World Transactions on Engineering and Technology Education
Vol.2, No.2, 2003

Using collaborative learning and peer assessment in an undergraduate engineering course: a case study

Małgorzata S. Żywno

Ryerson University
Toronto, Canada

ABSTRACT: The paper describes an undergraduate course in control where a student-centred educational model was successfully introduced. While the main focus is on developing better engineering design skills, collaborative approach also results in improved communication and critical judgement skills. Peer assessment empowers and motivates students, and classroom instruction is supported by educational technology. Student teams complete technical projects as well as research more general, control-oriented topics, and prepare multimedia presentations on them. Pearson's correlation coefficients between student and instructor ratings of the presentations suggest that peer review is a valid assessment tool. Student performance, instructor-course evaluations and student feedback suggest increased satisfaction with the process, greater motivation and improved learning outcomes.

INTRODUCTION

A rapidly changing knowledge base and the need for students to become life-long learners and develop critical thinking skills requires a more effective pedagogy, including the use of active, collaborative learning, instructional technology and different evaluation methods. Engineering educators thus increasingly embrace a learner-centred educational paradigm [1].

Learner-Centred Engineering Education

The traditional, instructor-centred paradigm is to transfer faculty's knowledge to passive students, who are classified through norm-based competition [2]. The new, learner-centred paradigm, while retaining the principles of learning objectives that are crucial from the point of view of professional, accredited curricula, is based on collaborative learning in the context of real-life applications [2]. Individual differences and cognitive theories are acknowledged and have implications for education, with varied assessment formats also promoted [2][3]. The new paradigm embraces instructional technology and the Internet applications that have the potential to change the balance of power in the classroom. They are seen as enhancing the capabilities of both the learner and the teacher, providing opportunities to lessen the reliance on lecturing and encourage complex interactions between learners, experts and content through the use of technology [4].

Active, collaborative learning has a measurable effect on student performance [5][6]. Group projects and presentations can be used to assess the ability of individuals to work as a team and to communicate ideas [7]. Peer and self-assessment are an important part of the student-centred paradigm, as they have the potential to empower and motivate students [3]. They can also provide information on students' performance not captured by traditional evaluation methods [8]. They encourage students to develop a deep learning approach to the subject and provide them with an opportunity for reflection that enables learning. However, peer and self-assessment can also be too demanding of students, time consuming and unreliable. Common identified problems include *friendship marking*, resulting in over-marking, *collusive marking*, resulting in a lack of differentiation, and *parasitic marking*, where students fail to contribute, but benefit from group marks [9]. Such problems can fuel instructors' reluctance to include peer-based assessments [6][7][9]. Clearly, the design of any peer review is crucial to its success and assessment criteria must be well identified and made explicit [6][9].

ELE829: SYSTEM MODELLING AND IDENTIFICATION

ELE829 course is an eighth semester course in the Electrical and Computer Engineering programme at Ryerson University, Toronto, Canada. It provides an introduction to modern system identification, which is essential to any advanced control design. Topics include: empirical, deterministic-stochastic modelling of unknown processes, the use of correlation analysis for diagnostics, model selection and validation, and different estimation algorithms. As a professional elective, it combines a fairly advanced theory with a practical design component. The subject matter, typically taught at a graduate level, presented a considerable challenge when offered in an undergraduate environment.

Initial Instructional Design

The course has a three-hour lecture/one hour tutorial format, favouring theory over practical applications. When it was first offered in 1995, the author attempted to reduce reliance on a traditional lecture and encourage a more active learning by combining lectures with discussions and hands-on work on computer simulations (eg *MATLAB* and *System Identification Toolbox*). Students could also prepare for classes and review the course material by downloading PostScript files from the departmental computer network. Since the challenging subject matter did not yield itself easily to a conventional testing format, course evaluation, while still based on individual effort, replaced conventional written tests with a portfolio-type assessment of tutorial work and assignments.

Changes in Instructional Design

The first two offerings of the course turned out to be a rather frustrating experience, for both students and the author. While final grades were high, students found the course overly difficult and reported spending a disproportionate amount of time working on it. The course was not offered in 1997 and 1998, allowing the author to rethink its philosophy and to make a conscious effort to firmly ground it in the tenets of a progressive, learner-centred education. In 1999, new instructional media were introduced, collaborative and communication skills were emphasised, and peer and self-assessment became part of the course evaluation. Materials on the departmental network lacked the *anytime-anywhere* flexibility and interactivity of the WWW and were replaced by a course Website with multimedia components. *WebCT*, a software package for online course management and delivery, enabled students to interact in discussion groups and presentation areas and access course materials and grades securely at their own time and pace. The supporting Website and asynchronous communications (ie e-mail, bulletin board) improved time on task and allowed more flexibility. Class time was divided between lectures and laboratories on a week-by-week basis, as required.

The emphasis in the course shifted towards developing better design skills through collaboration. Group, rather than individual, effort, became the basis for evaluations, which included a series of tutorials, a formal report on a major design project, an oral presentation, as well as a self-assessment exercise. The collaborative aspect reduced student anxiety regarding the perceived difficulty of the course. Class presentations and peer-assessments allowed students to hone their communication and critical judgement skills, nowadays more frequently expected by employers [6][9][10]. The course redesign proved very successful. As Figure 1 shows, its enrolment continues to increase.

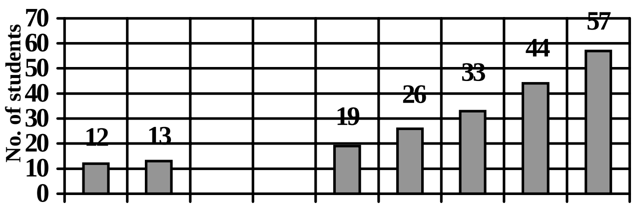

Figure 1: Enrolment trends in ELE829.

In 2002, in order to maintain the active learning aspect of the course despite the large class size, the four hours of contact time were split between two smaller tutorial groups, supported by online self-study. Following negative student feedback on the disappearance of the face-to-face discussion/lecture time, the common lecture was restored in 2003 and the increased enrolment (currently capped at 60) was accommodated by adding another section to the course.

The Introduction of Individual Review Marks

In 1999, the first year that peer-assessments were introduced, several cases of friendship marking were identified. Inflated and collusive marking in peer- and self-assessments was also encountered, although it did not have significant bearing on the overall fairness of the process due to its low weighting in the overall scheme. This underscored the fact that without a well-thought out design of peer-assessment, even senior level students can and do engage in strategies that are counter-productive to intended learning [6][9].

Thus, in 2000, a mechanism was introduced that encouraged and rewarded more objective evaluations. Students are now asked to evaluate their own group on a scale together with other peer-assessments, which gives them a more realistic look at their own effort. Weighting of peer assessments increased to 15% of the total grade, but now students receive individual marks for their reviews, also worth 15%. Guidelines are offered, but students are free to choose their own marking criteria. In 1999, when an absolute scale for the reviews was used, grading disparities occurred because of inexperience and of different personal criteria, or lack of them. To avoid that, students now convert their reviews into rankings, thus using a *relative*, rather than absolute, evaluation scheme. At the end of the semester, each student submits documented assessments of all group presentations, including his/her team's, ranking them from top to bottom. Next, the instructor's ranking is combined with the class average (at equal weights) to create a *template ranking*. The student's individual mark for participation in the review process is then computed on a sliding scale, using Pearson's correlation coefficient between the template and the ranking submitted by the student. A statistically significant correlation factor equal to or above the class median ensures a full individual mark. Individual reviews convergent with perceptions of the class as a whole, and of the instructor, earn higher review marks. Unreasonable reviews result in the reduced marks. Thus, while a collusive peer- or self-assessment may benefit a group by improving their overall peer assessment mark, any such *gains* will be offset by low individual marks for the colluding reviewers. In a course where correlation analysis plays a significant role in testing model validity, this approach provides an effective lesson in practical applications of statistics, as well as a built-in incentive for objectivity.

Within-Group Peer Assessment

In 2000 and 2001, teams consisted of two, rarely three, students, and all received common grades for their teamwork. The individual review mark provided the only differentiation of grades within a team. The instructor did not intervene in any workload arrangements and students chose coping strategies suited to their group dynamics. In 2002, the climbing enrolment necessitated a move further along the collaborative assessment route. While more tasks may be accomplished, it is more difficult to negotiate the division of labour among five or six partners than it is between two or three. Yet working effectively together, negotiating differences and being able to assess the contributions of others, as well as one's own, is an important aspect of collaborative learning, as well as a reality of the workplace [6][9][10]. Thus, in 2002, the assessment scheme was again modified to include peer-assessment *within the group*, which provided an additional differentiation of individual marks. At the end of the semester, students confidentially rate all team members, including themselves, on their contributions to the team accomplishments. The following ratings are used: *excellent*, *very good*, *satisfactory*, *marginal*, *unsatisfactory*, and *no show*, with appropriate definitions [6]. Examples to justify the rating have to be provided. The number of *excellent* and *very good* ratings is capped at two each, to avoid parasitic and non-discriminant assessments. The instructor uses these ratings to assign a proportion of the group mark to an individual. An average *very good* rating corresponds to a 100% share of the group mark. If a student receives a uniformly *excellent* evaluation from his/her peers for consistently going above and beyond of the call of duty, such as showing initiative, organising communications, tutoring others, etc, he/she receives a leadership bonus mark.

The course organisation and marking scheme is clearly explained at the beginning of the course to enhance ownership and to ensure that fairness and openness of the process are established. Students tend to express initial apprehension and distrust regarding peer and self-assessment. The instructor addresses their concerns by explaining the checks and balances on the peer review. These include the relative evaluation scheme and correlation analysis employed in the process. The instructor also engages students as their mentor. This includes help in conflict resolution when, occasionally, interpersonal dynamics and perceived inequities in carrying the workload cause team dysfunction. There have been no incidents yet of a severe dysfunction that would impede the team goals or would necessitate breaking the team up.

Group Presentations

As the enrolment, and thus group size, increased, additional tasks were added to the team effort. In 2000-2002, students mounted project reports on the course Website, developing useful Web skills in the process. In 2003, this was replaced by a new and ambitious component of the course, an Internet-based independent group research project on relevant control topics. In previous group presentations of the design project, all groups dealt with a problem of identifying unknown industrial processes, utilising techniques learned in the course. Currently, the presentations are more varied, as they deal with the research projects covering a wide range of topics. All have to include a demonstration of either an experiment, a hardware design or an original software simulation. Most are chosen from a list based on a review of relevant literature, but student-initiated and instructor-vetted topics are also permitted. The 2003 presentations included the development of a Graphical User Interface (GUI) for a control systems laboratory, a system identification using a GUI environment, fuzzy logic motor controller design, principles of gyroscopic stabilisers in Segway Human Transporter, control aspects of Honda ASIMO humanoid robots, controls of MagLev high speed trains and environmental controls at the International Space Station.

RESULTS AND DISCUSSION

Table 1 shows the Pearson's correlation coefficients between individual rankings and the template ranking. The average correlation between individual rankings and the template ranking has been moderate (r > 0.3) to strong (r > 0.6).

Table 1: Pearson's correlation between individual assessments and template rankings.

Pearson's r:	1999	2000	2001	2002	2003
Mean	0.635	0.666	0.698	0.434	0.604
Median	0.648	0.672	0.733	0.467	0.636
Minimum	0.048	0.146	0.212	-0.167	-0.133
Maximum	0.907	0.953	0.967	0.867	0.915
STD	0.203	0.186	0.195	0.273	0.251
No. Students	19	26	33	44	57
Significant *r*	44%	69%	66%	12%	49%

Except for 2002, almost half or more correlation coefficients were statistically significant. This suggests that students exhibit an understanding of their peers' strengths and weaknesses, and mark them fairly. The sharp drop-off observed in 2002 is most likely related to the reduced in-class hours, less guidance and a resulting lower confidence in the process. Following the restored lecture hours in 2003, all coefficients increased again.

As Table 2 shows, there were no significant differences in correlations between students who performed below the class median (BM) and those who performed above the class median (AM). This suggests that students who were academically weaker in the course exhibited as good critical judgement skills as those who excelled in it. Average review marks were 88%, 86%, 77% and 83% (2000-2003), attesting to generally consistent and fair peer-assessments.

Table 2: Differences in Pearson's correlations, BM vs AM students.

	2000	2001	2002	2003
BM Mean	0.687	0.661	0.360	0.540
AM Mean	0.644	0.734	0.526	0.663
ANOVA	F=0.333 df=25 p=0.569	F=1.154 df=32 p=0.291	F=3.890 df=43 p=0.055	F=3.548 df=56 p=0.065

Table 3 shows that, except in 1999, a strong and statistically significant correlation was observed between peer-assessments of the presentations and the overall course grade. A study of chemical engineering students similarly reported statistically significant correlation (r = 0.54, p = 0.0001) between peer ratings and course grades [6].

Table 3: Pearson's correlations between peer-assessment of presentations and final course grade.

1999	2000	2001	2002	2003
r=0.203 p=0.425 n=19	r=0.941 p=0.0005 n=26	r=0.808 p=0.0005 n=33	r=0.658 p=0.0005 n=44	r=0.636 p=0.0005 n=57

The author collected anecdotal evidence suggesting increased enthusiasm and creativity resulted from the introduction of the research project in 2003. Figure 2 shows the distribution of final grades in the course, normalised with respect to the number of students. While the averages have remained consistently high since the course's inception, a significant shift towards even higher final grades occurred in 2003, supporting the evidence of increased student engagement.

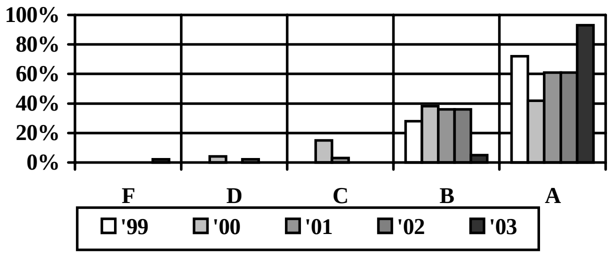

Figure 2: Final grades in ELE829.

Instructor-Course Evaluation (ICE), a standard assessment tool, is administered university-wide at the end of each semester. Figures 3 and 4 show ICE results regarding perceptions of the instructor's effectiveness in ELE829 and of the method of assessment used in the course. While consistently better than the departmental, faculty and university averages, they further improved after the course was redesigned in 1999, including the introduction of Web support for the course. This confirms findings elsewhere at Ryerson that the lecture period can be used to engage students in more participatory activities [11]. Further, it indicates that, when technology is properly integrated into the process, it effectively enhances learning [12].

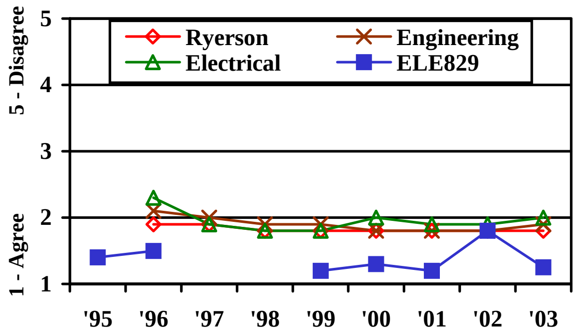

Figure 3: ICE: Faculty member was effective.

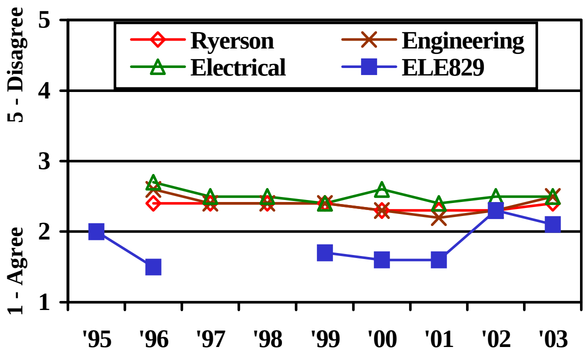

Figure 4: ICE: Course assessments provide good measure of student accomplishment.

However, a sharp spike can be observed in both evaluations, coinciding with the 2002 reduction of face-to-face contact between students and the instructor, resulting from the rapid increase in enrolment (Figure 1). While the course grades did not suffer (Figure 2), student satisfaction with the instruction (Figure 3) and their confidence in course evaluations (Table 2 and Figure 4) clearly did. All measures improved again in 2003 when the instruction hours returned to the previous levels. This suggests that personal interactions with the instructor are essential to the quality of undergraduate experience and are difficult to replace by an online environment, even if there are no significant differences in the student academic performance.

SUMMARY

Peer and self-assessment empowers students, provides them with an opportunity to reflect on what is being learned and how it is assessed and become aware of their strengths and development as a team, while the individual review mark introduced in this course encourages objectivity. A significant correlation between peer ratings and final course grades confirms the validity of peer assessments. Additional anecdotal evidence accumulates, with students describing the collaborative process as a rewarding learning experience, noting that gaining confidence in their judgement and interpersonal and communication skills acquired in the course as one of its most important benefits. Based on the statistical evidence and student feedback, the author concludes that peer assessment provides valuable insight into student performance and is a valid and reliable evaluation tool.

Progressive educational strategies implemented in this course, including the use of technology by the instructor, as well as by students, demonstrate that it is possible to meet the challenge of teaching a demanding course so that both high student performance and satisfaction are achieved. In 2003, despite continuing increases in student enrolment (to 57, from 19 in 1999) there was an observable improvement in the level of student engagement, which the author attributes to the introduction of a professionally relevant independent research project. Project presentations demonstrated a considerable creativity in demonstrations and a palpable student enthusiasm. However, the author's experiences also serve as a reminder that, while technology can act as a lever to improve the quality of undergraduate education, it should not be used as a tool to achieve efficiencies of scale in increasingly large classes.

REFERENCES

1. Catalano, G.D. and Catalano, K.C., Transformation: from teacher-centered to student-centered engineering education. *J. of Engng. Educ.*, 88, **1**, 59-64 (1999).
2. Smith, K.A. and Waller, A.A., *Afterward: New Paradigms for College Teaching*. In: Campbell, W. and Smith, K.A. (Eds), New Paradigms for College Teaching. Edina: Interactions Book Co. (1997).
3. Hargreaves, D.J., Student learning and assessment are inextricably linked. *European J. of Engng. Educ.*, 22, **4**, 401-410 (1997).
4. Bransford, J., Brophy, S. and Williams, S., When computer technologies meet the learning sciences: issues and opportunities. *J. of Applied Dev. Psychology*, 21, **1**, 59-84 (2000).
5. Felder, R.M., Felder, G.N. and Dietz, E.J., A longitudinal study of engineering student performance and retention. V. Comparisons with traditionally taught students. *J. of Engng. Educ.*, 87, **4**, 469-480 (1998).
6. Kaufman, D.B., Felder, R.M. and Fuller, H., Accounting for individual effort in cooperative learning teams. *J. of Engng. Educ.*, 89, **2**, 133-140 (2000).
7. Temple, B.K., Allan, M., Davidson, S., Silfver, R. and Neitenbach, J., The VIDEEO project: teaching cross-disciplinary subjects. *Proc. 4th UICEE Annual Conf. on Engng. Educ.*, Bangkok, Thailand, 293-296 (2001).
8. Pappas, E.C. and Hendricks, R.W., Holistic grading in science and engineering. *J. of Engng. Educ.*, 89, **4**, 403-408 (2000).
9. Pond, K. and ul-Haq, R., Learning to assess students using peer review. *Studies in Educ. Evaluation*, 23, **4**, 331-348 (1997).
10. Sullivan, M.E., Hitchcock, M.A. and Dunnington, G.L., Peer and self assessment during problem-based tutorials. *American J. of Surg.*, 177, **3**, 266-269 (1999).
11. Zywno, M.S., White, W.E and Northwood, D.O., Innovations in engineering education - Ryerson model for 21st Century. *Proc. 5th Baltic Region Seminar on Engng. Educ.*, Gdynia, Poland, 45-48 (2001).
12. Zywno, M.S., Enhancing good teaching practice in control education through hypermedia instruction and Web support. *Proc. 2002 Inter. Conf. on Engng. Educ.*, Manchester, England, UK (2002).