

An investigation into the student grading system in problem based learning

S.Venkatesan, S.Fragomeni and K.Zhang

School of Architectural, Civil and Mechanical Engineering, Victoria University, Australia.

srikanth.venkatesan@vu.edu.au (corresponding author)

Abstract — Problem Based Learning (PBL) is adopted as an alternative learning strategy in many universities across the globe. The current method of assessing student performance in PBL at Victoria University, Australia, utilises grading terminologies such as Distinction, High Distinction, Pass or Fail that follow the definitions adopted in conventional teaching approach. Note that student performance is usually well defined in the conventional teaching and evaluation methods. The correlation between the student grade (e.g. High Distinction) and the student performance (e.g top 10% student in the class) are usually higher. In PBL, the team dynamics play a very important role in the final grade received and as such a student might receive ‘High Distinction’ without sufficient demonstration of student competency or input. This paper investigates the performance of students enrolled in a PBL based engineering design subject. Data were collated based on the performance of about 100 students who participated in the Conventional Learning (CL) for weeks 1 to 6 and PBL in weeks 7 – 12 were analysed. The overall conclusion is that the present methodology of grading students in PBL setting does not distinguish individual student performance from the group. However, the methods seem to be appropriate for assessing group work. The paper concludes with recommendations for assessing and grading the student performance at a group level and at individual levels.

Key words: *Problem Based Learning (PBL), Conventional Learning (CL), Combined Assessment (CA), Victoria University (VU), student performance, student grades.*

1. INTRODUCTION

Today’s engineering industry expects graduates to possess a multitude of generic skills in addition to traditional engineering skills. Engineers Australia is also encouraging such an approach in its accreditation requirements and stage 1 competency standards [1]. Governments have also emphasised on “quality” in education and even the funding allocation is linked to the demonstration of education quality [2]. There is also an increased awareness amongst the students in their choice of universities, courses, programs etc., recognising these changing needs Victoria University (VU) in Melbourne, Australia, embraced Problem Based Learning (PBL) approach in the year 2006. Many other universities in Australia such as Central Queensland University, the University of Queensland have also embraced the PBL approach. RMIT University adopts a project based learning approach whilst the University of Swinburne includes industry based learning into its curriculum. The choice of alternative teaching strategies in place of Conventional Learning (CL) has helped reduce attrition rates in some universities not necessarily within Australia.

However, PBL in itself does not guarantee success. There is a wealth of evidence to demonstrate that unless adequate previous knowledge and guidance is provided this technique would be ineffective [3]. In addition, the measurement of success, learning outcomes, are difficult to assess [4]. There is also a wealth of evidence to demonstrate that PBL has been successful [5], even in the field of medicine [6,7,8,9]. A very wide volume of literature is available on PBL (other than the references noted). This is a clear indication of research needs to improve teaching, assessing and evaluation methods. Very little guidance is available on assessment methods for a subject that allows students with varied disciplines in engineering (e.g. Mechanical and Civil) undertaking the same subject. The authors are involved in teaching civil engineering design and project management subjects at VU. In this paper, the performance of about 100 students enrolled in a second year subject called “engineering

design” which includes concepts across civil and mechanical engineering design is analysed. Since students belonged to different disciplines in engineering, it was decided to teach fundamental concepts in both fields using a traditional lecture-tutorial based approach before the students could be invited to undertake the PBL approach. Thus weeks 1 – 6 involved teaching fundamentals of civil and mechanical engineering design. At the end of week 6 students were assessed based on a traditional exam. In week 7 students were invited to design an industrial structure that comprised of designing an over head crane (mechanical design) and a structure to house it (civil design). This problem had been chosen based on the requirements of a company interested in building an additional shed for its warehouse. An examination conducted at the end of week 6 revealed poor results – typically a “fail” grade, although the performance in portfolio assessment of student groups was far superior – atypical of “distinction” grade. (Refer Figure 1). Puzzled by these results, the lecturers (the first author of this paper and another colleague in mechanical department) choose to weight average the results of PBL assessment by 60% and the CL assessment by 40%. Thus the final grading presented a better and a fair assessment for students, not necessarily for the lecturers. (Figure 2). It can be noted that there was no scientific evidence to support our claim at that time and decided to investigate this issue later on.

Please note the grading procedures in Figure 1 & 2. It is to be noted that this procedure evaluates student performance in the conventional learning style. Thus the question of “is the student performance with a ‘high distinction’ grade in PBL the same as in CL style” arises. This is an important question, since if the grades are consistent across the two learning methods (namely PBL and CL) then the procedure of weighting them or even a comparison is meaningful. Given that PBL addresses criterion which are not necessarily the same in CL mode, it becomes imperative either to re-define the grades or at least specify the core attributes that each grade represents. Also, whether anyone agrees or not, it is clear that consultancy companies or universities would like to employ top ranking students. Traditionally the term ‘top students’ have meant students with higher grades according to the CL mode. Thus a student with a majority of ‘high distinction’ scores is preferred over a student with say ‘pass’ grades. However, the term ‘top students’ in today’s context must encompass a number of attributes that is consistent with the expectations of potential employers. This leads to another interesting question: What are the core attributes on which the student assessment has to be made? This could be an interesting debate and the answer may vary depending on the field, discipline, subject, topic etc., within the scope of this paper we decided to account for the main attributes specified by PBL and CL based on literature review. Furthermore, it would be interesting to work out the correlation of the results arrived using different assessment methods.

A very important point to be noted is that PBL assessments are evaluated on the group portfolio. i.e team dynamics play a very important role and a student may get a ‘high distinction’ grade without real input. Thus the final grade reflected could not have captured the real work. Whilst this is a well recognised problem, fortunately for us in this subject, we had asked students to design particular elements and include their original work in the portfolio. Results shown in Figure 2 are based on the group portfolio and thus do not reflect individual student learning. To eliminate this bias, we decided to remark the portfolio for each student based on his individual contribution to the report and also based on the re-review of exam papers.

In summary, the following three questions aroused:

1. Does the student performance grade in PBL and CL represent the same attributes?
2. What criteria need to be considered in the assessment of student learning in a quasi PBL setting (i.e. PBL plus CL in some combination)?
3. Is our earlier assessment of a design subject reliable?

The following sections of the paper describe the approach undertaken to evaluate student performance.

2. RESEARCH METHODOLOGY

Literature review was undertaken to identify key attributes or criteria emphasised by PBL and CL. Rubrics relating to the key criteria were also collated (e.g. rubrics for critical thinking skills [10]). It was decided to maintain the same sub-levels for each criterion (i.e four tiers) and a similar level of scale (i.e. one to four – typically of the form poor, average, good and excellent) across all the main criteria. The idea behind this normalising is that equivalent weights can be adopted for each of the criteria and scale – typically between 0 – 0.25, 0.26 to 0.5, 0.51 to 0.75 and 0.76 to 1.0 respectively for poor, average, good and excellent. In essence, broad definitions of each criterion were identified based on rubrics and weights allocated for each category. Thus student performances in terms of the grades could be expressed as shown in Equation (1):

$$P[G] = f(PBL_c, CL_c) \tag{1}$$

Where,

P is performance denoted by Grades such as High Distinction, Distinction, Credit, Pass and Fail and PBL_c and CL_c represent the key criteria in both PBL and CL methods.

Based on the above discussion, $f(PBL_c, CL_c)$ could be calculated by assigning weights (W) to each rubric score (W_i) and the performance in terms of marks and grades could be recalculated using a simple linear weighted average (LWA) as shown in Equation (2):

$$f(PBL_c, CL_c) = \left(\frac{\sum_{i=1}^n W_i}{n} \right) * 100 \quad (2)$$

Percentage of marks obtained using LWA for each student would then be compared with earlier assessments of PBL, CL and the combined assessment (CA) scores shown in Figure 2 to arrive at any meaningful conclusions.

3. LITERATURE REVIEW

Literature identifies PBL as an instructional methodology that uses real-world problems to promote critical thinking, problem solving and the acquisition of key concepts in the area of consideration [5, 11,12]. In PBL, students strive to find a solution to the problem by working in groups and the lecturer acts as the facilitator [3,5]. PBL is aimed to achieve or enhance or imbibe: critical thinking, problem solving, resourcing or research ability, communication skills, interpersonal skills, group work, team building, presentation skills, reflective learning, self-learning, student-centered learning, life-long learning, knowledge, integrity, ethics, environmental awareness, motivation, creativity etc., [5, 13] and these can be broadly classified into knowledge, skills and attributes. Based on our recent experiences in PBL and literature review and the nature of the subject taught, the following eight criteria were chosen for assessment: Critical thinking, problem solving, research capability, written communication, team player skills, professional presentation, reflective learning and peer assessment.

It is to be noted that PBL assessments can be of varied nature. A group portfolio or a report or even a conference paper, a completed model may be the final product chosen for assessing [14]. Authentic assessment can generally be categorised into performance assessment, portfolio assessment, reflection and self-assessment. Portfolio /reports might include descriptions of the problem, research undertaken, initial ideas, methodology, planning, calculations, site visits, practical experiments and so forth. The assessment made on the portfolio might be awarded to the whole group of students. In some instances, there may be a combination of tests, practical, portfolio and is usually weighted by the facilitator [3,15]. These can also be carried out at various phases of learning. For example tutors may assess student learning during mid-way to monitor progress and provide feedback to students.

The usual method of assessing the final PBL product is to award marks or grades. Given the nature of PBL assessment, rubrics have been used to describe the criteria and the level of performance [16]. Transparency and clarity is the main focus of rubrics. For this reason rubric descriptions were studied in great detail relevant to the eight criteria noted above. The number of sub-criteria and the grading scale varied amongst researchers. Therefore sub-criteria up to four levels with typical quality scales in the poor, average, good and excellent range were chosen to facilitate consistency in weighting. [10,16, 17, 18, 19]. In some cases these descriptions were re-defined with out major changes in the meaning.

In the references [1 to 15], it is also noted that traditional learning or conventional learning (CL) is the instructional methodology that has been in practice thus far, wherein subject concepts are taught in class rooms through lectures and tutorials. The lecturer is considered as the prime source of knowledge. "Knowledge" is the key word in this mode of learning and thus conceptual understanding, knowledge acquisition, retention, reproduction are expected of a student. Conceptual understanding and knowledge retention were chosen as the key criteria representing CL mode. One key feature of this CL is the exam where the students' understanding is tested within a short duration. The authors observe that this response under pressure is usually not tested in the PBL mode. This might be a key attribute to be embraced by PBL approach in future.

4. DATA ANALYSIS

Table 1 presents the complete data analysed for the paper. In the first column, a number has been assigned for each student since names have to be removed for privacy reasons. Total numbers of students enrolled were around 120. However, students absent totally or with incomplete works have been removed. Thus the final number of students considered for the analysis is 95. Out of this 50 of them are civil engineering students and 45 are mechanical engineering students. This also ensures an even distribution. Columns 2 – 11 provide information on the weights allocated for each student in each of the criterion. As noted earlier this weighting is based on re-marking the portfolio and a re-visit of the exam questions. In particular, exam questions assigned earlier were categorised into questions testing the conceptual understanding, knowledge retention and application. Thus weights for columns 5 and 6 are based on the performance in the exam. Weights awarded for other criteria were based on the individual contribution to the portfolio. (Columns 2,3,4,7,8,9,10 & 11). The work on design of components or elements carried out by each student was taken into consideration in awarding these weights. It can be noted that critical thinking, problem solving and to a certain extent written communication can also be evaluated using the performance in the exams. For this reason our judgement included a generic assessment of a student in both modes of learning. Mean of these weights were calculated in column 12. Using Equation (2), the weighted scores (LWA) were obtained in column 13. In order to facilitate comparisons, scores obtained from our earlier assessment on PBL, CL parts and the combined assessment scores (CA) are also presented. Some cells are highlighted to show a very reasonable match in the LWA assessment with that of the earlier assessment. In all, about 1/6th of the student assessments were very close. In general, the mean score estimated by LWA approach is less by about 10% in comparison with EA approach (Figure 3). This shows that our earlier combined assessment shown in Figure 2 is reasonable. A correlation based on the linear regression had been carried between LWA scores with PBL, CA and CL scores (Figure 4).

The above table also helped in identifying the performance of the whole class against each of the assessment criterion. Thus mean values for each of the criteria could be calculated as shown at the last row of the table. This now facilitates a generic comparison of the performance of each student against the whole class. In order to justify this statement, performance of one top student, one above average student and one below average student is shown in Figure 5. It can be seen that Figure 5 gives a clear indication of the level of skills in these three students across the criteria.

5. DISCUSSION

The above method of analysis and the results invoke many useful discussions. At first, the results obtained using LWA confirm that our earlier assessment using combined scores CA were reasonable. The difference is about 10% less in LWA and the grade change is between a credit to a pass. This might be due to the fact, that LWA scores are based on individual performance. One strong reason for the poor performance in exams is that students might not have performed well in the other discipline. It is well known that all civil engineers cannot be good mechanical engineers. Thus the knowledge background would have played a part in poor exam scores. When students designed elements and components pertaining to their knowledge background, their scores have been higher. This performance variation in PBL and CL also suggest that the assessments should not be restricted to strict criteria to a particular mode of learning. Perhaps including criteria across the two learning modes is a better idea as presented in this paper. Only ten criteria have been considered in this paper. For example, we have not considered self learning – a very important criterion advocated in PBL mode. Such a criterion had to be excluded because sufficient background had been provided before assigning the task. Research capabilities have been included primarily for the reason that the younger generation relies on internet search and discussion with peers. We expected that students should not only look for specific information in lecture notes and recommended text books but must extend their knowledge through good research. The criterion of peer assessment had been included since it provided input on the individual contribution in terms of the rubric. Engineers are expected to present their findings or arguments to clients and for this reason the criteria of professional presentation criteria has been considered. It is also assumed that knowledge retention is not required for PBL since it advocates life long learning. Similarly abstract understanding of the concepts need not be tested in PBL and not all of the concepts within the scope of subject taught could be tested. Conventional exam style assessment is better indicator of these criteria.

The linear weighting system adopted herein needs to be discussed. This is an approximate approach. An improved way must include a scheme for each criterion. That is one would rate critical thinking with much higher weightage as compared to say peer assessment. Also the weighting would vary across the sub-levels. Thus a more complex procedure of weighting would provide a better evaluation of student performance. It is of a

positive note that the procedure explained above can accommodate this multi-level assessment. It is quite clear that some of the criteria are inter-related (e.g. critical thinking and problem solving) – critical thinking can enhance problem solving. This inter-relationship between different criteria can be studied using techniques like fuzzy logic. If an abstract core of the criteria would be required this can be ascertained by means of testing on individual categories. However, this exercise would need to be conducted within the scope of the learning mode and subject content. For the purposes of student performance evaluation, the level of rigour is up to the lecturer or facilitator.

As shown in Figure 5, student performance can be measured against the average class performance. This provides a qualitative assessment. It is interesting to note that the top student performs about an average level in terms of written communication and is a reasonable team player. On back analysis we found that this student's first language is not English and this might be a probable reason in low written communication score. However, the performance in conceptual understanding, knowledge retention, peer assessment and reflective learning are very good. The above average student seems to match the top student in some criteria and generally stays on par with average class performance. The below average student has scores well below class average and at best manages to scrape up with average class performance. Considering the class performance it can be seen that they were good team players capable of assessing their peers. The comparison grading chart in Figure 5 makes a lot of sense now. Given the list of criteria assessed, potential employers can typically evaluate the relevance of the applicant in relation to the position applied. Figure 5 can be modified to reflect the scores of the top performer and that of the student in question. Thus the clarity between class room assessment and industrial world can be ensured. Interestingly this format is flexible. i.e. student performance can be evaluated across his group (can be meaningful for students), in each subject or across the years. It will be useful to analyse the student response for this format of assessment. It is to be noted a slightly less rigorous format of Figure 5 would already exist in practice elsewhere. One of the significant points in our analysis and in Figure 5 is that the highest weight score of 1 has been awarded to some students. This may need theoretical correctness and a better way of assessment would have allocated say 0.9 or 0.95 to a top performer. Interestingly couple of students scored well in the LWA as compared to their previous assessment. This might be due to the lower group performance despite a lone bright individual contribution. Thus analysing based on individual contribution to the group work provides a better reflection of student capabilities.

The correlations analysed between the LWA scores with that of the PBL scores, CL scores and CA scores have not resulted in higher values. This is due to PBL scores data and CL scores data displaying a large scatter. The only point of interest is to identify which of the assessment methods would better reflect student performance. As expected the CA scores ended up with better correlation coefficient.

It is of interest to note that die-hard fans of PBL might claim that the method of teaching adopted in this subject can't be categorised as PBL since in PBL the 'problem' comes first and initiates the conceptual learning. There may also be a claim that the methodology adopted is "project based learning" since the concepts are taught first and a problem provided to learn and to evaluate. Given that VU has recently adopted PBL, the structure of PBL for a subject is usually decided by the lecturer(s). Furthermore, the poor performance in exams at the end of week 6 is an indication that the students have not understood the concepts that well. Interestingly, students reported (in their reflective journal) that they better understood the equations and concepts at the end of the course. Therefore the paper has been written under the premise that the work belongs to a quasi-PBL nature or a hybrid-PBL approach.

6. CONCLUSIONS AND FURTHER WORK

The analysis presented above helped in answering the basic research questions set out in section 1. A brief summary is presented herein:

- Student performance needs to be assessed on criteria that embed Problem Based Learning and Conventional learning.
- Grading schemes such as high distinction, pass or fail need to be specified using the criteria considered.
- Grading schemes assigned using PBL must reflect individual student performance even if analysed using group work.
- The current practice of assessment may be improved by more transparency in the way of providing rubric based descriptions, clarity on the type of assessment, criteria, scales of evaluation and other measures made explicit.

The data analysed herein and other sets of data from similar subjects are planned for further work on applying a multi-level weighting scheme. Comparison grading as shown in Figure 5 would be circulated to students and other academics and their response typically to a questionnaire would be surveyed.

7. REFERENCES

- [1] EA (2004). "Australian engineering competency standards - stage 1 competency standards for professional engineers", Engineers Australia, Canberra. accessed at <http://www.engineersaustralia.org.au>
- [2] Hadgraft, R. and Holecek, D. (1995). "Towards total quality in engineering education", *Int. Journal of engineering education*, 11 (1), 8-13.
- [3] Sweller, J. Clark, R.E. and Kirschner, P.A. (2004). "Why unguided learning does not work: An Analysis of the Failure of Discovery Learning, Problem-Based Learning, Experiential Learning and Inquiry-Based Learning". http://projects.ict.usc.edu/itw/materials/clark/DiscoveryAndProblem_Based_Learning.pdf
- [4] Nield, T. (2004). "Defining, measuring and maintaining the quality of problem-based learning", *Proceedings of the Australian Universities Quality Forum 2004* <http://www.auqa.edu.au/auqf/2004/program/papers/Neild.pdf>
- [5] Hadgraft, R. G. (2004). "Integrating engineering education – key attributes of a problem-based learning environment", *Proceedings of the 2005 ASEE/AaeE 4th Global Colloquium on Engineering Education*, Sydney, Sept 2005, paper 312.
- [6] Bergmann, H., Dalrymple, KR., Wong, S. and Shuler, C.F. (2007). "Investigating the relationship between PBL process grades and content acquisition performance in a PBL dental program", *J Dent Educ.* 71(9): 1160-1170-2007
- [7] Barrows, H. S., "A taxonomy of problem-based learning methods", *Medical Education*, Vol. 20, 1986, 481-486.
- [8] Harden, R. M. (1986). "Approaches to curriculum planning", *Medical Education*, Vol. 20, pp. 458-466.
- [9] Walton, H. J. & Matthews, M. B. (1989), "Essentials of problem-based learning", ASME Medical Education Booklet No. 23, ASME Office, Level 10, Ninewells Hospital and Medical School, Dundee DD1 9SY, UK.
- [10] Facione, P.A. and Facione, N.C (1994). "Holistic critical thinking rubric", PAF49:R 4.2:062694, California academic press.
- [11] Riberio, L.R.C and Mizukami, M.G.N (2005). "Student assessment of a problem-based-learning experiment in civil engineering education", *J. prof. issues in engineering education*, 131 (1).
- [12] Williams, W. and Pender, G. (2002). "Problem based learning approach to construction management teaching", *J. prof. issues in engineering education*, 128 (1).
- [13] Ning, C.C. (1995). "Undergraduate academic programme: planning, development, implementation and evaluation", *Int. Journal of engineering education*, 11 (3), 175-184.
- [14] Murray-Harvey, R. and Slee, P. (2000) "Problem based learning in teacher education: Just the beginning!" *Proceeding of the annual conference of the Australian Association for Research in Education*, Sydney Australia, 4-6 December.
- [15] Wang, Y. (2003). "Using PBL in Teaching Analytical Chemistry." *The China Papers*, July, 2003.
- [16] Andrade, H. (2001). "The effects of instructional rubrics on learning to write", *J. current issues in education*, 4 (4).
- [17] Rubric for assessment: "Integrated problem solving model", accessed at <http://www.big6.com/showarticle.php?id=39>.
- [18] Problem based learning: university of Delaware, accessed at <http://www.udel.edu/pbl/>
- [19] Bronson, P., Ng, A. and Wong, K.K. (2007). "Design and implementation of a peer assessment tool for problem based learning in engineering", *Proc. of the 2007 AaeE conf. Melbourne, Australia*.

8. TABLES

TABLE 1. Student performance assessment – including weighted approach LWA, PBL scores, conventional learning (CL) scores and combined scores (CA).

9. LIST OF FIGURES

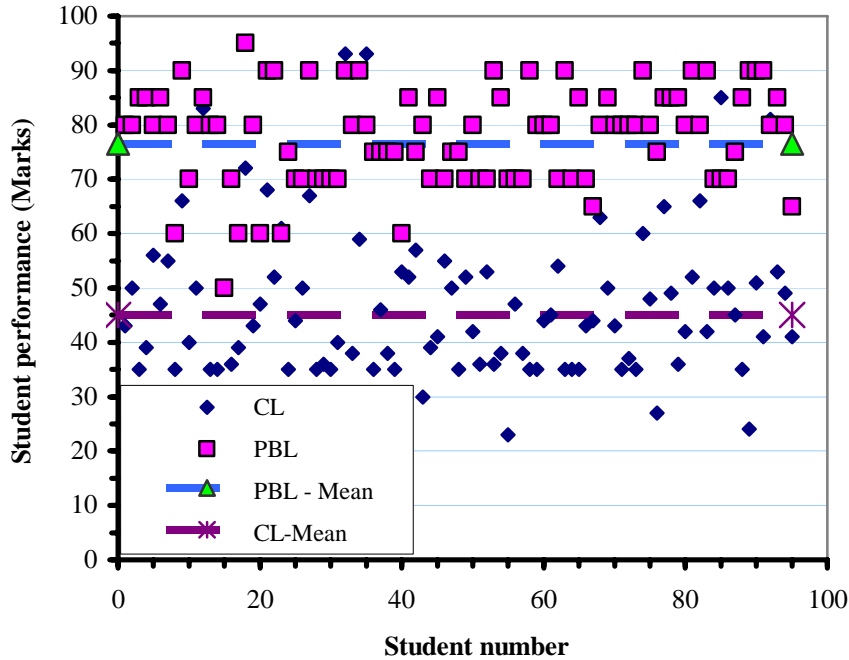


FIGURE 1. Student performance in an engineering design subject – in PBL scores and Conventional Learning (CL) scores

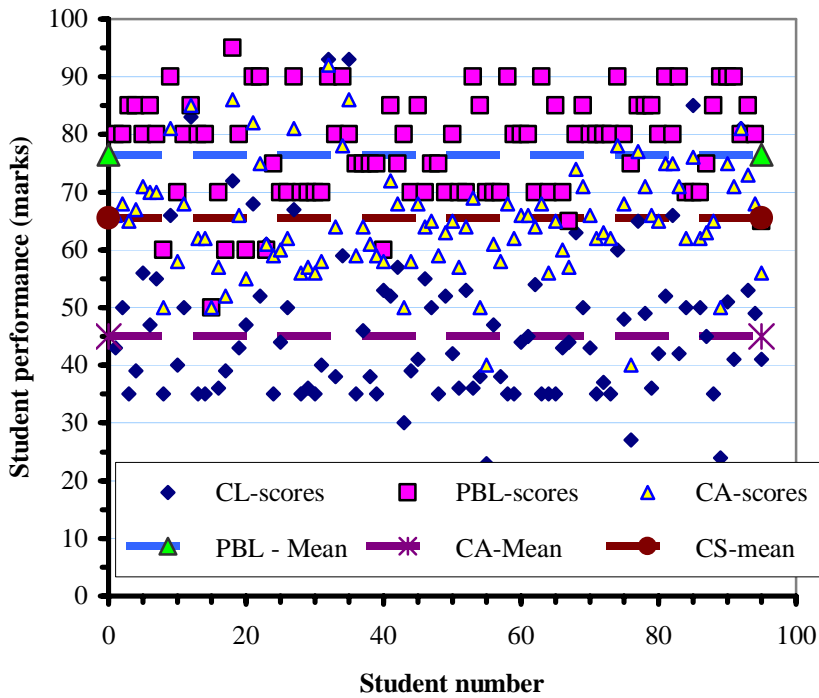


FIGURE 2. Student performance in an engineering design subject – in PBL scores, Conventional learning (CL) scores and Combined assessment scores (CA)

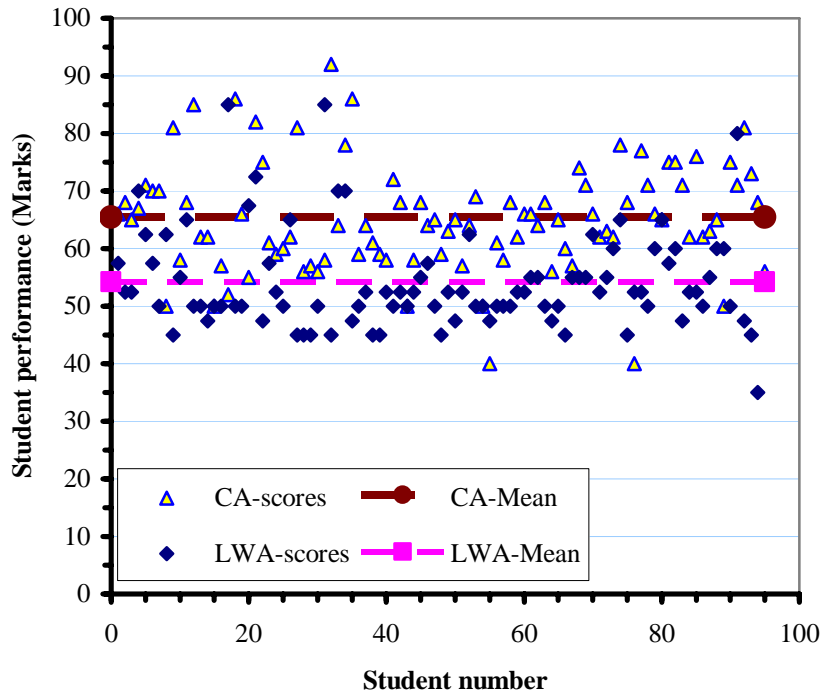


FIGURE 3. Student performance in an engineering design subject – in combined assessment scores (CA) and linear weighted average (LWA) scores

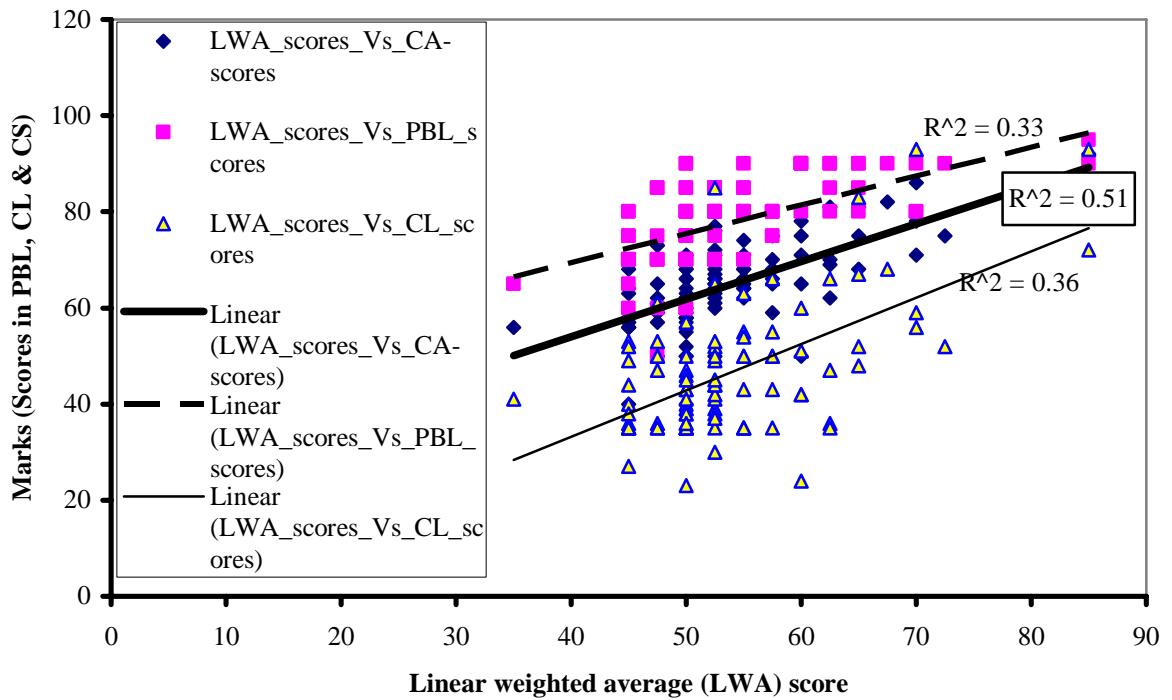


FIGURE 4. Correlation analysis between LWA scores with PBL scores, CL scores and CA scores

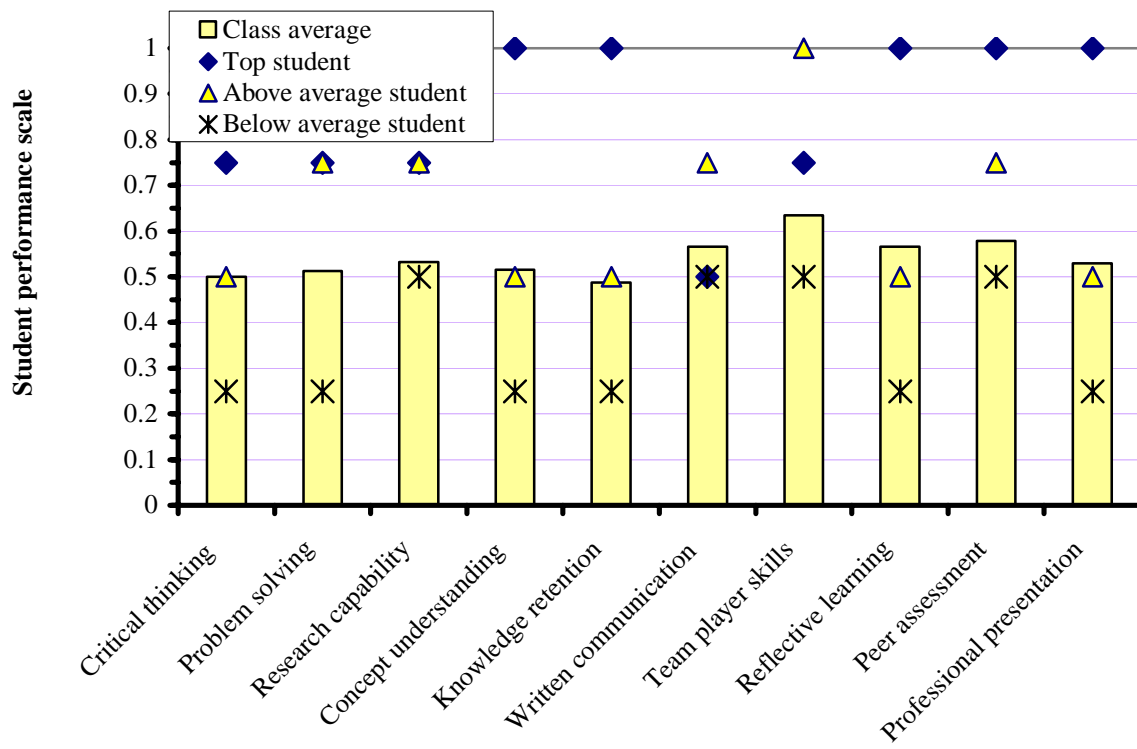


FIGURE 5. Comparison grading of student performance in a quasi-PBL setting