Supporting peer assessment of individual contributions in groupwork

Richard Raban, Andrew Litchfield
Faculty of Information Technology
University of Technology, Sydney

The ability to assess the work of others is a core attribute for most professionals. To develop this graduate attribute in our students requires the learning of self and peer evaluation, feedback, and review skills. This paper discusses the changing design of peer assessment and the impact of a new groupwork support tool within a capstone undergraduate subject with large student numbers - Systems Development Project – in the Faculty of Information Technology at UTS.

Since 1998 by implementing different support strategies for peer assessment of individual contributions the distribution of the students marks has markedly widened, and now more reflect the reality of differing team member contributions. This substantial change has occurred with the use of an online tool which supports the development of student evaluation, feedback and review skills when peer-assessing individual contributions to large group projects. In use since 2004 the groupwork support tool is called Team Contribution Tracking (TeCTra).

Keywords: designing peer assessment, online learning support, developing graduate professional attributes

Introduction

In many disciplines higher education courses include significant capstone subjects involving projects that require large student teams. When facilitating peer assessment with a holistic approach (Schechtman, 1992; Schechtman & Godfried, 1993) the common assessment strategy for groupwork of allocating the same or almost the same mark to all team members (Rosen, 1996; Lejk & Wyvill, 2001; Kennedy, 2005) is not adequate as the project tasks are extensive, the teams are large in number (more than 4 members), extend for the whole semester and groupwork can constitute 100% of the final student assessment. The subject coordinator has limited opportunities to observe and assess the complex group and teamwork dynamics that are taking place. A peer-assessment strategy is required which is ideally formative, diagnostic and summative (Goldfinch, 1990; Gatfield, 1999). This ideal has been difficult to achieve (Lejk & Wyvill, 2001; Li, 2001) and remains as an important and unresolved feedback and assessment issue.

Peer assessment has been shown to support not only students' learning but also improve their understanding of the assessment processes themselves (Bloxham & West, 2004). Peer assessment is also required to assess individual contributions to group assignments (Johnston & Miles, 2004). The development of the evaluation, feedback and review skills required to peer assess these complex teamwork processes is a key learning objective of such large project-based capstone subjects. These are skills every professional should possess and be able to use for different purposes. It is also important for the novice professional to experience being on the receiving end of peer-reviews and assessment and to learn to benefit from any feedback received.

Peer-assessment for assessing individual contributions to groupwork is controversial not only because it can produce 'unreliable' results caused by the inexperience of the student assessors and often by rather undifferentiated marks (Kennedy, 2005). Also the labour intensive processes the subject coordinators have to administer are problematic (Clark et al., 2005). This paper addresses these concerns and presents a peer-assessment strategy and online tool for the peer-identification of students' individual contributions in a large groupwork-based capstone subject in software system design.

The presented learning and teaching strategy and online tool requires the students to rate and comment on each other on a weekly basis. This task is informed and supported by evidence of the work done and

outcomes achieved by each student. The strategy creates a formative, diagnostic and summative assessment environment in which the students can learn the skills of peer-assessing their peers using quantitative ratings and qualitative comments. This peer-assessment strategy has delivered greater differentiation of student marks than those reported in the literature and experienced by the authors in the period before the introduction of the TeCTra online tool. The online tool and system for data collection, presentation and calculating contribution factors has released the subject coordinator from the enormous work otherwise required to process any similar paper-based strategy.

Subject description

Systems Development Project (SDP) is a capstone subject in the Bachelor of Science in Information Technology at UTS with 350-400 students each year. The degree has three years of course work and a year of industry training. SDP is taught in the second semester of the second year and aims to prepare the students for industrial training in the third year. Before undertaking SDP the student has completed three semesters of IT education in programming, systems design and development, networking and information systems. During SDP the students experience working in a team and learn how to apply their prerequisite knowledge to a practical system development problem. During the project they develop a system from specifications to a working software product.

SDP involves groups of 10 students in a major project that takes 50% of their study time (12 credit points) for a full time student for one semester of 15 weeks. Groups have a great degree of autonomy. They are responsible for planning and allocating project tasks and organizing work in the groups. Academic tutors, usually project managers from industry, are subject Project Managers responsible for overseeing the groups' progress and attending to problems with group dynamics and project work.

There are two milestones in the project, a mid-semester review and a final review, and each produces 50% of the final assessment. The two assessments comprise of a peer review (worth 40%) and a staff review that assesses written submissions (60%).

Peer assessment in the project

Students are required to undertake a number of peer-assessment activities. Firstly, they review other groups' work at the two project milestones of the mid-semester review and the end-of-semester final review. Secondly the groups are asked to assess individual contributions to the project made by the team members. This assessment is done formatively and progressively during the semester, and then summatively during the mark allocation of the mid-semester and final reviews.

During the peer reviews each group assesses an oral presentation given by another group. The presentation takes 20 minutes and is followed by 10 minutes of question and answer time. The reviewers make their assessment against a set of given criteria that the designers were to achieve through their solutions. During the presentation each member of the reviewing team does their own assessment of the presented solutions. The group discusses the individual marks and consolidates them into a group assessment which is given to the presenters and accounts for 40% of the total mark. There is a requirement that the marks given to the other group are properly justified and both the advantages and disadvantages of the presented designs are assessed.

The project outcomes as assessed by the peer and staff reviews produce an overall mark for the group effort. This mark is then multiplied by the number of students in the group and the result becomes a pool of marks that the group members must distribute amongst themselves according to their assessment of individual contributions to the project. Guided by instructions given to them in the assessment policies and procedures a meeting of all the team members is convened to discuss the mark allocation. The groups are advised to start the meeting with a round of statements by the team members about their respective contributions to the project. Then through discussion and negotiation the group arrives at an allocation of the marks that all team members can agree on. The results are then presented to the Project Manager, a staff member, for approval. Once the consensus on the mark allocation is confirmed the individual marks are accepted.

Supporting peer assessment of individual contributions

In capstone subjects with large groupwork projects students are often given responsibility to allocate individual marks according to the perceived individual contributions made by each team member. Time and again this responsibility has proved too difficult for the students to dispense properly resulting in an equal distribution of marks irrespective of the actual contributions (Rosen, 1996; Lejk & Wyvill, 2001; Kennedy, 2005; Raban & Litchfield, 2006). As a result, good students are dissatisfied with their summative grade and marks while those students who chose to do very little receive undeserved rewards.

In the SDP capstone subject the groups of 7 to 10 students experienced the same problem. In recognition of this, the students were given an increasing level of support in peer assessment across the 8 years in which the subject has been offered. There are three distinct periods in which peer assessment of individual contributions was assisted in different ways. These are:

- Summative assessment of contributions without on-line support (years from 1998 to 2001);
- Summative assessment of contributions with time recording (years from 2002 to mid-2004); and
- Formative and summative assessment of contributions with time recording and weekly ratings (years from mid-2004 to now).

Development of peer-assessment support was gradual and the results of different approaches closely monitored. The ability to differentiate final marks by the groups was used as a measure of the impact. To make the analysis statistically significant, it was performed only for semesters with 10 or more groups. For each group, a coefficient of standard deviation of the final individual marks was calculated. It was used as an indicator of to what extent the group was able to align marks with individual contributions. For each semester a graph showing the percentage of groups that differentiated their contributions by 0-5%, 6-10%, 11-15%, 16-20% and 21%+ were plotted.

Summative assessment of contributions without on-line support

In the years from 1998 to 2001 the students had to rely on their own records and recollections of individual contributions in allocating individual marks. The only support given to the groups was a set of rules and policies that spelt out a range of good practices for peer assessment. Occasionally groups were not able to reach a consensus and a staff member was called in to break a stalemate in the mark negotiations. It has to be stressed however that the academic tutors would never engage in the actual assessment of contributions. Instead the tutor would assist the group to choose an acceptable method of assessing contributions and then assumed the role of an impartial facilitator of the method's implementation. As a result in semesters Spring 1998, Spring 1999 and Spring 2001, the distribution of peer marks were diversified as shown in Figure 1. Note the thick line that shows the period total and represents the overall trend in the period.

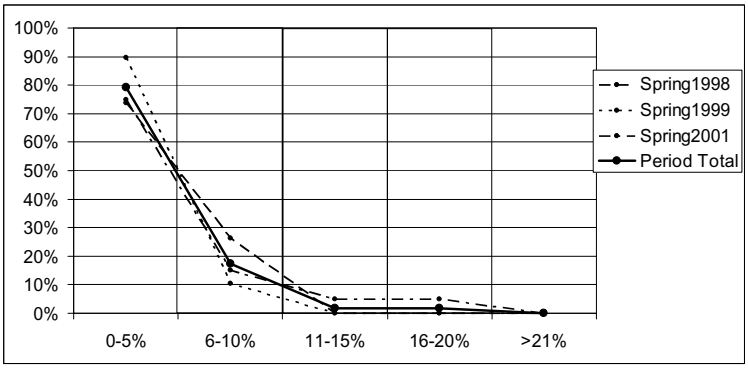


Figure 1: Period with no online support

The graph shows that between 75% to 90% of all groups opted for almost equal mark distribution. It was an expected result in line with similar cases reported in the literature (Rosen, 1996; Lejk & Wyvill, 2001;

Kennedy 2005). This nearly equal distribution of marks was hardly plausible as in groups of 10 students one can expect a wide-range of individual contributions.

Summative assessment with time recording

In the years from 2002 to mid-2004, in order to better support peer assessment of individual contributions the students used an online tool for recording individual time spent on the project. The time records were collected on a weekly basis, stored in the system and made available to all the team members for perusal as in Figure 2. While reporting the hours, the students had to state which task and what type of work the hours were spent on. In Figure 2 the student Jennifer Law spent a total of 15 hours working as a Project Leader and on Requirements Specification tasks engaging in management, development, documentation and quality review.

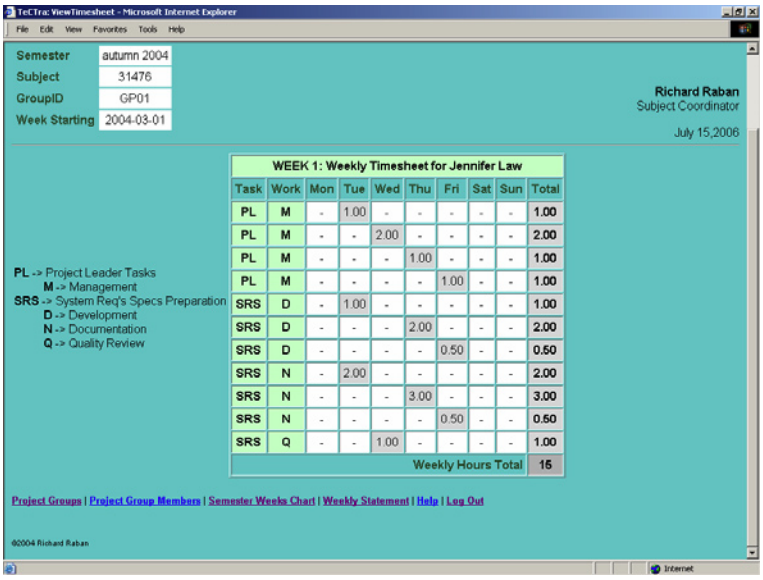


Figure 2: Time records available for peer assessment

The time records made individual efforts visible to the team members and thus could be used to inform the process of assessing individuals. It ensured that all work from early attempts, possibly no longer visible in the final product and easily forgotten, could be taken into account in the summative mark allocation.

As a result of the use of time records in semesters Spring 2002, Spring 2003 and Autumn 2004 the distribution of peer marks were diversified as shown in Figure 3. Note the thick line that shows the period total and represents the overall trend in the period.

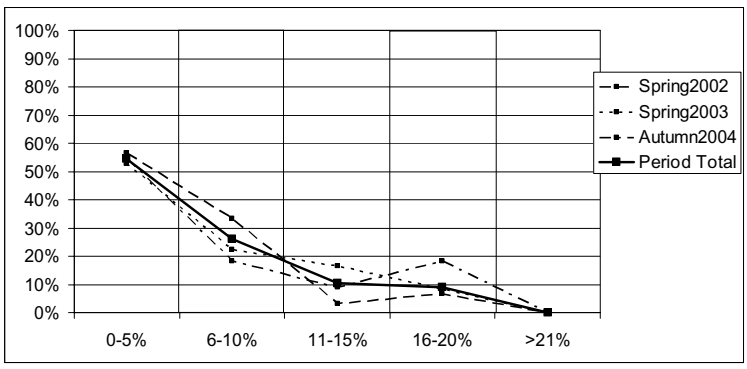


Figure 3: Period with time records available

An analysis shows that the time recording tool reduced from 75-90% to 55% the percentage of the groups electing an easy way out by giving everybody equal or almost equal marks. Providing the students with evidence of who was doing what and how much time was spent on the tasks empowered a greater number of groups to diversify marks. It was an improvement on the previous situation yet the general pattern of most groups was still in the 0-5% band indicating ongoing difficulties with the peer-assessment of individual contributions. In both periods when an holistic summative peer-assessment strategy was used, the results were consistent with those presented by other authors (Rosen, 1996; Lejk & Wyvill, 2001; Kennedy 2005).

Apparently being informed about time spent does not easily translate into peer-contribution ratings as time-records do not take into account the quality of work and the level of participation in leadership, motivating team members or organising team work. The results achieved by (Lejk & Wyvill, 2001) seem to confirm this argument. It was demonstrated that the summative category-based approach to peer-assessment in group projects produced a wider and smoother distribution of individual marks than the summative holistic one. The holistic approach produced a lot of almost equal marks with only extreme cases of over-or-underperformance reflected in mark differentiation. However the likely cause of this mark distribution is the fact that minor differences in contributions are not easy to quantify especially if there is a considerable time-lapse between the work done and its assessment. Only very poor or outstanding efforts seem to be recognised and reflected in the distribution of marks.

Formative and summative assessment with time recording and weekly ratings

In the three semesters from mid-2004 to 2005 the shortcomings of a time-based peer-assessment strategy have been addressed with the design and implementation of an online groupwork tool called TeCTra. The tool facilitates:

- time recording - now with an additional feature to allow recording of deliverables that were produced during the time reported;
- a confidential feedback system that allow team-members to make comments on individual contributions directly to each other;
- ratings of individual contributions that were done on a weekly basis; and
- a progressive calculation of weighted weekly contribution-factors that indicate how the group rated individual contributions on a weekly basis.

Weekly ratings make peer-assessment much easier as the assessed work and group interactions are still remembered by the team-members. It also takes away the stress of summative assessment that contributes a large portion of the final mark and therefore creates a lot of tension within the group. The students can be more honest in making judgements on their peers. Contribution factors calculated for each week and weighted contribution factors calculated for a period of time give quantitative diagnostic feedback to each team member regarding their standing in the group. Additionally, qualitative comments on the work done by team members can be exchanged facilitating formative feedback that explicitly raises groupwork issues with the students. The students are not only given quantitative feedback on their perceived contributions but also receive qualitative feedback on their group work efforts.

Each reporting week every team member was required to rate the contributions of all team members as above normal (3), normal (2), below normal (1) or no contributor (0). This scale is similar to the one proposed in Goldfinch (1994) with the exception of not including a hindrance to the group (-1) rating. To inform the peer-judgement, each week's deliverables and time-spent was displayed as shown in Figure 4. Additionally, each rating could be accompanied by feedback to the person rated indicating perceived problems or good points of their participation. The feedback was only visible to those to whom it was addressed without disclosing who sent it. In Figure 4, the student Bernard Greenspan sent comments to four team members. The feedback to the student Patrick Coble shown in Figure 5 includes Bernard Greenspan's comment without disclosing who sent it. No other student could see the thus confidential comment.

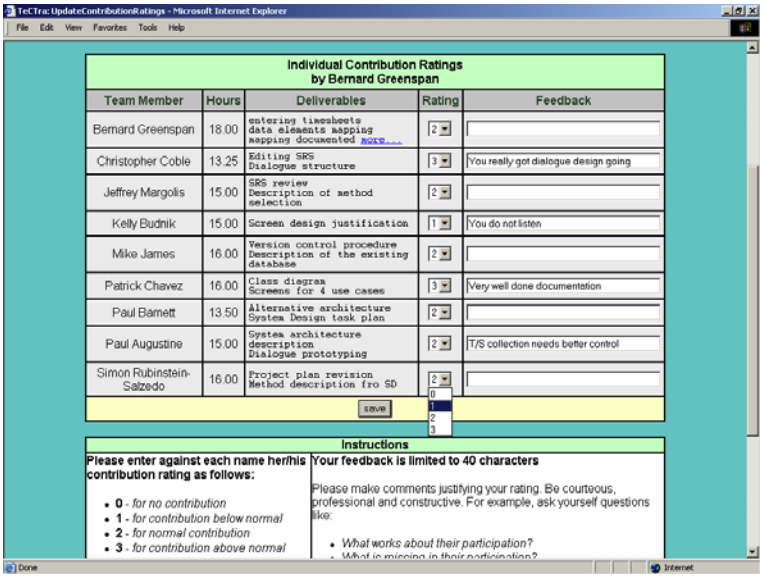


Figure 4: Weekly peer ratings entry screen

In Figure 4 ratings for the week are combined into Weekly Member Contribution Factors of team members. The ratings of each team-member are totalled and an average rating calculated, then each individual total is divided by the average and multiplied by 10. The consequent Weekly Member Contribution Factor reflect the whole group's perception of individual contributions informed by the time records, ratings and comments and also reflects the quality of groupwork.

The Weekly Member Contribution Factors always represent a distribution of a fixed-pool of marks among the team-members so the Factor is indicative of contribution levels. In the subject SDP the pool of contribution points is 10 times the number of team-members. If everybody contributed equally, all team members would have a Weekly Member Contribution Factor equal 10. Therefore, an individual contribution factor below 10 indicates a below average contribution, and a factor above 10 indicates a contribution better than average.

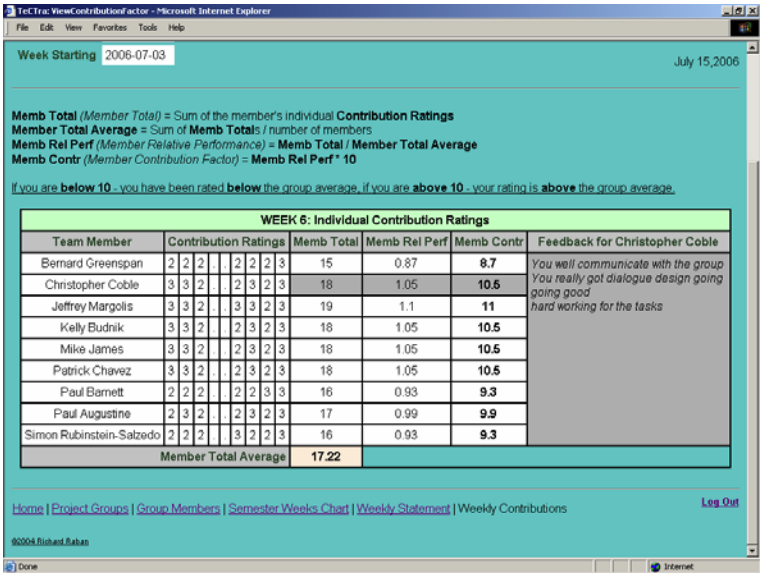
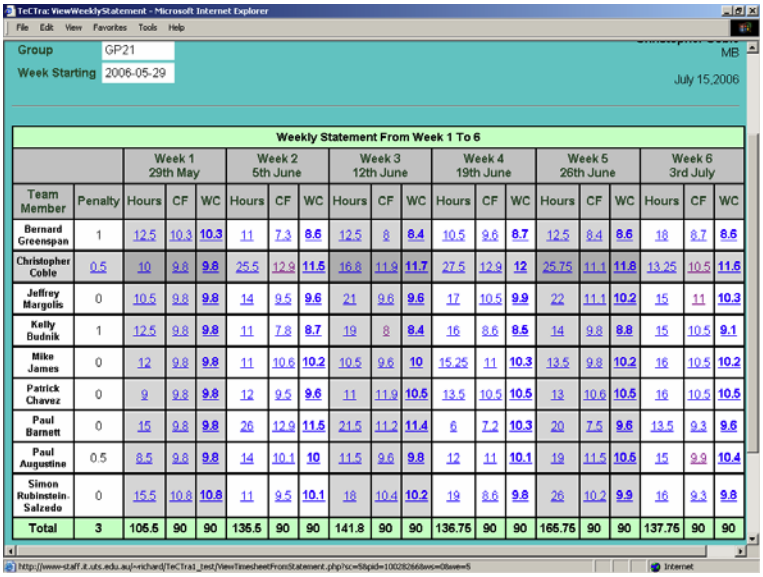


Figure 5: Weekly contribution factors

Time records (Hours column), Weekly Member Contribution Factors (CF) are shown for all team members across all the weeks already reported on (see Figure 6). Additionally, Weighted Contribution

Factors are calculated progressively across the period of time showing the members' standing in terms of contributions for a period of time. In these calculations, Weighted Member Contribution Factors (WC) are weighted by the total weekly hours spent on the project by the team. The rational behind it is that a week with 100 hours spent is expected to contribute to the project roughly half of what a week with 200 hours spent do.

On the screen in Figure 6 Christopher Colbe appears to be the top contributor of the group with a Weighted Contribution Factor of 11.6. This progressive peer-assessment of individual contributions allows team-members to see how their contributions are perceived by the whole group and to modify their interactions with the group in order to improve their rating and standing within the group.


TeCTra: ViewWeeklyStatement - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Group GP21

Week Starting 2006-05-29

MB

July 15,2006

Weekly Statement From Week 1 To 6

		Week 1 29th May			Week 2 5th June			Week 3 12th June			Week 4 19th June			Week 5 26th June			Week 6 3rd July		
Team Member	Penalty	Hours	CF	WC	Hours	CF	WC	Hours	CF	WC	Hours	CF	WC	Hours	CF	WC	Hours	CF	WC
Bernard Greenspan	1	12.5	10.3	10.3	11	7.3	8.6	12.5	8	8.4	10.5	9.6	8.7	12.5	8.4	8.6	18	8.7	8.6
Christopher Coble	0.5	10	9.8	9.8	25.5	12.9	11.5	16.8	11.9	11.7	27.5	12.9	12	25.75	11.1	11.8	13.25	10.5	11.6
Jeffrey Margolis	0	10.5	9.8	9.8	14	9.5	9.6	21	9.6	9.6	17	10.5	9.9	22	11.1	10.2	15	11	10.3
Kelly Budnik	1	12.5	9.8	9.8	11	7.8	8.7	19	8	8.4	16	8.6	8.5	14	9.8	8.8	15	10.5	9.1
Mike James	0	12	9.8	9.8	11	10.6	10.2	10.5	9.6	10	15.25	11	10.3	13.5	9.8	10.2	16	10.5	10.2
Patrick Chavez	0	9	9.8	9.8	12	9.5	9.6	11	11.9	10.5	13.5	10.5	10.5	13	10.6	10.5	16	10.5	10.5
Paul Barnett	0	15	9.8	9.8	26	12.9	11.5	21.5	11.2	11.4	6	7.2	10.3	20	7.5	9.6	13.5	9.3	9.8
Paul Augustine	0.5	8.5	9.8	9.8	14	10.1	10	11.5	9.6	9.8	12	11	10.1	19	11.5	10.5	15	9.9	10.4
Simon Rubinstein-Salzedo	0	15.5	10.8	10.8	11	9.5	10.1	18	10.4	10.2	19	8.6	9.8	26	10.2	9.9	16	9.3	9.8
Total	3	105.5	90	90	135.5	90	90	141.8	90	90	136.75	90	90	165.75	90	90	137.75	90	90

Figure 6: Formative and summative assessment of contributions.

As a result of using TeCTra in semesters Spring 2004, Autumn 2005 and Spring 2005, the distribution of peer-marks show only about 20% of groups distributing marks close (0-5%) to an even distribution. Note the thick line that shows the period total and represents the overall trend in the period.

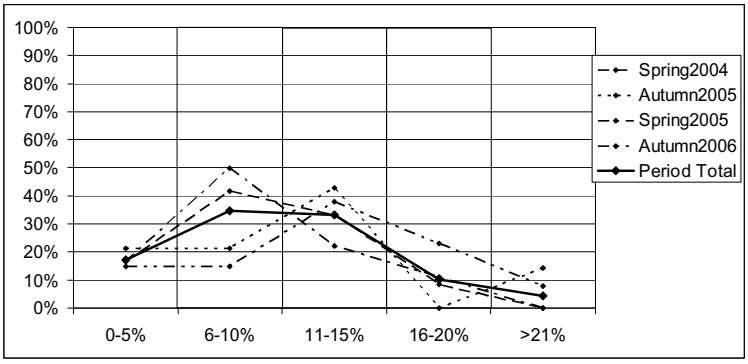


Figure 7: Period supported by TeCTra

The introduction of TeCTra produced a dramatic change in the students' peer-assessment of individual contributions. Only 20% of groups allocated marks with little or no diversification and 65-75% of groups diversified marks by 6-15%. This is a more accurate reflection of the expected range of students' individual contributions within a large groupwork project. As the students were not obliged to use the TeCTra calculated individual Weighted Contribution Factors in allocating their marks, it seems that progressiveness and visibility of peer evaluation, feedback and review empowered individuals to claim

their 'rightful' share of the marks. Non-performers were exposed early through the formative assessment of their peers, and they had two options, either to improve or to accept lower summative marks.

Conclusion

Figure 8 combines the graphs of the distribution of peer marks for the three peer assessment approaches discussed in the paper. The results demonstrate that without TeCTra's online support the students were not capable of reflecting individual contributions in the marks allocated to team-members and an equal distribution of marks was given to 75-90% of their peers. The visibility of individual work on the project provided by the online time-records improved the situation by reducing the percentage of groups with a near-equal mark allocation to about 55%. This result proved that reliable evidence of individual efforts empowered team members to claim better marks and the groups were willing to accept resulting mark differentiation. The most significant change in peer-assessment mark distribution occurred with the introduction of the current TeCTra system that has facilitated peer evaluation, feedback and review assessment processes. An equal distribution of peer-marks is now only about 20% and the distribution has become significantly wider and a better reflection of the variety of individual contributions to large groupwork project outcomes.

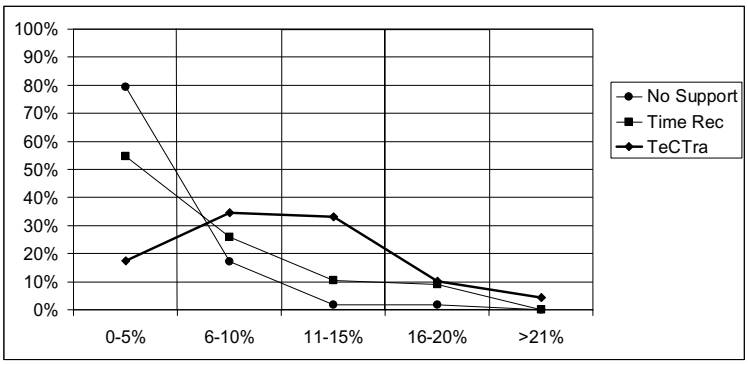


Figure 8: Overview of the changes in the distribution of peer-marks due to changes in peer-assessment strategies from 1998 to 2005.

TeCTra provides visibility of individual efforts and outcomes. Apart from time records collected in the earlier system, TeCTra also records deliverables produced. While rating their team members, the students were presented with all individual results produced in the week being assessed. It ensured that the rating process was evidence-based.

TeCTra supports peer evaluation, feedback and review – both a quantitative rating and qualitative comment – throughout the duration of the project and thus formatively and positively influences individual contributions and behaviours within the team. This improved capacity for peer-review facilitates diagnostic attributes and thus significantly influences the overall project management process and outcomes.

TeCTra supports the development in students of the ability to evaluate, give feedback, review and assess the work of others, to make professional judgments, to articulate well-justified decisions and to communicate in a non-confrontational manner to their peers – core skills and graduate attributes for most novice professionals. Knowledgeable yet inexperienced individuals are supported to act professionally and take responsibility for and accept the consequences of their own contributions to large groupwork projects.

TeCTra is relatively simple for the students and the staff to operate and avoids complexities and additional work that present in other online tools (Clark, 2000). The online tool's user-friendliness is important as increasing academic teacher workloads leave minimal time for the administration of elaborate self-and-peer assessment methods and tools (Fisher, 1999).

There is still a question about whether TeCTra produces marks that do reflect the true individual contribution of each team-member. The students are not under obligation to use TeCTra contribution factors for their peer-mark allocation and indeed the majority of groups choose not to use the TeCTra contribution factors. Yet there has been no return to the previous practice of allocating marks close to an equal distribution. It has to be concluded that the online tool did make the difference. It seems to have changed the group dynamics although the mechanisms behind that are not quite clear and will be investigated through student usability evaluations in future research.

References

Beasley, R.E., & Vila, J.A. (1992). The identification of navigation patterns in a multimedia environment: A case study. *Journal of Educational Multimedia and Hypermedia*, 1(2), 209–222.

Bloxham, S. & West, A. (2004). Understanding the rules of the game: Making peer assessment as a medium for developing students' conceptions of assessment. *Assessment and Evaluation in Higher Education,* 29(6), 721-733.

Clark, N., Davies, P. & Skeers, R. (2005). Self and peer assessment in software engineering projects. *Proceedings of the 7th Australasian Conference on Computing Education*. Newcastle, Australia

Fisher, R. (1999). *Academic Workloads and Assessment.* http;//www.tedi.uq.edu.au/conferences/teach_conference99/papers/fisher.html

Gatfield, T. (1999). Examining student satisfaction with group projects and peer assessment. *Assessment and Evaluation in Higher Education*, 24(4), 365-377.

Goldfinch, J. (1994). Further developments on peer assessment of group projects. *Assessment and Evaluation in Higher Education*, 19(1), 29-35.

Johnston, L. & Miles, L. (2004) Assessing contributions to group assignments. *Assessment and Evaluation in Higher Education,* 29(6), 751-768.

Kearsley, G. (2004). *Explorations in Learning & Instruction: The theory into practice database.* http://www.gwu.edu/~tip/ [viewed 14 Jun 2004].

Kennedy, G.J. (2005). Peer-assessment in group projects: Is it worth it? *The Australasian Computing Education Conference 2005*, Newcastle, Australia.

Li, L. (2001). Some refinements on peer assessment of group projects. *Assessment and Evaluation in Higher Education,* 26(1), 5-18.

Lejk, M. & Wyvill, M. (2001). Peer assessment of contributions to a group project: A comparison of holistic and category-based approaches. *Assessment and Evaluation in Higher Education*, 26(1), 19-39.

O'Shea, T. & Self, J.A. (1983). *Learning and teaching with computers*. Englewood Cliffs: Prentice-Hall Inc.

Raban, R. & Litchfield, A. (2006). Peer assessment in large group projects: forming professional attitudes in IT students. *Uniserve 2006.*

Rosen, C.C.H. (1996). Individual assessment of group projects in software engineering: A facilitated peer assessment approach. *Proceedings of the 9th Conference on Software Engineering Education*. Daytona Beach, Florida.

Schechtman, Z. (1992). A revised group assessment procedure administered in several academic settings. *Journal of Personal Evaluation in Education*, 6, 31-39.

Schechtman, Z. & Godfried, L. (1993). Assessing the performance and personal traits of teacher education students by group assessment procedure: A study of concurrent and construct validity. *Journal of Teacher Education*, 44(2), 130-138.

Underwood, J. (1997). Breaking the cycle of ignorance: Information technology and the professional development of teachers. In D. Passey & B. Samways (Eds.), *Information Technology: Supporting change through teacher education*. (pp.155–158). London: Chapman & Hall.

Author contact details

Richard Raban, Faculty of Information Technology, University of Technology, Sydney, PO Box 123, Broadway, NSW 2007, Australia. Email: richard@it.uts.edu.au.

Andrew Litchfield, Faculty of Information Technology, University of Technology Sydney, PO Box 123, Broadway, NSW 2007, Australia. Email: ajl@it.uts.edu.au.

Copyright © 2006 Raban, R., Litchfield, A.

The author(s) assign to ascilite and educational non-profit institutions a non-exclusive licence to use this document for personal use and in courses of instruction provided that the article is used in full and this copyright statement is reproduced. The author(s) also grant a non-exclusive licence to ascilite to publish this document on the ascilite web site (including any mirror or archival sites that may be developed) and in electronic and printed form within the ascilite *Conference Proceedings*. Any other usage is prohibited without the express permission of the author(s). For the appropriate way of citing this article, please see the frontmatter of the *Conference Proceedings*.