Teaching and assessing team work in Machine Design

Kristiaan Schreve [1]
[1]*Department of Mechanical and Mechatronic Engineering, Stellenbosch University South Africa.*
[1]kschreve@sun.ac.za

Abstract

Demonstrating successful ability to do team work is one of the required ECSA (Engineering Council of South Africa) outcomes for engineering students. In this paper, some interventions taken at Stellenbosch University to motivate third year machine design students to perform better in their team tasks are presented. In the past students gave each other a confidential buddy rating. This was a single number that was supposed to capture all aspects of team work. Four interventions were applied to both improve the students' performance in teams and also obtain a better assessment thereof. The objective of implementing this was that students will have a clear understanding of what behaviour is expected of them in the team and also that they will gain some basic tools to better function in a team. The interventions were: 1) presenting a lecture on basic aspects of team work, 2) teams were require to submit evidence of their team work, 3) a rubric was introduced to score each student's performance and 4) using an A1 or A0 sized hand sketch of the concept to facilitate communication. The paper compares the aggregate results over the past four years. The results are encouraging, but not conclusive.

Introduction

Demonstrating the ability to work in a team is one of the 10 outcomes for engineering programmes required by the Engineering Council of South Africa (ECSA). It is equally recognised world-wide by leading bodies such as the Accreditation Board for Engineering and Technology (ABET) (Seidel, 2004) and institutions such as the University of California (Delson, 2004), Stanford (Brereton, 1995), and Massachusetts Institute of Technology (Graham *et al*., 2007). The reason is very simply that engineers in practice work in teams. Also, Graham *et al*. (2007), found that team work is often an excellent environment where students can grapple with concepts, relating it to their lived experience. For mechanical engineers, Machine Design is very often used as the vehicle to teach and assess team work. The typical project usually allows a fair degree of creativity, but also often requires a systematic design approach (Graham *et al*., 2007) and demands that students apply technical knowledge that they would typically have acquired in the first years of engineering education.

Very often students will work in teams of 3 to 5 students (Delson, 2004; Graham *et al*., 2007, Raucent, 2001). Teams can be formed in various ways, e.g. randomly, based on results from previous courses, mixing up teams from previous projects, topic based, external selection or allowing them to form their own teams. Team work can be done at any time during the formal education period.

It is not unusual that students would have had little or no instruction in team work (see e.g. Leiffer *et al*., 2005). The expectation is then that students will learn how best to perform through experience. However, this approach is cautioned against. Leiffer *et al*. (2005) found the following problems:

> Lack of team cohesiveness, communication breakdown, ineffective leadership, unequal effort from members, or missing deadlines.

Assessing team work is challenging because it is difficult to determine the contribution of individual team members, determine the team's effectiveness, preventing students from getting

a free ride (Raucent, 2001) and preventing students from avoiding specific tasks e.g. CAD modelling, report writing or analysis tasks (Seidel, 2004). A range of assessment tools is used e.g. direct or video observation (Graham *et al*., 2007), confidential peer assessment (Seidel, 2004), questionnaires (Graham *et al*., 2007), oral interviews (Seidel, 2004) regular quizzes to ensure individual participation (Leiffer *et al*., 2005) and peer-review evaluation process (PREP) reports (Graham *et al*., 2007). The selection of the appropriate assessment tool will depend, amongst other things, on class size (Graham *et al*., 2007, using direct observation had an instructor to student ratio of 1:4 to 1:2) and length of the task (Raucent, 2001, recommends observation also for short problems rather than longer duration design problems). With larger classes, a tool is needed that assesses at an appropriate level of detail with a manageable workload. For this reason, some form of confidential peer evaluation is often used, possibly in combination with some of the other tools.

This paper will review Stellenbosch University's Department of Mechanical and Mechatronic Engineering's past experience with team work in a third year (6th semester) Machine Design course. Some adjustments to the teaching and assessment process were recently made and the results are compared with previous years. While the paper cannot be conclusive about the results because the data is limited and it is very difficult to isolate the parameters that contribute to the learning experience, the results are encouraging. The changes have since been implemented in several other courses.

The paper is structured as follows. In the next section, the past experience with team work is presented. The interventions taken to improve the situation is presented next followed by a section on the results. The last section presents the conclusions.

Previous Assessment of Team Work in Machine Design

During the years from 2009 to 2011, when the author was involved with the Machine Design module, students were required to design a gearbox. It was typically a forward and reverse gearbox that could switch under load, thus a clutch was also required. Students worked in groups of 3 to 5 and normally they were allowed to form their own teams. They stayed in the same team for the duration of the semester. The project spanned the full semester. In 2011 there were 33 teams with 3 instructors. The team submitted two reports. The first was submitted just past the middle of the semester and the last at the end of the semester. Assessment of the first report was mostly formative. Based on the two reports as well as the drawing pack a final mark for the team was calculated.

A confidential peer evaluation (buddy rating) was done during the final test after the completion of the project. They simply had to give a single rating for each other. They could rate each other above or below 100%, depending on their contribution. Students rated themselves as well as all team members. Based on all the ratings, a single rating was calculated for each student. Buddy ratings were normalised around 100%, thus some students could be rated more than 100%. There was a minimum threshold, 90%, above which students had to be in order to satisfy the requirements of the outcome. They were well informed of the entire process.

At this point in their engineering education, students would have bone at least two other Machine Design projects in previous modules where they were exposed to team work as well as a similar assessment strategy.

The project mark was then multiplied by the buddy rating to get the individual score. This was done to acknowledge the fact that some students may have made a better contribution to the final result than others.

The instructor moderated all the buddy ratings. If there were any anomalies or students that did not meet the threshold, these groups were called for an oral interview. At least two instructors as well as the entire group attended the interview. During the interview, the instructors would

endeavour to judge the contribution of each member by first of all finding out that everybody had a fair chance to contribute, that they were disciplined in attending team meetings and meeting deadlines, responsibilities were clearly communicated and that they delivered their milestones. Students often presented evidence of their own work in the form of design journals and sometimes minutes of meetings or copies of electronic communication.

Anomalies in the buddy ratings included a significant disagreement between team members' ratings or if one team member rated him-/herself very high.

Up to that point in their formal engineer education they would not have had any instruction on team work. Team work and project management is taught in a module in the 7th semester. Due to various reasons, this module cannot be moved to an earlier opportunity.

The team work learning experience was not exceptionally bad, but the number of teams failing each year due to team related issues was concerning. The author also felt that we owed it to the students to better prepare them for this experience.

Interventions

In order to improve the team work learning experience, four interventions were undertaken in 2012 in the same Machine Design course. Team formation and the nature of the project remained the same as previous years. The interventions were:

1. One lecture was presented on basic aspects of team work, such as holding meetings, allocating work, keeping minutes, basics of group dynamics, basics of leadership, allocating and balancing work within the team, etc. Delson (2004) also used one lecture on team work instruction. Our lecture was broadly based on the work of Miller (2011). The lectures did not include the material on team types, personalities and relationships. These topics are obviously important, but time did not allow their inclusion. The success of this is not possible to judge directly, however the impression based on interaction with the student throughout the semester was that they were better informed of what was expected of them regarding team work. This intervention will not be discussed further in this paper.
2. Teams were required to submit evidence of their team work (meeting minutes and meeting attendance sheets) as part of their project report. This was formally assessed as part of their report mark. From past experience, it was known that students lacked in the basic disciplines of keeping minutes, attending meetings, clearly communicating responsibilities and even just being aware of team activities. Team management skills are also emphasised by Delson (2004).
3. A rubric was introduced that was used to score each student's performance (Table 1). The rubric has four categories with four score levels. They were given this rubric at the beginning of the course. The rubric was taken from Nitko (2004) and essentially assesses team behaviour. It has the advantage that it formally states what good team behaviour is and reinforces the work covered in the team work lecture. It is therefore a more comprehensive assessment than the buddy rating. In 2012 the rubric was used in parallel with the buddy rating system and only to moderate the buddy rating since the instructors were not yet convinced that it will be a useful tool.
4. Rather than the first report submitted after the middle of the semester, teams were asked to submit a hand sketch of their concept on an A1 or A0 sheet. They could include any number of views and could also include any written description or calculation, as long as it fitted on the page. This was submitted after the third project week and returned one week later with written feedback. The assessment was purely formative and no mark was allocated. The drawing's purpose was twofold. Firstly to get the students working early on and give them quick feedback. Secondly to serve as a communication tool between the instructor and students as well as within the team. This intervention will not be

discussed in the rest of the paper, but was deemed a great success since all through the semester, when students had discussions with the instructors this drawing was always the focus of the discussion. It avoided them committing to the structured approach inherent in CAD too quickly and hopefully facilitating a free and creative exchange of ideas.

Table 1. Peer Evaluation Rubric. Based on Nitko (2004).

Learning Target	**Fail 0-1**	**Poor 2**	**Average 3**	**Excellent 4**
Works towards the achievement of team goals	Does not work towards team goals or actively works against them.	Communicates commitment to the team's goals, but does not carry out assigned roles.	Communicates commitment to the team goals and effectively carries out assigned roles.	Actively helps to identify team goals and works hard to meet them.
Demonstrates effective interpersonal skills	Does not participate in team interaction, even with prompting, or expresses ideas and opinions in a way that is insensitive to the feelings and knowledge base of others.	Participates in team interaction with prompting. Expresses ideas and opinions without considering the feelings and knowledge base of others.	Participates in team interaction without prompting. Expresses ideas and opinions in a way that is sensitive to the feelings and knowledge base of others.	Actively promotes effective team interaction and expresses ideas and opinions in a way that is sensitive to the feelings and knowledge base of others.
Contributes to team maintenance	Does not attempt to identify changes or modifications necessary in the team process, even when prompted, or refuses to work towards carrying out those changes.	When prompted, helps identify changes or modifications necessary in the team process, or is only minimally involved in carrying out those changes.	Helps identify changes or modifications necessary in the team process and works toward carrying out those changes.	Actively helps the team to identify changes or modifications necessary in the team process and works toward carrying out those changes.
Effectively performs a variety of roles within the team	Rejects opportunities or requests to perform the required role in the team.	Does the bare minimum for the assigned role.	Makes an attempt to perform more than one role within the team and contributes to secondary tasks.	Effectively performs multiple roles within the team.

The scores for each student in the rubric were added up to a possible maximum total of 16. This was reworked to a normalised percentage. Therefore, if all the students did not rate the same, the maximum rating would be more than 100%. Again, students scored themselves as well as all other team members. The peer evaluation and buddy rating was done at the same time and was confidential. Again, when a team member failed to achieve the threshold required to demonstrate the outcome or if there were anomalies in the ratings, the groups were called for an oral interview.

Results

The outcome was encouraging, but not conclusive. Table 2 presents an analysis of the buddy ratings of the past four years as well as the new buddy rating calculated according to Nitko's

(2004) rubric. In the three years preceding 2012 the buddy rating was calculated as described at the beginning of this paper, while in 2012 the old method was used in parallel with Nitko's (2004) rubric (the last row with 2012*).

The table uses four metrics to compare the methods of peer evaluation.

1. It frequently happens that every member in a team would rate all members 100%. This is perhaps more likely to happen in teams formed by association of friends, as is the case here. It is unlikely that a significant proportion of the class would all score 100%, thus implying that they worked equally hard and performed equally well in a team. Thus, the number of teams where all members scored 100% is reported in the second column of Table 2.
2. The percentage of teams where someone failed to achieve the outcome is reported in the third column. From an educational point of view, this number should be as low as possible.
3. A reasonable spread in the data would also be an indicator that the measure is working well, since it would reflect the expected diversity in a large class. Two measures were used to measure this. In the fourth column, the number of teams, as a percentage of the total number of teams, where a minimum and maximum rating differed by more than 20 percentage points are shown.
4. In the fifth column, for the same reason as point 3 above, the standard deviation of the ratings is reported.

It is encouraging that, in terms of teams with all members rated 100%, the new method resulted in the lowest score since it shows that students possibly gave more thought to the evaluation and considered more aspects of the contributions.

The year 2012 had the lowest failure rate for this outcome. It is hoped that it is a result of the interventions discussed here, but it should be acknowledged that many factors play a role; therefore it is not possible to draw this conclusion. The last row has two numbers here. This is to show that only 2% of the groups had an outright failure, while in 7% of the groups, at least one member rated one or more colleagues below the required performance.

The spread in the data using the new method is also better, when looking at the last two columns, than the simple buddy rating when looking at the 2012 results, but not as good as in 2009 and 2010. The reason for the relatively large spread in the data in those years is not clear.

Table 2. Analysis of Buddy Ratings.

Year	All 100%	Failures	Range > 20%	Standard Deviation
2009	42%	9%	12%	4.9%
2010	45%	10%	16%	4.7%
2011	27%	6%	6%	3.3%
2012	24%	2%	2%	3.6%
2012*	17%	2% / 7%	5%	7.6%

While this sample is small and therefore not statistically significant, the results seem to indicate that the buddy rating calculated with Nitko's (2004) method, combined with the other interventions, resulted in a better distribution of the data and fewer students failed the buddy rating evaluation. It shows that this method gives the students the opportunity to give a finer

rating. Clearly they use this opportunity. Also, the interventions resulted in the least number of problems with group interaction in the past four years.

Figure 1 shows a good correlation between the normalised rubric based peer evaluation and the buddy rating. The correlation coefficient is 0.894. This may be interpreted that the rubric does not really give a significantly better final score than the buddy rating. However, the value of the rubric is very much educational. As stated earlier, it gives the students a clear indication of what behaviour is acceptable in a team, it encourages them to think critically about the performance of the rest of the team as well as themselves and it gives the instructors a better tool to isolate problems within a team.

Two teams came to the instructors before the end of the project to discuss team work related problems. It was interesting to note in the ensuing interviews that the rubric was used to explain the problem to the instructors.

There is a clear outlier in Figure 1. This group was called for an interview. Again the rubric was used to underscore the problems the team had.

When interpreting the data in Figure 1, it should also be remembered that there was an 85% threshold to demonstrate achievement of the outcome. Students would not have rated their colleagues lower than the threshold unless they deemed their performance unsatisfactorily. (The threshold was lowered from 90% to 85% in 2012 to encourage the students to give a larger spread in the ratings.)

A further advantage of the rubric, in terms of outcomes based education, is that satisfactory performance is no longer related to a somewhat arbitrary number, but it is now criterion related. The column marked “Fail” would indicate that a student did not perform satisfactorily.

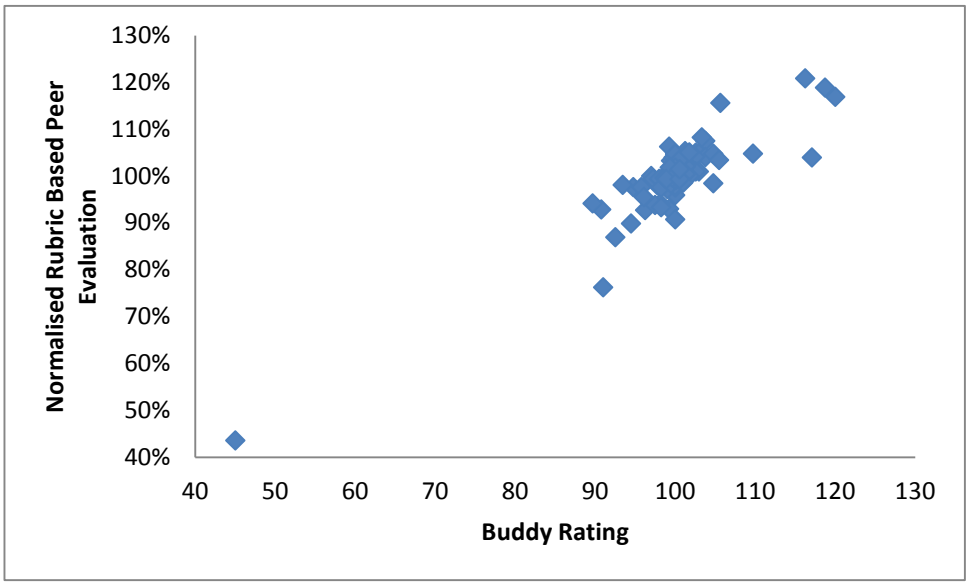

Figure 1. Correlation between Normalised Rubric Based Peer Evaluation and Buddy Rating of 2012

One of the problems with this system is the amount of effort required to capture and process the peer evaluation data. (See also Delson, 2004). There is not a lot of time to do this, because firstly it must happen during one of the busiest periods for academic staff and secondly quick feedback is needed since the outcome must be published so that teams may have an opportunity to appeal the rating. All groups with failures, anomalies and appeals must then be interviewed. This takes a lot of time and must be finished before students leave campus. In future a tool such

as the CATME website (https://engineering.purdue.edu/CATME) may solve this problem, but have not yet been tested here.

The project documentation was scored with a rubric. One item on the rubric was used to evaluate the team work documentation (second intervention). The rubric item is shown in Table 3. The results for the class on this outcome are: Average 64%, standard deviation: 18%, minimum 20% and maximum 100%. Five teams performed poorly (scores below 5). Clearly there is much room for improvement in this area. However, it is encouraging that in oral interviews, some form of team work documentation was often used to support arguments.

Table 3. Rubric Item for Assessing Team Work Documentation.

Outcome	
Team work is well documented.	
Performance	
Poor 0-4	Significant number of required documents not present. Documentation is generally incomplete
Average 5-7	All required documents are included. Documentation in general complete.
Excellent 8-10	Documentation complete. Division of work is clear. Clear minutes.

Conclusion

The results were encouraging and this system is already being used in other courses as a result of this trial. From interaction with students, it appears that the interventions had some impact on their learning experience. The rubric as such is not necessarily an indicator of individual performance in a team, but it does have educational value.

Similar systems are also being implemented in other courses in the engineering programmes at Stellenbosch. The method therefore seems to be generally suitable for assessing project based team work in engineering programmes. There is certainly much room for improvement.

One problem with the current system is that only one peer evaluation is done after the project is finished. While some teams do report problems within the team to the instructors, this is not always the case and cannot be relied upon. A better way may be to have more frequent (formative) evaluations (Leiffer *et al*., 2005). The self-assessment may be more valuable as a learning tool during such formative assessments.

A further shortcoming in this system is that the team's performance as a whole is nowhere assessed. The rubric is used to assess individual performance. The question remains, how well did the team perform. A good project report does not necessarily imply that the team performed well as a unit. This issue is also raised by Leiffer *et al*. (2005). They propose the following metrics, but give no clear method of assessing these: organisation, communication, decision making, team unity, ability to resolve conflicts, individual time spent, time spent in meetings, individual action items completed, team milestones achieved, peer evaluation, private comments to instructors.

References

Brereton, M., Sheppard, S. & Leifer, L. (1995). Students Connecting Engineering Fundamentals and Hardware Design: Observations and Implications for the Design of Curriculum and Assessment Methods, *Proceedings of Frontiers in Education Conference* (Vol. 2, pp. 4d3.1-4d3.7)

Delson, N. (2004). Creating a Positive Introductory Design Experience, *The NCIIA 8th Annual Meeting, March 18-20*, 59-65.

Graham, M., Slocum, A. & Sanchez R. M. (2007). Teaching High School Students and College Freshmen Product Development by Deterministic Design with PREP. *Journal of Mechanical Design*, 129, 677-681.

Leiffer, P. R., Graff, R. W. & Gonzalez, R. V. (2005). Five Curriculum Tools to Enhance Interdisciplinary Teamwork, *Proceedings of the American Society for Engineering Education Annual Conference & Exposition*.

Miller, D. (2011). Brilliant Teams: What to Know, Do and Say to Make a Brilliant Team, 2nd Edition, London: Prentice Hall.

Nitko, A. J. (2004). Educational Assessment of Students, 4th Edition, Upper Saddle River: Prentice Hall.

Raucent, B. (2001). Introducing Problem-based Learning in a Machine Design Curriculum: Result of an Experiment. *Journal of Engineering Design*, 12(4), 293-308.

Seidel, R. (2004). An Integrated Approach to the Teaching of Engineering Design, *International Conference on Engineering Education and Research* (pp. 423-430)

Acknowledgements

The author would like to thank his colleagues for participating in this endeavour to improve the team work learning experience. They have done the enormous task of assessing the project reports and have encouraged and supported the above mentioned initiatives.