Self and Peer Assessment: Development of an online tool for team assignments in business communication and architecture

Jan Fermelis, Richard Tucker

Deakin University Australia

Introduction

The increasing use of team assignments within higher education is well documented, fuelled by desires to facilitate reflective and collaborative learning, develop generic teamwork skills for graduate employment and reduce the grading workloads of faculty staff. Self-and-peer-assessment (SAPA) is presented as a fair, valid and reliable method of producing information about ongoing team processes. This information can function as feedback to team members throughout the team project and as data to instructors attempting to assess the team process and teamwork skills and for the purpose of fairly individualising team grades. The manual implementation of SAPA protocols can work effectively for smaller classes, but the volume of data produced can become excessively time consuming for larger classes. For very large classes delivered on multiple campuses by multiple instructors it becomes necessary to turn to a more efficient system. Computer-assisted SAPA offers a solution to these problems.

This paper discusses the early stages of a project evaluating and developing an alternative prototype on-line SAPA model in cross-disciplinary and cross-faculty contexts. The software was originally developed over a three year period at one Australian university for the purpose of more fairly assessing the collaborative team studio design projects of undergraduate architecture and design students. The current project extended the developmental SAPA pilot to include a large undergraduate business communication course undertaken by commerce students. A combined total of nearly 2000 students were involved, all of whom were participating in collaborative team assignments.

The intention of this paper is to discuss the further development of the SAPA software tool as a potential solution to unfair shared assessment of team assignments in a business communication course. The use of the SAPA tool within this new and very different pedagogical context revealed flaws but also resilience, flexibility and promise in the face of unexpected hurdles. The paper also provides reflective insights into the successful implementation of a cross-faculty and cross-disciplinary, collaborative pilot project.

Background

The escalating inclusion of teamwork assignments in higher education has been driven by a range objectives; including i) to use peer learning to improve the overall quality of student learning; ii) to help develop specific generic skills sought by employers; and iii) to reduce the workload involved in assessing, grading and providing feedback (James, McInnis, & Devlin, 2002). In addition to perceptions about a lack of perceived relevance and overuse of group work, a major disquiet for students is that team assignments may not fairly assess individual contributions (James et al., 2002; Li & Campbell 2008). As a response to these concerns,

self- and peer-assessment (SAPA) has been advanced as a valid and reliable alternative and/or supplement to teacher-only assessment of individual contributions to group work (Falchikov & Goldfinch, 2000; Sluijsmans, Dochy, & Moerkerke, 1999).

Peer-assessment has been shown to promote independent, reflective, critical learning (Somervell, 1993), to enhance in students the motivation for participation (Michaelsen, 1992) and to encourage students to take responsibility for their learning (Rafiq & Fullerton, 1996). Moreover, online SAPA systems have been found to solve problems of confidentiality (Freeman & McKenzie, 2002) and improve assessment efficiency (Lin, Liu, & Yuan, 2001) (Freeman & McKenzie, 2002).

Other potential benefits of SAPA have been widely acknowledged. These include: promotion of effective teamwork (Brown, 1995); development of professional skills in self-reflection on behaviour (Sluijsmans et al., 1999); resolution of self-rater problems when self-assessment only is used (Freeman & McKenzie, 2002); development of professional graduate attributes for working in multidisciplinary teams and lifelong learning; shifting of the student's role from passive receiver to active participant in learning; and explication to students of learning objectives and desired performance levels (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).

A major disadvantage already acknowledged with most manual or paper-based SAPA models is that they are extremely time consuming to implement. Any instructor time saved with grading team rather than individual assignments, could consequently be lost in the individualisation process (Fermelis 2006). Computer-assisted SAPA has been identified as an emerging trend (Topping, 1998) which has the potential to reduce the problems associated with the volume of assessment and feedback and the time required for the manual implementation of SAPA, particularly in connection with large classes (Ballantyne, Hughes, & Mylonas, 2002, Fermelis, 2006). Online SAPA systems could also improve efficiency (McGourty et al., 1998) whilst ensuring confidentiality (Freeman & McKenzie, 2002).

Whilst different aspects of team assignments have received considerable attention from business communication instructors, little research has been conducted in relation to SAPA instruments. Useful studies have been published on the impact of social style on peer evaluation in project teams (May & Gueldenzoph 2006) and more recently on the effectiveness of rater training in order to reduce social style bias in peer evaluation (May 2008). This paper attempts to extend the literature on both team assignments and SAPA by documenting the development and implementation of an online SAPA instrument, whilst it was being customised in response to the needs of an undergraduate business communication cohort.

As with any assessment mechanism, SAPA needs to more than simply efficient and confidential. Validity and reliability in relation to the assessment of team assignments have indeed already been recognised as major concerns to business communication instructors (Fermelis 2006; May & Gueldenzoph 2006; May 2008). The systematic implementation, tailored development and evaluation of the SAPA tool in connection with a business communication team assignment would test both the efficiency and the integrity of the tool

for the valid and reliable individualisation of business communication team assignment grades.

The online SAPA tool

The online SAPA tool was originally developed in connection with the assessment of architecture team projects in one Australian university. Previous research into implementation of the tool produced two broad conclusions: that the quality of student work increased when continuous peer assessment was used to assess individual contributions, instead of all team members receiving a uniform grade; and that students greatly preferred continuous online peer assessment of an individual's contribution rather than all team members being allocated the same mark. (Tucker, in press). This prior research also demonstrated that the introduction of a more participatory forum for student-centred assessment facilitated reflective learning and the development of the diverse interpersonal and conflict management skills necessary for working in teams.

The SAPA online tool piloted here required students to rate and comment on themselves and each other on a weekly basis for the purpose of identifying uneven teams and individualising member scores where appropriate. As with other online SAPA systems, (Raban & Litchfield, 2007), a parallel aim was to create a formative, diagnostic and summative assessment environment in which students were encouraged to develop peer-assessing and feedback skills by making quantitative ratings and qualitative comments and to become cognizant of how their own work was being perceived and assessed by their peers. When awarding scores and ratings, students were asked take into consideration actions and behaviours such as whether each member attended meetings and tutorials, actively communicated with team-mates, responded to others' messages, participated in decision-making, completed work they offered or were designated to do, contributed work of the required standard and/or form, met agreed deadlines and shared the workload.

Making SAPA entries

In the pilot, students were required to make five weekly online SAPA entries throughout the respective six-week team assignment periods. Students made their entries by logging on to a password protected website that was accessed via the university on-line study portal. Students were able to change their entries at any time before each time window expired. Architecture students who made at least four assessments received their complete score after it had been individualised. Those who made less than four assessments were penalised by having 2% of the team score deducted from their individualised score for each missed peer assessment. For business communication students it was decided unfair to directly penalise any for having missed entries, because of the procedural challenges which emerged with implementing SAPA with such a large and complex cohort, and unexpected access difficulties for some students. However, the number of entries made was used by the instructor during individualisation as one indicator of team member participation and engagement in the assignment task and process. In effect, students were slightly rewarded for having made their SAPA entries, rather than being penalised for not having done so.

Students were asked to make three different assessments. The first asked them to award each team member a Relative Contribution score of between 0.5 and 1.5 in consideration of the quantitative distribution of fair workload. The total of all Relative Contribution scores was equal to the total number of members in the team. This score was backed up by a second

assessment that asked students to rate the qualitative Individual Performance of their team members' work on a five point Likert scale. Likert evaluations are commonly used to rate aspects group experiences (Ellis & Hafner 2005), allow for the coding of responses and the subsequent statistical analysis of possible patterns of bias in student assessments. The Individual Performance assessment encouraged students to consider the quality, as oppose to the quantity, of each other's contribution. These two scores were combined to produce a Multiplicative Scaling Factor (MSF), or holistic evaluation of each member's contribution. The combination of these two modes of peer assessment acted to avoid peer over-marking, which has been identified as a problem common to many peer assessment methods (Falchikov 1986; Freeman & McKenzie 2000). The third assessment measure asked students to record comments on the performance of their peers. This was included firstly, to elucidate for course co-ordinators anomalies or unexpected final evaluations and, secondly, to develop in students the evaluation, feedback and reflective skills that are key learning objectives of teamwork projects. All MSF scores and comments were made available to each member of the team after expiration of the entry window. However, these were rendered anonymous by randomising the order in which they appeared. It was hoped that students who completed the qualitative feedback section, even if they themselves did not themselves receive constructive or informative comments, might be motivated to improve their performance (Dominick et al. 1997).

At the end of each weekly assessment, and at the conclusion of the team assignment, an assessment matrix was automatically generated for each team. Before the calculation was made, all self-assessment marks were removed from the matrix to negate the possible bias of self over-marking (Zhang, Johnston & Bagic Kilic 2008). The matrix calculated the MSF for each team member which indicated the evenness of respective efforts within the team and which would be used to individualise marks if there was evidence of significant inequity within the team. The MSF was calculated as follows:

1. Individual Total Peer Assessment (ITPA): total of each team member's Relative Contribution (RC*) plus Individual Performance (IP**) scores
2. Team Total Peer Assessment (TTPA): total of all team members' ITPA scores
3. Team Mean Peer Assessment (TMPA): TTPA divided by the number of team members
4. Multiplicative Scaling Factor (MSF) for each student: ITPA divided by TMPA

Where:
* RC scores are restricted to between 0.5 and 1.5.
** IP scores are between 1 and 5.

Implementing the SAPA pilot

Integrating the SAPA pilot into course assessments

Assessment practices and procedures at universities should ensure quality control, fairness, consistency, comparability and equity. However, the development of the SAPA model and the testing of its variations presented challenges to these requirements. The first of these was that internal grant used to fund the pilot required the study to obtain and report results within the one calendar year. The second challenge related to internal comparability of assessment policies within the university which demanded that consistent assessment procedures be implemented within each course cohort within any calendar period.

The collaborative solution created by the research team turned these two experimental obstacles into an improvement. The research program was extended to include one semester beyond the calendar year to create the control group and to thereby achieve three comparative cohorts: (i) SAPA with both quantitative and qualitative feedback: (ii) SAPA with no peer feedback; and (iii) no SAPA. Revising the program enabled research integrity, funding restrictions and comparability of assessment policies to all be satisfied. Phase 3 data was to be collected via online questionnaires, which could generate statistical analyses without any need for a funded research assistant.

The primary focus of assessment should be to encourage, direct and reinforce learning and match the learning objectives associated with any course. Assessment should also be capable of indicating achievement, maintaining standards and providing certification whilst remaining as transparent as possible. As this project was simultaneously investigating students' perceptions of the SAPA tool whilst that tool was being used to help assess students' work, integrity and transparency were critical. Students were thus provided with a precise explanation of how a SAPA tool calculates individual contribution, made aware of the pedagogical intent of the model and were fully informed of the research aims of testing that model. Students were also made aware that whilst it was compulsory for them to use the SAPA tool during the course of their team assignment, that they were free to choose whether or not to complete any research questionnaires in association with the pilot.

Phases of the research program

The total SAPA project involved three broad phases.

In Phase 1, two pilot applications were implemented in core second-year courses within architecture and in business communication, with each testing a peer feedback model providing students with ongoing quantitative and qualitative peer feedback. This phase also included the collection and collation of entry questionnaires that collected demographic information as well as elicited student's perceptions of team work and its assessment prior to their use of the online SAPA application in connection with their team assignment.

In Phase 2, architecture students in pedagogically equivalent courses were required to use SAPA in the same way as in Phase 1. Shortly before the commencement of this phase, the business communication course chair role was unexpectedly assumed by a different instructor, who insisted that SAPA be made optional for students in connection with their team assignment. Nonetheless, for all architecture and business communication students this phase again included the collection and collation of entry and exit questionnaires.

In Phase 3, both courses acted as control groups with students completing their team assignments without SAPA, and instructors reverting to the non-individualised assessment strategy of allocating the same or almost the same mark to all team members. Once again student perceptions of their team experiences will be elicited via entry and exit questionnaires.

Nature of the team assignments

The business communication course involved in this pilot study is a compulsory, one-semester course constituting twenty five percent of a full time load at second year level in a three year Bachelor of Commerce award. Offered twice each year, it is studied by up to 2000

students each year on three campuses within the state of Victoria in Australia, in off-campus mode and at partnership campuses in Singapore and Malaysia. Up to fourteen different members of staff were involved in course delivery at any one time, with uniform teaching materials and a strict comparability of assessment protocol. Approximately 60% of the cohort comprised full fee-paying, international students, primarily from South East Asia, China and the Sub-continent.

For their team assignment, business communication students worked in teams of four over a period of up to six weeks on an integrated oral and written report and reflective journal, worth 35% of their total score for the course. Teams operated with considerable autonomy and were responsible for planning and allocating component tasks and for combining their research. They operated individually, collaboratively and co-operatively to research a chosen scenario and were required to communicate regularly, face to face and/or electronically. The students presented an oral report in a three-minute presentation (10%). Team members were assessed individually (90%) but received a uniform teamwork component (10%). They were then required to produce a formal written report of 4000 words (20%). This document received a team score, which could then be individualised if there were consistent evidence of uneven contributions by different team members. Students were also required to maintain an individual reflective journal during this period (5%). For this task they analysed their team experiences in light of what they had learned from reading, lectures and tutorial exercise.

Within the business communication curriculum, teamwork skills were already included within the curriculum as a learning objective, topic and in the form of strategies for the evaluation of others' work and communication of positive and effective feedback. In addition, the business communication course required students to maintain a journal in which they reflected on their teamwork experiences and skills. However these team-related skills are rarely formally included within architecture, construction and design courses. The teaching of teamwork skills as well as tutorial exercises that explored the students' abilities to assess the work of others were thus also introduced into the architecture and design curricula for those students included in this study in order to prepare them more fully for their SAPA obligations.

The SAPA model was trialled in two architecture and building courses. The first was a second-year level design course studied on one campus by 120 architecture and dual degree architecture/construction management students. The design course included a team assignment, worth 37% of the course marks, which required 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 was a second-year level building environmental studies course studied by 170 architecture and dual degree architecture/construction management students. A five-person team assignment throughout the duration of the course, which is worth 50% of the course marks, required students to assess the environmental performance of a house designed by an eminent local architect. After assessing the house, the teams redesigned it for greater energy and recourse efficiency by using passive heating and cooling techniques and through the careful selection of construction materials. At the end of the semester, the teams presented the house, designed for greater sustainability, to its original architect.

For all courses within the project, each team received a raw team score for their project or assignment document. This score was then to be individualised if there was evidence of unequal contributions by different team members. Individualisation decisions would be made

by the course co-ordinator on the basis of SAPA numerical ratings, freed from the considerable time required to process similar paper-based strategies (Fermelis 2006). Individualisation decisions could supplemented if necessary by SAPA comments, or by any instructor knowledge of a particular team's activities.

Student participants

In Phase 1 of the study, the business communication cohort consisted of 185 self-selected teams. These were 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. Whilst undeniably desirable, instructor formation of teams was deemed excessively problematic in light of the size and complexity of the cohort. The vast majority of these teams had four members. 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, but to discourage the option of working with friends, their choice of team-mates was restricted to pools of twenty-five students. Such restrictions have been shown to encourage diversity within design teams that results in a more challenging learning environment.

Entry and exit questionnaires

All business communication and architecture students in Phase 1 were offered an entry questionnaire at the beginning, and an exit questionnaire after completion of the team assignment (Cheng & Warren 1997). These questionnaires were based on a combination of questionnaires used in earlier studies (Ballentine; Hughes & Mylonas 2002; Cheng & Warren 1997 (which in turn was based on Burnett & Cavaye 1980); Davies 2000; Lejk & Wyvill 2002; Sivan 2000; and Walker 2001). The two-page entry and exit questionnaires were identical apart from use of the present tense in the entry and past tense in the exit. The entry questionnaire elicited student demographics, followed by questions on students' attitudes towards teamwork and towards peer assessment. All responses to these questions were chosen from a 5-point Likert scale. Students were also able to include unstructured written comments to open questions about team assignment work and its assessment. The questionnaires were labelled with each student's personal student number to enable the student's entry and exit responses to be paired together.

In Phase 2, architecture students were also offered both questionnaires, and business communication students exit questionnaires. The absence of entry questionnaires for these students was viewed as unproblematic because the size and nature of the cohort were equivalent to that in Phase 1, and the number and high percentage of completed questionnaires collected in that first phase meant that the data already collected could be considered representative. In Phase 3 when SAPA was not used, there was only a need for a single questionnaire to be administered, and business and architecture students were offered this opportunity online

Data contained in the entry and exit questionnaires from all three phases continue to be analysed, and are to be reported in detail elsewhere (Tucker, Fermelis and Palmer in press).

Preliminary results

The cross-faculty research team

An unanticipated but fascinating dimension of the project was the way in which the experiences of our cross-faculty and cross-disciplinary research team paralleled those of many student assignment teams (Fermelis, Tucker & Palmer 2008). The investigators quickly progressed through the forming and storming stages. The team then proceeded to function with amazing equanimity and productivity during the norming or emergence phase, which rapidly morphed into the fourth performing stage. The researchers quickly raised questions, settled on solutions and achieved consensus on a myriad of procedural matters in an atmosphere of trust and mutual respect. Refinements to the SAPA software were suggested and quickly implemented, comprehensive student instructions created and refined, folders and discussion sites created within the university's Blackboard learning environment, the questionnaires designed and hard copy mail-outs to off-campus and off-shore students completed.

At the time of writing the research team is still performing well, collaborating amicably and productively on analysing the later phases of the pilot and on generating publications for a variety of academic audiences. Only one potentially fatal hiccup has been encountered during this entire pilot project, which created a situation of shared danger which ironically functioned as a bonding stimulus to uninte the research team even more strongly together.

Student compliance with SAPA requirements

In this paper results are restricted to a detailed discussion of the implementation of Phase 1 and an overview of Phase 2.

As indicated above, whilst all students in Phase 1 were instructed to make SAPA entries throughout the life of their team assignment, neither of the two cohorts demonstrated complete compliance with this instruction. Business communication students made 13,453 entries and missed 8338 entries, with 54 students (7%) making no entry at all. In the architecture cohort, all students in Phase 1 made at least one entry using on-line SAPA; students made a total 1406 entries and missed 359 entries. It is relevant to note the relationship between participation rates to the differing reward/penalty contexts created for cohorts in the two phases. In Phase 1, the business communication students were rewarded for compliance and demonstrated 65% completion, whereas the architecture students, who were penalised for not making entries, demonstrated a participation rate of 82%. In Phase 2, students in the second architecture cohort were rewarded for their participation and achieved 57% completion of maximum entries. In this same phase, an unexpected change in the business communication course chair resulted in SAPA participation being decreed voluntary. Some instructors continued to strongly encourage their enrolled students to make SAPA entries, whilst others consistently reminded their students that this was not compulsory and would result in neither reward nor penalty. In Phase 2, a total participation of only 25% by business communication students was achieved.

Student participation rates became a matter of great interest to the researchers throughout the research project. One of the challenges was to maximise the number of SAPA entries made by business communication students, a complex, multi-instructor, multi-campus cohort, in comparison with a far smaller, single instructor architecture cohort. Within Phases 1 and 2, the highest student participation rates existed where SAPA was compulsory, with penalties for students who failed to make entries. Next highest was when students were required to participate and rewarded for doing so. The lowest rates occurred where SAPA was voluntary.

Our data therefore strongly suggests that if SAPA is to be used for the purposes of individualisation of student grades, student entries need to be compulsory, with a penalty built into the assessment matrix formula for student non-compliance.

Individualisation of student grades

The weekly and final scores received by a student each week indicated how their received ratings compared to the team average. If their rating was less than 1 they were considered to be performing more poorly than the average team performance; if their rating was greater than 1, they were considered to be performing better than the average. Previous pilot trials of a prototype of the SAPA model indicated that SAPA rating ranges of between 0.8 and 1.2 were the norm (Tucker, in press). Any greater range of SAPA student scores within any team was interpreted as indicating uneven member contributions, with the magnitude of the range indicating the magnitude of the unevenness. In teams with a MSF rating range greater than an identified amount, the MSF was then applied as a multiplier of the raw team score, to individualise each team member's final score for that assignment task. Effectively, this process redistributed assessment points from underperforming team members to their better performing team-mates.

On the professional judgement of the architecture instructor, it was decided that marks would only be individualised with his course if the range of SAPA ratings in a team was greater than 0.15 (such that the lowest MSF 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 (MSF rating 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 business communication cohort, individualisation of student results proved problematic because the SAPA ratings appeared to reveal far greater inequity. Of the 185 teams, 142 produced a ratings range of 0.15 or less, with one team scoring a range of 1.228 (MSF from 0.584 to 1.812). Because of the compulsory nature of the course, and size and diversity of the cohort, it had been predicted that the business communication cohort might be less compliant with making SAPA entries and the data on number of entries made certainly bore this out. It had also been anticipated that the business communication students could have less positive teamwork experiences than the architecture students, and so the larger range in MSF ratings was expected. However, after closer analysis, the MSF rating scores from some teams appeared contradictory and quite contra-intuitive to the instructor. Relatively low MSF scores were achieved by a number of students who had in fact made frequent and detailed entries, and who appeared to be more actively and positively engaged in the entire assignment process. These scores were at times lower than those of team-mates who had made fewer entries, received less complimentary peer comments and definitely appeared to be less actively engaged. After the instructor had examined the teams' SAPA separate ratings and comments more closely, a pattern finally began to emerge. If a student

failed to make an entry, SAPA assigned a bland default entry. In teams where the numbers of entries made by the four team members were widely disparate, these lower default entry scores effectively penalised the team members who had entered more encouraging scores, and whose self-entries were disregarded.

For the 185 business communication teams, it was therefore deemed necessary for the instructor to analyse the full set of SAPA ratings and student comments for each separate team in great detail in order for their assignment to be individualised fairly. Because time was limited, it was necessary to focus on a manageable number of teams, so the ratings range was increased to 0.3 which resulted in 72 teams being examined. For on-campus students, these constituted 35% of teams, but for the off-campus students the figure was 59%, suggesting that teamwork equity may be more problematic for teams functioning at a distance rather than it is for face to face 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 deserving individualised scores, but the inconsistency of the data, skewed multiplier scores and time constraints necessitated a conservative approach.

Ironically, whilst SAPA did not provide the accuracy and time-savings expected for the business communication cohort, much was learned for the future. In the previous implementation of online SAPA with architecture and building students, no student had made no or few SAPA entries and so there had been no skewing of the MSF. This new and unanticipated finding from Phase 1 of the project was precisely the type of weakness the investigators were hoping would be revealed from implementation with the large, complex and less co-operative business communication cohort. A more robust and sophisticated SAPA MSF multiplier was accordingly devised and tested for use in Phase 2.

Student response to individualised team assignment grades

Business communication students were anecdotally observed to exhibit an improvement in optimism and class spirit during group assignments using the on-line SAPA model, compared with previous semesters when no, or more rudimentary forms of peer assessment, had been used. Increased maturity and confidence in many students as the assignments progressed was also apparent. Numerous students reported that the 'pressure valve' SAPA provided throughout the project allowed groups to function well despite unequal levels of skill and contributions. The SAPA model can thus be seen to have allowed students to better tolerate the fact that their peers might not have the same learning and assessment aspirations, as they realised that they were more likely to receive the grades they deserved in their team assignments without having to cover for team mates.

Student acceptance of their individualised results was even more positive than had been anticipated. Perhaps tellingly, not one student made a complaint about the individualisation of their mark within architecture, and even within the business communication students, only two teams registered any complaints. The first of these was quickly resolved once the instructor had reminded the student of the consistently poor ratings and comments made by his team mates. The second complaint was from an off-campus student who felt that her team's individualisation had been too conservative. Hence the acceptance of individualised team assignment scores was almost universal within the two experimental cohorts.

Implications for instructors

For instructors committed to the notion of individualised assessment for team assignments, the SAPA tool is conceptually a highly attractive instrument. It enables instructors to take into account student experiences and perceptions of their team processes during assessment in addition to the team assignment product. Whether it be a business communication report or an architectural model, this possibility introduces far greater depth and validity to student grades as well as enhancing the student learning experience. SAPA reduces many of the disadvantages of team assignments, including unfair common grades, undetected social loafers, underperforming team members, stress and inequitable workloads for more diligent team members. This is a tool which facilitates learner reflections and the development of teamwork skills in addition to providing a numerical score for use by the instructor in individualising student scores, as well as considerable quantitative and qualitative evidence to support shared or individualised grades.

However, a common aim of team assignments, as mentioned above, is to help reduce instructor workloads. The pilot raises very real questions about whether SAPA saves time for the instructor, when one factors in the considerable amount of time and effort required to motivate and train other instructors, induct students, construct the online team sites, monitor student compliance and deal with other matters, both expected and unexpected. The issue of time is of particular concern in the context of a large, complex course, multiple campuses, multiple mode, multiple instructor courses. However, it is important to remember that an overt aim of the pilot had been to trial, adapt and further develop the SAPA tool. This process was at times both painful and extremely time-consuming, although the improvements made would of course produce time-savings in future implementations.

Additional refinements to the tool are both needed and possible, and would further reduce the instructor workload. Students could create and populate their own online team sites. Increased student compliance with making rater entries could be facilitated if online blocking mechanisms could deny other avenues of student online access until after their SAPA ratings had been entered. Compliance and training would also become less problematic if university policies integrated the SAPA tool into all team assignments in all courses. The SAPA process would then quickly become both familiar and routine. Training would only need to be completed once, and trained students would operate as trainers for their peers.

Issues for further investigation

When analysing the questionnaire data and results of the further stages of our research, 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, ability and gender, as suggested by Topping (1998). Later analyses will also examine the response to SAPA relative to the different feedback options in the three different experimental phases. In light of the high proportion of international students who study business communication, we are particularly keen to investigate the impact of students' cultural background on perceptions of team assignments and SAPA by analysing the data in terms of ethnicity, first language, domestic/international student status.

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 SAPA entries, that a significant amount of student self-censorship may have been applied to their online entries. Although qualitative feedback comments were anonymous, at least some on campus students in small teams appeared unwilling to openly criticise peers who might have been able to deduce which team-

mates made which comments. The more mature and less tolerant off-campus team members appeared less fearful of being frank in their feedback, possibly because they were unlikely to ever meet face to face. Within Phase 1, off-campus business communication students tended to me more critical of their team-mates than were their on-campus equivalents, with 59% of teams having a range of MSF greater than 0.3, but only 35% of on-campus teams being in that same category. Whether such variations were the result of off-campus/on-campus student differences, or the effect of the other demographic factors mentioned above, remains to be determined.

The potential for student collusion within a SAPA system had previously been suspected by the instructor whilst using a paper-based SAPA system to manually individualise business communication team scores (Fermelis 2006). Another example of this occurred during Phase 1 within the architecture cohort when the ratings and comments of two students who were known to live together were extremely similar to each other, and yet very different to the ratings of the third member of their team. After investigation by the course chair, the two agreed that perhaps their entries had exaggerated the differences in contributions and agreed to an adjustment. This once again demonstrates the importance of the instructor carefully monitoring what the SAPA data is showing throughout the assignment period.

The impact of self-assessment on SAPA ratings is a similarly problematic issue. There is contradictory evidence within the literature that students may inflate their self-ratings (Goldfinch 1994; Zhang et al. 2008). For this project it was decided that students should make self-ratings but that these would be excluded from the assessment matrix automatically generated by the SAPA software. It is hoped that further analysis of the SAPA data, to compare MSF's which exclude self-assessment, with those which include self-assessment may further illuminate this issue.

The significance of number of entries made, collection of continuous scores and comments versus one final, summative overall or global rating is another area worthy of closer examination. Paper-based devices previously used within business communication had been based on the collection of one data set only (Fermelis 2006), whereas the previous use of online SAPA with architecture students had always included multiple entries. In Phases 1 and 2 of this pilot, five separate entries were made with the final entry being overtly tagged as a global one. Again, comparative analyses of the data should illuminate this issue.

A final factor worthy of more detailed investigation is that of the possible avoidance of extreme response styles by members of some cultural groups (Chen, C., Lee, S., & Stevenson 1995; Chun, Campbell & Yoo 1974). The attitudes of international students to team assignments, especially those from Asian countries with non-western thinking styles, have already received some research attention (Li & Campbell 2008; Lin, Liu & Yuan 2001). However, this is an important area in which our understanding needs to be improved, as it is vital in terms in the current context of increased internationalisation of higher education, that this type of rater bias be eliminated in situations comprising homogeneous and/or heterogeneous teams made up of students from differential cultural backgrounds.

Conclusion

This pilot has demonstrated that the SAPA tool trialled is sufficiently resilient and robust to adapt to the needs of a business communication team assignment in a very large, complex, multi-cohort, multi-instructor context. Student perceptions of the SAPA tool and its ability to

facilitate reflective student learning and reduce student anxiety about uniform team assignment grades will become clearer once entry and exit questionnaires have been analysed. However, the time and effort invested by the project team produced innovative and effective software and procedural refinements. The pilot study was thus successful in identifying and strengthening weaknesses in the SAPA tool, and in producing a more refined, valid and reliable instrument. Further enhancements are still required, but the tool has demonstrated great potential for providing evidence to assist business communication and other instructors in the individualisation of team assignment grades.

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 for the duration of this 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), 427-441.

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

Burnett, W., & Cavay, G. (1980). Peer Assessment By Fifth Year Students of Surgery. *Assessment in Higher Education*, 5, 273-278.

Chen, C., Lee, S., & Stevenson, H. W. (1995). Response Style and Cross-cultural casmparisons of rating scales among East Asian and North American Students. *Psychological Science*. 6(3), 170-175.

Chun, K., Campbell, J. B., & Yoo, J. H. (1974). Extreme Response Style in Cross-Cultural Research: A Reminder. *Journal of Cross-Cultural Psychology*. 5(4), 465-480.

Cheng, W., & Warren, M. (1997). Having Second Thoughts: Student Perceptions Before and After A Peer Assessment Exercise. *Studies in Higher Education*, 22(2), 233-239.

Davies, P. (2000). Computerized Peer Assessment. *Innovations in Education and Teaching International*, 37(4), 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), 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), 287-322.

Fermelis, J. (2006) *Making Teamwork work*. Proceedings of the 71st Association for Business Communication Annual Convention, San Antonio, Texas, USA.

Fermelis, J., Tucker, R., & Palmer, S., (2008) Fair Assessment of Team Assignments in Business Communication and Architecture, *Business Communication Quarterly*,71(3), 339-345.

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), 551-569.

Goldfinch, J. (1994). Further Developments in The Peer Assessment of Group Projects, *Assessment & Evaluation in Higher Education*, 19(1), 29-35

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), 569-577.

Li, M. & Campbell, J. (2008) Asian students perceptionsof group work and group assignments in a New Zealand tertiary institution. *Intercultural Education*, 19 (3), 203-216

Lin, S. s. J., Liu, E. Z. F., & Yuan, S. M. (2001). Web-based peer assessment: feedback for students with various thinking-styles. *Journal of Computer Assisted Learning*, 17, 420-432.

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.

May, G.L. (2008) The effect of rater training on reducing social style bias in peer evaluation. *Business Communication Quarterly*, 71(3), 297-313.

May, G.L., & Gueldenzoph, L.E. (2006) The effect of social style on peer evaluation ratings in project teams. *Journal of Business Communication*, 43(4), 4-20.

Michaelsen, L. K. (1992). Team learning: a comprehensive approach for harnessing the power of small groups in higher education. *To Improve the Academy*, 11, 107-122.

Raban, R., & Litchfield, A. (2007). Supporting peer assessment of individual contributions in groupwork. *Australasian Journal of Educational Technology*, 23(1), 34-47.

Rafiq, Y., & Fullerton, H. (1996). Peer assessment of group projects in civil engineering. *Assessment & Evaluation in Higher Education*, 21, 69-81.

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), 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), 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), 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, 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), 266-269.

Topping, K. (1998). Peer Assessment between Students in Colleges and Universities. *Review of Educational Research*, 68(3), 249-276.

Tucker, R. (in press). 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., Fermelis, J. & Palmer, S. (in press). Designing valid and reliable peer assessment in e-learning environments, Spratt, C., & Lajbcygier, P. (eds.) *ELearning and Advanced Assessment Technologies: Evidence-Based Approaches* (Advances in Information Communication Technology Education Book Series), IGI Global Publishing.

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), 39-56.

Tucker, R., & Rollo, J. (2006). Teaching and Learning in Collaborative Group Design Projects. *Journal of Architectural Engineering & Design Management*, 2(Teaching and Learning in the Built Environment), 19-30.

Walker, A. (2001). British psychology students' perceptions of group-work and peer. *Psychology Learning and Teaching*, 1(1), 28-36.

Zhang, B., Johnston, L., & Bagic Kilic, G. (2008). Assessing the reliability if self- and peer rating in student group work. *Assessment & Evaluation in Higher Education.* 1-12

Biographies

JAN FERMELIS
Jan has made significant contributions to the strategic priorities of Deakin University. She has received many awards for her passionate teaching of business and academic communication, and efforts to improve the learning outcomes of international students. In recent years she has worked to improve her research profile and has recently commenced doctoral studies in the area of intercultural business communication.

RICHARD TUCKER
Richard worked for five years in professional practice before completing his PhD in architecture in the UK. Although he is an Early Career Researcher, he has been awarded many grants and prizes and already built an impressive publication record. He is particularly interested in architectural teaching and learning in relation to design projects and in collaborative research projects focusing on group design projects.