplement to teacher-only assessment of individual contributions to group work (Nancy Falchikov & Goldfinch, 2000; Sluijsmans, Dochy, & Moerkerke, 1999).

There are many other potential benefits of SAPA. These include promoting effective teamwork (Brown, 1995); developing professional skills in self-reflection on behaviour (Sluijsmans et al., 1999); overcoming self-rater problems when self-assessment only is used (Mark Freeman & McKenzie, 2002); developing professional graduate attributes for working in multidisciplinary teams and lifelong learning; shifting the student's role from passive receiver to active participant in learning; making learning objectives and desired performance levels explicit to students (McGourty, Dominick, & Reilly, 1998). Applications of SAPA have been reported in a wide range of academic/discipline contexts (Topping, 1998), including teacher education (Sluijsmans, Brand-Gruwel, & Van Merriënboer, 2002), computer programming (Sitthiworachart & Joy, 2003), architecture (Tucker & Rollo, 2006), medicine (Sullivan, Hitchcock, & Dunnington, 1999) engineering (Brown, 1995; McGourty et al., 1998) and business (Fermelis 2006).

As early as 1998 it was noted that computer-assisted SAPA was an emerging trend (Topping, 1998). Technology was seen as the solution to reducing the problems of volume of feedback and time required for the manual implementation of SAPA with large classes (Ballantyne, Hughes, & Mylonas, 2002, Fermelis 2006). There have been reports of other benefits of online SAPA systems which include confidentiality (Mark Freeman & McKenzie, 2002) and improved efficiency (McGourty et al., 1998).

SAPA pilot in architecture and business communication

The research project introduced here addresses weaknesses identified in the current uniform assessment structures of team assignments that are highly unpopular amongst both instructors and students (Tucker 2007). These weaknesses impair the teaching of effective collaborative learning and team-working skills. In addition, the pilot seeks to implement a pedagogic strategy which can reduce resource-intensive individual assessment in the current context of increasing student to staff ratios.

Previous implementations of online SAPA had produced three broad conclusions: first, that students greatly preferred continuous online peer assessment of an individual's contribution rather than all team members being allocated the same mark; second, that the quality of student work increased when student individual contributions were assessed, instead of all team members receiving a uniform grade; and third that the introduction of a more participatory forum for student-centred assessment, helped facilitate reflective learning between peers.(Tucker 2007). Students appeared simultaneously empowered to develop the diverse interpersonal and conflict management skills necessary for working in teams.

The above findings need to be systematically evaluated, tested in other courses, and the SAPA tool further developed to enable tailoring to the specific pedagogic requirements of other disciplines and other forms of team assignments. The intention of this paper is to discuss the early findings of such a pilot and to suggest indicators and benchmarks yielding best practice methods. The use of the SAPA model within a new and very different pedagogical context revealed flaws but also resilience and flexibility in the face of unexpected hurdles. We shall now briefly consider how the design of our research and use of SAPA were influenced by our funding restrictions and by the learning and assessment policies within our institution.

Assessment transparency, comparability and pedagogical intent

Assessment practices and procedures within higher education should be fair, ensuring consistency, comparability and equity to all students. However, the fair and equitable development of a SAPA model and the testing of its variations in itself presented challenges. For instance, Deakin's internal funding body of teaching and learning research projects requires projects to be completed within one calendar year. This collaborative project had originally intended to study the following four variations of SAPA: i) no peer feedback; ii) quantitative peer feedback only; iii) qualitative feedback only; and (iv) both quantitative and qualitative feedback. These were also to be compared to a control group which had not used SAPA at all. However, the one year restriction reduced the scope of the research design and data collection to variations on the SAPA model across two test cohorts of equivalent demographics, pedagogic challenges and teachers within the 2007 academic year.

It had also been originally envisaged that business communication teams might be given the choice of what type of feedback they preferred. It is arguable that the self-selection of feedback type would not differentiate the underlying process of adjusting the initial team grade by peer assessment of that individual's contribution assessment. However the researchers questioned the fairness of team feedback choices on the grounds that research validity could become impaired due to the biases inherent in self-selection. In recognition of the possible short and long-term advantages to those students who receive peer feedback (Topping, 1998), it was decided not to risk disadvantaging any student by offering different forms of feedback within the same cohort. This decision meant, however, that only two feedback options could be compared within the one academic year, and, moreover, that this would leave no control group. The collaborative solution created by the research team turned this experimental obstacle into an improvement. The research program was extended by one semester beyond the calendar year to create the control group and four comparative cohorts were achieved: (i) SAPA with both quantitative and qualitative feedback: (ii) SAPA with quantitative feedback only; (iii) SAPA with no feedback; and (iv) no SAPA. All cohorts in the first and second semester would complete entry and exit questionnaires prior to and after their use of SAPA. In the first semester of 2008, entry and exit questionnaires would be offered to a control cohort who would not use SAPA. This revised program enabled research integrity, funding restrictions and comparability of assessment to all be satisfied.

The primary focus of assessment within higher education should be to match learning objectives and to facilitate learning. Assessment should also be capable of indicating achievement, maintaining standards and providing certification whilst remaining as transparent as possible. Assessment transparency in this project required students to be provided with information about the pedagogical intent of the model, the research aims of testing that model and a precise explanation of how SAPA calculates individual contribution.

Teamwork skills were already included as a learning objective within the business communication curriculum together with strategies for effective feedback. However these objectives are rarely included within architecture, construction and design courses. Tutorial exercises that developed students' interpersonal and team skills and explored their abilities to assess and comment on the work of others were thus also introduced into the curriculum for all students included in this study. For transparency, detailed presentations were also made to all students on the pedagogical intent of the SAPA model.

Research program

The research program consists of the following five broad phases scheduled to occur in three teaching semesters between April 2007 and June 2008:

1. Implementation of two pilot applications given to 2007 cohorts enrolled in core second-year courses within architecture and in business communication, each testing a model providing students with ongoing quantitative and qualitative peer feedback. This phase also includes the collection and collation of entry and exit questionnaires to collect demographic information as well as canvass students' opinions of teamwork and its assessment prior and subsequent to the implementation of SAPA.
2. Evaluation of the first two pilot applications through a comparison of entry and exit questionnaires.
3. Implementation of a SAPA model that includes quantitative feedback only for the architecture/construction management course and no feedback at all in the business communication course. All courses in this third phase are pedagogically equivalent to those from the first phase and phase will again include the collection and collation of entry and exit questionnaires.
4. Evaluation of the second two pilot applications. This fourth phase is equivalent to the second, but involves cohorts of different students.
5. Analysis of entry and exit questionnaires given to the 2008 cohorts enrolled in the same second year architecture and business communication courses that were the focus of the first phase. Students participating in this final phase will act as control groups by making no use of SAPA and reverting to the non-individualised assessment strategy of allocating the same or almost the same mark to all team members.

Nature of the team architecture and business communication assignments

Two architecture and building courses are trialling the SAPA model in 2007. The first is a second-year level design unit studied by 120 architecture and dual degree architecture/construction management students. The design unit comprises two individual assignments and one team assignment, worth 37% of the course marks, which requires teams of three students to collaboratively design small-scale dwellings in one of three sites remotely located (without grid electricity or water) in one of three Australian climatic zones. The second course is a second-year level building environmental studies unit studied by 170 architecture and dual degree architecture/construction management students. Their semester-long team assignment is worth 50% of the course marks and requires teams of five students to assess the environmental performance of a house designed by an eminent local architect. After assessing the house, the teams must redesign it for greater energy and resource efficiency and sustainability. At the end of the semester, the teams present their houses to the original architect.

Business Communication is a compulsory one-semester course at second year level in a three year Bachelor of Commerce award. It is offered twice each year, and is studied by 1800 students on three Victorian campuses, in off-campus mode and at partnership campuses in Singapore and Malaysia. Up to fourteen different members of staff are involved in course delivery at any one time, with uniform teaching materials and a strict comparability of assessment protocol. Approximately 60% of the cohort comprises full fee-paying, international students, primarily from South East Asia, China and the Sub-continent. The aim of the unit is to broaden students' understanding of the complexity of communication, their awareness of the skills and strategies required for effective communication within a range of contemporary professional business contexts, and to provide opportunities to develop those skills within businesslike contexts. For the team assignment, students are required to work individually, collaboratively and co-operatively in teams of four over a period of six weeks and to communicate regularly, face to face and/or electronically whilst researching and reporting on a chosen scenario. Teams are responsible for planning and allocating component tasks and for combining their research. The students present a formal team oral report via three-minute individual presentation (10%), which are assessed individually but with a uniform teamwork component. They are then required to produce a formal written report of 4000 words (20%). This document receives a team score, which can then be individualised. Students are also required to maintain an individual reflective journal during this period (5%), analysing their team experiences in light of what they have learned from their reading, lectures and tutorial exercises on three related course topics – Interpersonal Communication, Teamwork and Decision-making and Intercultural Communication.

Self and peer assessment in the team assignments

The SAPA online tool discussed in this paper requires students to rate and comment on themselves and each other on a weekly basis. As with the TeCTra online SAPA system, (Litchfield & Raban 2007, Raban & Litchfield, 2007), this SAPA tool aims to create a formative, diagnostic and summative assessment environment. In this environment, students are encouraged to develop peer-assessment and feedback skills by making quantitative ratings and qualitative comments and to become cognizant of how their own work is being perceived and assessed by their peers. The online tool collects rich quantitative and qualitative information which can be made available to students and instructors during the team assignments and also frees instructors from the considerable time necessary for processing similar paper-based individualisation tools (Fermelis 2006).

Individualisation of team scores

Within the SAPA tool, students are asked to rate contributions to the assignment made by themselves and by each member of their team. Students are informed that the purpose of this is to assist their teachers to identify teams whose members contributed unevenly, so as to more appropriately individualise students' scores. When awarding scores and ratings, students are asked to consider many indicators of contributions such as whether or not they attended team meetings and tutorials, actively communicated with team-mates, responded to others' messages, participated in decision-making, completed work they offered to do or were designated, met agreed deadlines and shared the workload. In addition they could consider the quality and quantity of contributions to the assignment report, drawings and/or models. Different contributions ebb and flow through the life of any student team, reaching a climax in the lead up to submission due dates, so it was considered appropriate to suggest indicators of positive team member behaviour to assist students in making fair ratings and comments.

Each assignment then receives a team mark following assessment and moderation by instructors and course co-ordinators, which can then be individualised if there is evidence of unequal or uneven contributions by different team members. This decision is prompted primarily on the basis of quantitative SAPA rating scores and statements of task allocation within assignment documentation. Whilst other qualitative information, such as SAPA peer comments, student feedback and tutor feedback may be used for triangulation, the purpose of SAPA is to both alert the instructor to uneven teams and to also save time.

Making SAPA entries

In all units piloting the SAPA model, students were required to make five weekly assessment entries between the beginning and the submission of their six-week team assignments, with entries themselves regarded as an indication of active participation in the assignment. Architecture students who made at least four assessments received their complete score after it had been individualised, but those who made less than four assessments had 2% of the team mark deducted from their individualised score for each missed peer assessment. The size and multi-campus, multi-cohort nature of the business communication students introduced several unexpected complexities including browser problems, computer incompatibility and slight variations in instructor instructions, making it clearly unfair to directly penalise any student for having missed entries. However, number of entries made was used as an indicator of participation and engagement in the assignment process. The SAPA tool was hosted within the university on-line study portal, and populated with student enrolment details, before students were able to log on and change their default passwords. Students had a window of four days (around the weekend) to make their assessment and were able to change their entries at any time before each time window expired, although some of the complications mentioned above meant that these time windows had to be extended on three occasions.

Within SAPA, students were asked to make three different assessments. The first was to award themselves and each other team member a relative contribution score of between 0.5 and 1.5. The clear intent of this was to encourage students to consider the question of quantitative relative contribution or equitable team workload distribution. This score needed to add up to the total number of members in the team. If a student believed that all team members contributed evenly, they should award each member a rating of 1. Or, if they believed that one member had contributed more than the other three, they might award that student 1.3 and the other three team members 0.9 each. This first measure was complemented by a second that asked students to rate the individual performance of each other, using a drop-down menu, on a five point, multiple-response Likert scale evaluation. Likert evaluations are commonly used to rate aspects group experiences (Ellis & Hafner 2005), and allow for the coding of responses and the

subsequent statistical analysis of possible patterns of bias in student assessments. This Likert evaluation aimed to encourage students to consider the quality of each other's contribution; it was translated into a numeric evaluation that could be used in combination with the quantitative relative contribution assessment to arrive at a holistic evaluation of each member's contribution. The combination of these two modes of peer assessments also avoided peer over-marking, which is a problem identified within many peer assessment methods (Falchikov 1986; Freeman & McKenzie 2000). The third SAPA measure asked students to offer comments on the performance of their peers. This was included firstly to elucidate for unit co-ordinators anomalies or unexpected final evaluations; secondly, to develop in students the evaluation, feedback and reflective skills that are key learning objectives of teamwork projects; and thirdly to create an opportunity for underperforming team members to become motivated to self-improve. It was also thought that students who completed the qualitative comments section, even if they themselves did not themselves receive constructive or informative comments, might incidentally become motivated to improve their performance (Dominick et al. 1997). During this first phase of the SAPA project, all scores and comments became visible to each member of the team after expiration of the entry window. However, all scores and comments were made anonymous by randomising the order in which they appeared.

Multiplier scores

At the end of each weekly assessment, and at the conclusion of the team assignment, an assessment matrix was generated for each team. For each student a multiplier of the team grade was calculated. The multipliers could be used to individualise marks. Before the calculation was made, all self-assessment marks were removed from the matrix to negate the bias of possible self over-rating, although the investigators are keen to analyse future data from the pilot to either confirm or reject this hypothesis of self over-rating. The Peer Assessment Multiplier (PAM) is calculated as follows:

1. Group Mean Peer Assessment (GMPA) is total of all students' Individual Contribution Marks[1] AND Individual Performance Marks[2] is divided by the number of students in a group.
2. The Individual Total Peer Assessment (ITPA) is the sum of each individual student's Individual Contribution Marks and Individual Performance Marks.
3. The Peer Assessment Multiplier (PAM) is equal to ITPA divided by GMPA
   Where:
   [1] Individual Contribution Marks are between 0.5 and 1.5.
   [2] Individual Performance Marks are between1 and 5.

The weekly and final scores received by a student therefore indicate how their peers' ratings compare them to the team average. If their rating is less than 1 they are considered to be performing below the average team performance; if their rating is greater than 1, they are considered to be performing above the average. Students are therefore advised to aim for a multiplier score of close to 1 or greater. The range of multiplier scores within any team functions as an indicator of that team's level of evenness, in that a team with similar or closely ranged multiplier scores is deemed to be even and functional but a team with widely varying scores suggests unevenness and dysfunctionality. Previous pilot trials of a prototype of the SAPA model have indicated that SAPA ratings of between 0.8 and 1.2 could be the norm (Tucker, in print).

Entry and exit questionnaires

Students were asked to complete a voluntary two page entry questionnaire at the beginning of the first phase of the study, and a similar two-page exit questionnaire after completion of the team assignment (Cheng & Warren 1997). These questionnaires were a synthesis of questionnaires used in earlier studies (Ballentine; Hughes & Mylonas 2002; Cheng & Warren 1997 (which in turn is based on Burnett & Cavaye 1980); Davies 2000; Lejk & Wyvill 2002; Sivan 2000; and Walker 2001). The two questionnaires were identical apart from use of the present tense in the entry and past tense in the exit. The entry questionnaire began with a section to elicit student demographics, followed by eight questions on students' attitudes towards team-work and a third section questioned attitudes towards peer assessment. All responses to the second and third sections were provided by a 5-point Likert scale with 1 for 'strongly agree' through to 5 for 'strongly disagree'. The investigators decided to expand the scope of the questionnaires to allow students to add written comments after each section on what they liked the most and least about team assignment work and its assessment, and what changes they would like to see made to each. This was to encourage comments to be related to how students were feeling at the time that they made their rating, rather than at some later time when they might revert to stereotypical replies. The

questionnaires were labelled with each student's personal student number to enable the student's entry and exit responses to be paired together.

Student participants

In phase 1 of the study, the architecture cohort consisted of 37 groups of three, one group of two and one group of four. The architecture teams were self-selected by students, with their choice of team-mates restricted to pools of twenty-five students to discourage the option of working with friends. Such restrictions have been shown to encourage diversity within design teams that results in a more challenging learning environment (Tucker & Reynolds, 2006), because heterogeneous contexts expose the learner to multiple perspectives based on the diverse backgrounds and experiences of the other members of the team.

The business and law cohort consisted of 185 self-selected teams, formed after a series of tutorial task exercises designed to heighten students' awareness of the team member qualities required for effective team assignments and to introduce students to each other. Students were encouraged to form heterogeneous teams, but the numbers of students, tutors and classes involved made it impossible to prevent friends from forming teams. The vast majority of these teams had four members.

Preliminary results

At the time of writing, results are restricted to the software refinements introduced to improve management of the large cohort and an overview of the SAPA ratings for the students in phase 1 of our project.

Software and procedural refinements

In addition to cosmetic changes introduced during the first phase of the pilot, many helpful refinements to the SAPA tool were introduced prior, during and after this phase. They are listed below.

1. Student data entered by username and identification number as a default
2. Instructor coded student teams coded by campus and class number and team name for easier access and analysis of data
3. Instructor entered student team members by keying in individual student usernames
4. Software prompt introduced to confirm student usernames to prevent keying in errors
5. Students prompted to change/create own password upon registration
6. Model team created for student demonstration purposes
7. Students received a screen message indicating successful entry of ratings and comments
8. Students permitted to edit ratings and comments within each entry period
9. Students were encouraged to regard their final entry as a global one, which could then account for half of the multiplier score
10. Instructors able to quickly download complete listing of all student teams and range of multiplier scores
11. Instructor able to download information on any team including final multiplier score and randomised student comments
12. Instructor able to quickly download complete information on any team including individualised rating scores and comments by and about any team member

It became clear that important but more sophisticated refinements are still required, and the investigators have already introduced two of these before the commencement of phase two:

13. Instructor access to reset entry time periods
14. Instructor access to add or delete student names from any team at any time

Procedural improvements were also made during phase one, generally in recognition of the pragmatic difficulties of implementing the pilot with such a large cohort is such a complex educational context. As is undoubtedly the case with many pedagogic innovations, implementing a pilot such as this requires the full understanding and co-operation of all teaching staff in order to enthuse and motivate students. Although all students were provided with paper documentation explaining SAPA, e-copies of the instructions and questionnaires were also posted within the university's Blackboard online learning environment. In the case of some off-campus students located overseas and in remote parts of the county, this was a more reliable method of dissemination than the hard copy documents they were sent by

conventional mail. A SAPA Discussion folder was also established to provide a forum for students to ask and answer questions about the software and for the instructor to identify and monitor unexpected problems. In some cases it proved a more prudent though time consuming exercise to telephone students to resolve difficulties, rather than to respond online. It was also found helpful to email students weekly reminders to make their SAPA entries, and secondly to include the weblink within those email messages. Finally it was considered helpful to provide students with some information about the early findings and to thank them for their participation in the pilot. This included an indication of the number of students and entries made, the number of teams and individuals whose scores were individualised and an overview of what had been learnt as a result of their efforts.

Overview of SAPA ratings

All the first-semester architecture students made at least one entry using our on-line SAPA. In total, the cohort made 1406 entries and missed 359 entries. On the professional judgement of the instructors, it was decided that their marks would only be individualised if the range of SAPA ratings in a group was greater than 0.15 (such that the lowest rating subtracted from the highest rating was greater than 0.15). Thus, 21 out of 39 teams had their marks individualised, or 64 out of 117 students. The average range for the cohort was 0.275, and the highest range of ratings for a team was 0.688 (from 1.295 to 0.607). For this team, the tutor mark was 63% and the individualised marks therefore became 38%, 69% and 82%. Such a range might not be acceptable for some assessors, and this highlights a possible danger of implementing such a SAPA model. As Sharp notes, how great the numerical differences are in the ratings that students use to reflect unequal contributions will vary between and within groups (2006). Thus it may be necessary to multiply the SAPA rating by a value that can vary from group to group so that the range of individual marks is satisfactory (Sharp 2006). Our SAPA tool had hoped to minimise such ratings differences by restricting the contributions ratings to between 0.5 and 1.5, but in retrospect it may be that this range is still too wide.

For the 185 teams of business communication students, 72 produced SAPA ratings where the multiplier score range was greater than 0.3. This range was wider than for the architecture students, but was deemed comparable because the final entry rating by business communication students was not identified as a global one and not included as half of the multiplier score. For on-campus students, these constituted 35% of teams, but for the off-campus students the figure was 59%. For each of these teams, SAPA separate ratings and comments were then examined in more detail and individualisation of the team scores then completed by manually referring to a combination of SAPA scores, comments, information relating to which students had completed which specific tasks within the Written Report document, assessor observations and tutor feedback provided about problem teams. Of the 72 uneven teams identified earlier, a total of 29 teams had their results individualised, with score increases and decreases of up to 20%. Many more teams were strongly suspected as being less than even and their team members of deserving individualised scores. However inconsistency of the data and time constraints necessitated a conservative approach. One off-campus student took the time to state that she felt the approach might have been too conservative, so the investigators plan to examine the multiplier formula closely and introduce possible refinements. This will decrease inconsistencies which had hitherto not emerged with the architecture students and produce a more robust and sophisticated SAPA multiplier for phase two.

The potential for self over-rating is mentioned above. The potential for collusion is another SAPA contaminant which is not to be ignored and which had previously been suspected by the instructor whilst manually individualising business communication team scores (Fermelis 2006). Within one team in the current architecture cohort, the ratings of two students were very closely matched, and markedly different from the ratings of the third student in their team. The two students were known to live together and the instructor strongly suspected that they had colluded with each other on their ratings to reduce the score of their third team member. After investigation by the unit chair, the two agreed that perhaps their entries had exaggerated the differences in contribution and agreed to an adjustment. This once again demonstrates the importance of the instructor carefully monitoring throughout an assignment what the SAPA is showing, and the need for triangulation and qualitative comments.

Implications for further research

When analysing the feedback and results of the further stages of our research, the investigators plan to statistically analyse the questionnaire data. The variables that will be considered for both assessors and assessees will include familiarity and experience in peer assessment, geographical and/or cultural origin, chronological age, year of study, academic ability and gender (Topping, 1998). In addition to this, our research will also analyse the students' experiences with SAPA particularly in relation to the different

feedback options that our model will compare, and in consideration of the variable of team size and student perceptions of their team work. Even at this early stage of the project, it is clear from anecdotal evidence and the uncritical nature of comments made by students in their assessments and to their instructors, that even if qualitative feedback comments are anonymous, on campus students in small teams are unwilling to openly criticise peers who might easily deduce which team-mates made which comments. Off-campus students are likely to be more mature, more focused on the purposes and nature of their studies and less tolerant of under-contributing team members. They are also unlikely to ever meet face to face, and may therefore be less fearful of being frank in their SAPA comments.

Conclusion

Despite the size of the classes, marks were available to architecture and building students on the day after students made their final SAPA rating. This was far earlier than in previous years because accessing peer assessment factors via the on-line software tool took substantially less time than the collation and calculation of individual paper-based assessments by unit chairs. This meant that students were able to access both their team marks and their individualised marks for the team assignments only a few days after the assignment submission. For the business communication students in the current study, results of the team assignment were released much less quickly, due to the larger number of teams and more manual adjustments required and to unexpected staff absences. It is encouraging that the acceptance of individualised team assignment scores was almost universal within the two experimental cohorts. Not one student made a complaint about the individualisation of their mark within architecture. Even within the business communication students, only two teams registered any complaints, which were quickly resolved. Increasing academic workloads within higher education leave little time for the administrations of either manual paper-based or more elaborate self and peer assessment methods. The ease both of students making online SAPA entries during their team assignments and of instructors quickly able to access online ratings and individualise student results where required, are more important than ever.

The findings of our project to date thus support the positive contribution of on-line self-and-peer-assessment within student teamwork assignments. Instructors observed an improvement in student satisfaction and class spirit during group assignments using the on-line SAPA model in comparison from that experienced in equivalent past assignments which used no or more rudimentary forms of peer assessment. Increased maturity and confidence in many students as the assignments progressed was also apparent. Numerous students reported that SAPA provides a ‘pressure valve’ throughout the team projects which allowed teams to function harmoniously despite unequal levels of skill and contributions. The SAPA model can thus be seen to have allowed students to be tolerant of differential learning and assessment aspirations of their peers. Through this toleration, and other mechanisms that will only become clear once entry and exit questionnaires have been analysed, our on-line SAPA tool seems to have changed for the better the group dynamics seen in teams collaborating in the architecture and business communication courses under study.

Acknowledgement

The authors would like to acknowledge the input of Catherine Reynolds of the School of Behavioural Science at Melbourne University who has been a meticulous and diligent research assistant throughout this pilot project.

References

Ballantyne, R., Hughes, K., & Mylonas, A. (2002). Developing Procedures for Implementing Peer Assessment in Large Classes Using an Action Research Process. *Assessment & Evaluation in Higher Education*, 27(5), pp. 427-441.

Brown, R. W. (1995, 1-4 November). *Autorating: Getting Individual Marks from Team Marks and Enhancing Teamwork*. Paper presented at the 25th Annual Frontiers in Education Conference, Atlanta, Georgia.

Burnett, W., & Cavay, G. (1980). Peer assessment by fifth year students of surgery. *Assessment in Higher Education*, 5, pp. 273-278.

Cheng, W., & Warren, M. (1997). Having Second Thoughts: student perceptions before and after a peer assessment exercise. *Studies in Higher Education*, 22(2), pp. 233-239.

Davies, P. (2000). Computerized Peer Assessment. *Innovations in Education and Teaching International*, 37(4), pp. 346-355.

Dominick, P. G., Reilly, R. R., & McGourty, J. (1997). *Incorporating student peer review and feedback into the assessment process*. Paper presented at the Best Assessment Processes in Engineering Education: A Working Symposium, Terre Haute, Indiana.
Ellis, T. J., & Hafner, W. (2005). *Peer Evaluations of Collaborative Learning Experiences Conveyed Through an Asynchronous Learning Network*. Paper presented at the Conference Proceedings Hawaii International Conference on System Sciences (HICSS-38), Big Island, Hawaii.
Falchikov, N. (1986). Product Comparisons and Process Benefits of Collaborative Peer Group and Self-assessment. *Assessment & Evaluation in Higher Education*, 11(2), pp. 146-166.
Falchikov, N., & Goldfinch, J. (2000). Student Peer Assessment in Higher Education: A Meta-Analysis Comparing Peer and Teacher Marks. *Review of Educational Research*, 70(3), pp. 287-322.
Fermelis, J. (2006) *Making Teamwork work*. Proceedings of the 71st Association for Business Communication Annual Convention, San Antonio, Texas, USA.
Freeman, M., & McKenzie, J. (2000, 2-5 July). *Self and Peer Assessment of Student Teamwork: Designing, implementing and evaluating SPARK, a confidential, web based system*. Paper presented at the ASET-HERDSA 2000 Conference, Toowoomba, Qld.
Freeman, M., & McKenzie, J. (2002). 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. 551-569.
Freeman, M. Hutchinson, D., Trevealen, L. & Sykes, C. (2006), *Iterative learning: Self and peer assessment of group work*, Proceedings of the 23rd Annual ASCILITE Conference: Who's Learning? Whose Technology?, pp. 257-266.
James, R., McInnis, C., & Devlin, M. (2002). *Assessing Learning in Australian Universities*. Melbourne: Centre for the Study of Higher Education and The Australian Universities Teaching Committee.
Lejk, M., & Wyvill, M. (2002). Peer Assessment of Contributions to a Group Project: student attitudes to holistic and category-based approaches. *Assessment & Evaluation in Higher Education*, 27(6), pp. 569-577.
Lin, S. S. J., Liu, E. Z-F., & Yuan, S. M. (2001). Web-based peer assessment: feedback for various thinking style students. *Journal of Computer Assisted Learning*, 17, pp. 420-432.
Litchfield, A. & Raban, R. (2006), *Team Contribution Tracking System (TecTra) for assessment of individual contributions in groupwork.* Proceedings of the 23rd Annual ASCILITE Conference: Who's Learning? Whose Technology?, p. 984.
McGourty, J., Dominick, P., & Reilly, R. R. (1998, 4-7 November). *Incorporating Student Peer Review and Feedback into the Assessment Process*. Paper presented at the 28th Annual Frontiers in Education Conference, Tempe, Arizona.
Michaelsen, L. K. (1992). Team learning: a comprehensive approach for harnessing the power of small groups in higher education. *To Improve the Academy*, 11, pp. 107-122.
Raban, R., & Litchfield, A. (2007). Supporting peer assessment of individual contributions in groupwork. *Australasian Journal of Educational Technology*, 23(1), pp. 34-47.
Rafiq, Y., & Fullerton, H. (1996). Peer assessment of group projects in civil engineering. *Assessment & Evaluation in Higher Education*, 21, pp. 69-81.
Sharp, S. (2006). Deriving individual student marks from a tutor's assessment of group work. *Assessment & Evaluation in Higher Education*, 31(3), pp. 329-343.
Sitthiworachart, J., & Joy, M. (2003, 9-11 July). *Web-based Peer Assessment in Learning Computer Programming*. Paper presented at the 3rd IEEE International Conference in Advanced Learning Technologies, Athens, Greece.
Sivan, A. (2000). The Implementation of Peer Assessment: an action research approach. *Assessment in Education*, 7(2), pp.193-213.
Sluijsmans, D. M. A., Brand-Gruwel, S., & Van Merriënboer, J. J. G. (2002). Peer Assessment Training in Teacher Education: effects on performance and perceptions. *Assessment & Evaluation in Higher Education*, 27(5), pp. 443-454.
Sluijsmans, D. M. A., Dochy, F., & Moerkerke, G. (1999). Creating a Learning Environment by Using Self-, Peer- and Co-Assessment. *Learning Environments Research*, 1(3), pp. 293-319.
Somervell, H. (1993). Issues in assessment, enterprise and higher education: the case for self-, peer and collaborative assessment. *Assessment & Evaluation in Higher Education*, 18, pp. 221-233.
Sullivan, M. E., Hitchcock, M. A., & Dunnington, G. L. (1999). Peer and self assessment during problem-based tutorials. *The American Journal of Surgery*, 177(3), pp. 266-269.
Topping, K. (1998). Peer Assessment between Students in Colleges and Universities. *Review of Educational Research*, 68(3), pp. 249-276.
Tucker, R. (2007). The Impact of Assessment Modes on Collaborative Group Design Projects. In S. Frankland (Ed.), *Enhancing Teaching and Learning through Assessment: Embedded Strategies and their Impacts* (Vol. 2, pp. 72-85). Hong Kong: The Assessment Resource Centre, The Hong Kong Polytechnic University.

Tucker, R., & Reynolds, C. (2006). The Impact of Teaching Models on Collaborative Learning in the Student Design Studio. *Journal for Education in the Built Environment*, 1(2), pp. 39-56.
Tucker, R., & Rollo, J. (2006). Teaching and Learning in Collaborative Group Design Projects. *Journal of Architectural Engineering & Design Management*, 2, pp. 19-30.
Walker, A. (2001). British psychology students' perceptions of group-work and peer. *Psychology Learning and Teaching*, 1(1), pp. 28-36.

Jan Fermelis, Faculty of Business and Law, Deakin University, 221 Burwood Highway Burwood Vic 3125, Australia. Email: jan.fermelis@deakin.edu.au.
Richard Tucker, Faculty of Sciece and Technology, Deakin University, 1 Gheringhap Street, Geelong Vic 3217, Australia. Email: richard.tucker@deakin.edu.au.
Stuart Palmer, Institute of Teaching and Learning, Deakin University, 1 Gheringhap Street, Geelong Vic 3217, Australia. Email: stuart.palmer@deakin.edu.au.

Please cite as: Fermelis, J,. Tucker, R. & Palmer, S. (2007) Online self and peer assessment in large, multi-campus, multi-cohort contexts. In *Providing choices for learners and learning. Proceedings ASCILITE Singapore 2007*. http://www.ascilite.org.au/conferences/singapore07/procs/fermelis.pdf