AC 2010-1013: "SURVIVOR" MEETS SENIOR PROJECT

Glen Dudevoir, United States Air Force Academy

Andrew Laffely, United States Air Force Academy

Alan J. Mundy, United States Air Force Academy

“Survivor” Meets Senior Project

Abstract

We have all seen the formation of tribes and cliques on the latest edition of the television hit Survivor. Has Survivor mentality invaded your senior projects as well? For the last ten years or so, engineering programs nationwide have, with varying degrees of success, tried to incorporate the ABET-required outcome of “ability to function on multidisciplinary teams.”[1] While recognizing that all engineers must function on such teams in the real world, implementing them in the academic context poses substantial challenges. Nonetheless, I would argue that it is either extremely difficult or impossible to evaluate how well we have accomplished this outcome without putting students into the situation where they must actually work with members of other disciplines to accomplish a significant design. Even after such teams have been formed, the vagaries of academia, and the real world, interfere with our ability to measure student function, or dysfunction, within these teams. This paper presents a short history of efforts at the United States Air Force Academy, focusing on the projects sponsored by the Department of Electrical and Computer Engineering and the assessment techniques used to evaluate our students’ ability to function on multidisciplinary teams.

One method that we and others have used to assess the performance of students on teams is a peer rating scheme. There are a variety of approaches.[2-10] One component is usually a weighting system used to enforce reasonable grades and prevent the inevitable effect of team members all wanting to get good “grades” in the component of team performance. This is usually accomplished by limiting the average grade given to some value or renorming the student submissions to establish the desired average and variance. [2, 3] This can be effective, but if the variance is changed significantly, it may hide problems in team members working effectively together. A particularly difficult problem to deal with is the formation of sub-groups within the project team. In our experience this most commonly occurs with engineers from various disciplines forming a clique or tribe and rating the members of their tribe highly while penalizing those from other disciplines. When this occurs, the team cohesiveness and focus on the goal of completing the design successfully can be compromised, as team members try to insure that the majors in their own discipline get better grades than those in other disciplines. We have named this phenomenon the “Survivor” effect. We will discuss our approach to defeating this divisive aspect of assessing the performance of students on multidisciplinary teams.

Overview

In the early 1990s, many engineering programs accomplished the ABET requirement for design experience using a senior design project involving a single student. Students were typically enrolled in a single “Design Project” course and completion of the design and course constituted success in the required element. As ABET attention to industrial concerns increased, the importance of working as teams of engineers became apparent. Early on, these teams were limited to engineers, or more correctly students, from a single discipline or engineering major. Although students had to work together to accomplish the design goal, the scope of the design projects led to very highly integrated teams with a narrow focus and left little room for formation of sub-teams competing for team resources.

ABET's adoption of "Engineering Criteria 2000" (now just "Engineering Criteria"), introduced an entirely new dynamic.[1] The new criteria required students to demonstrate the ability to function on multidisciplinary teams. This outcome requires that students work with (or demonstrate the ability to work with) members of other disciplines on a projects where both (all) disciplines are required for success of the project. Moreover, the ABET requirement for assessment requires the program to assess the students' accomplishment of this outcome.

Assessing Individual Performance on Team

In the early days of team projects, before the requirement for (assessing) multidisciplinary teams, the common motivation for separating individual performance from team performance was to provide faculty members the ability to discriminate between the contributions of individuals to the overall success of the team, and eventually award each team member a grade indicative of both team success and individual contribution.[2-4] Unless the faculty member observed a substantial amount of the team's effort, the faculty member had to depend in part on the team members' assessment of their teammates' contributions. An important factor was that, in general, team members had similar, if not the same, backgrounds and preparation for a narrowly scoped project.[2-4] Team members were schooled in the same approaches to design and understood what their teammates were doing, even if was harder or easier than their individual part of the project.

As we evaluated individual contributions to these projects, it was most often sufficient to have team members assign scores to each other for each phase of the project and then average those grades. Initially, we had individuals assign scores to themselves as well and calculated a simple average of the grades assigned to each team member. After some analysis of scores assigned by students, we decided that some controls were in order. The initial steps were to stop students from scoring themselves and to adjust the average score for all teams to an A-. This technique worked reasonably well. It was transformed into a percentage of the team grade. Each student had 100 points to give to each of his/her teammates. The total points where capped at 100 times the number of students with you on the team. If you wanted to increase one person's score you had to decrease another's. Once the students were decoupled from having themselves in the matrix and the total number of points they could award were capped, we started to see grades that looked very much like those we might have awarded ourselves based on smaller number of observations. We did, however, notice that on some teams the student averages had a larger variance than others, and that there was a substantial difference in variances of the scores awarded by different students, sometimes even on the same team.

At the time we had not thought of the variance of the student averages being indicative of how the team was functioning as a whole, but we *were* concerned that students were incorrectly interpreting the grading rubric and awarding grades with higher variance than appropriate in some situations. Much more common however were teams where students apparently got colluded on the assignment of grades and adopted an approach that awarded the same grade to everyone. Occasionally, a strong student would not go along with this approach, but since they no longer graded themselves there was little they could do other than boost or drop other student grades.

A New Dynamic: Multidisciplinary Teams

As we attempted to comply with ABET's evolving Engineering Criteria, we felt that the only way we would be able to assess our students' ability to function on multidisciplinary teams, and hence improve our preparation of the students to do so, would be to actually see them performing in such an environment. The first obstacle in the effort was finding design problems that met two criteria: 1) The design had to be of sufficient scope that it required engineers from different disciplines in order to be successful, and 2) It had to be capable of being completed by reasonably sized teams of students within a single academic year. (Initially we only had one semester, but quickly realized that we could not routinely meet the first constraint with a single semester course.) This is a challenging (design) effort for faculty members. One early project that I supervised was a solar-powered vehicle. We had mechanical engineers working on the chassis, suspension, steering and braking subsystems and electrical engineers working on energy collection and storage, a battery driven motor, and instrumentation. The two groups had to come together at the power train, consisting of the transmission and drive components.

First efforts at integration between faculty members were shaky, let alone integration between students in different majors. Each major had its own ideas as to what the design process should look like including when to evaluate the students and on what design products. Once we finally agreed on a schedule, a set of common products and what design reviews ought to look like, we had to consider how to develop grades for individuals on the team that reflected more than just the team success. Initially, some programs were reluctant to have their students evaluated by students in other majors (from other departments), but we eventually decided to continue the same peer evaluation technique described above.

We typically scheduled peer evaluations surrounding three events each semester. In the fall, students evaluated each other at the conclusion of requirements definition (system requirements review), after completion of a preliminary design (PDR) near the middle of the fall semester and at the completion of detailed design and presentation of a critical design review (CDR) to faculty mentors near the end of the semester. In the spring semester, students accomplished peer evaluations at project status reviews (PSRs) after four and eight weeks, and at the end of the semester with coincident with the system verification review (SVR). Figure 1 shows the average of team standard deviations for each review for the '08, '09 and '10 capstone classes. The results of the first evaluations (SRR) were typical. Students were still learning each other's first names and we reluctant to insult anyone. Most of the ratings reflected little, if any, variance in student contributions. The second round of ratings did not produce substantial changes. Consistent with past results, we began to see an increase in variance of student ratings as the students became more comfortable with their teammates. At the end of the third rating period, we began to see some unusual results. The second semester ratings confirmed our suspicions that team members were devaluing the contributions of students in different majors than their own. Discussions with the students revealed that this occurred primarily because they did not understand what the students in the other majors were doing, and consequently felt they could not fairly assess the level at which they were performing. Several steps were taken for the class of 2010 to reduce this as will be discussed later in the document.

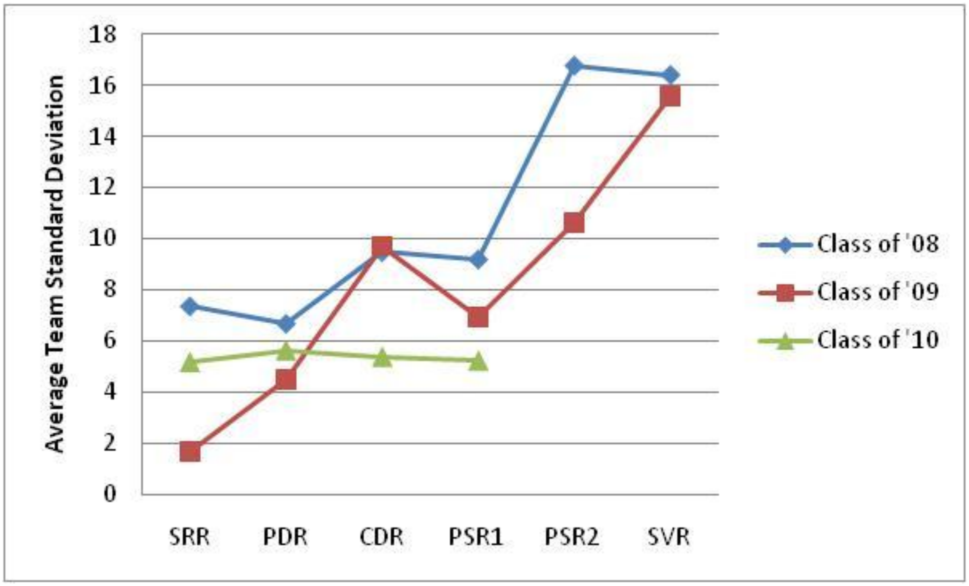

Figure 1. Average Standard Deviation of Peer Evaluation Scores per Team

Evaluating Student Performance on Multidisciplinary Teams

Eventually, we began to think about how to assess not only how well the students were doing on their own, but how well the students we performing as a part of the team, and in particular how well they were integrating their efforts with those of students from other majors. As one might expect, teams that were successfully integrating the efforts of all engineers were performing the best. Of course, differences in the difficulty of the projects also had an impact on the perceived performance of team members. In most cases, the students seemed to have difficulty taking into account the difficulty of another sub-team's tasks in evaluating the members of that sub-team who most frequently were from other engineering majors.

There were cases where a weak engineer was placed as the minority discipline on a team of strong students. In these cases the weak students where often ostracized in terms of the partition of valuable work. As a result their contribution continued to decrease.

We observed that the teams that were the most successful in integrating disciplines almost always had the lowest variance in their peer evaluations. If the group was functioning effectively, some members would have lower scores, but they would be consistent across team members from different majors. Conversely, teams that did not work well together, split up along disciplinary lines both in the tasks they took on and in how they rated each other. Team member A, an electrical engineer, might have high ratings from all the electrical engineers, but low ratings from the mechanical engineers on the same project. In general, over the '08 and '09 class years, there was a 12.8 and 5.5 point difference between the ratings given to students from the same department and those from different departments in the final peer evaluation. This put students in the minority at a disadvantage. The worst teams demonstrated even more dramatic numbers, 54, 35, and 24 points difference.

At times we observed weaker electrical engineers, feeling they had no chance to complete for high peer grades from the other EEs, abandon their EE teammates and take on tasks for which they were less qualified in the mechanical realm. Not surprisingly, they frequently improved their peer ratings from the mechanical engineers who were happy to have another set of hands to complete their parts of the project. This provided an example where the team variance went down, indicating better overall functioning as a team and better success of students participating on multidisciplinary teams, when in fact the students were switching roles for all the wrong reasons. The team fared worse because of the student decision to improve their grade at the expense of the overall team performance.

Adding a Subjective Component

Based on these results, we decided to add another component to the peer evaluations and another component to the grade awarded by course faculty and project mentors. In addition to asking the students to grade each other, we asked them to provide support for the grades they awarded which would be provided to the other students anonymously. We did this by asking them to provide for each other member of the team two strengths and one area for potential improvement. If they awarded a grade lower than B (82) to a fellow student they were encouraged to provide additional areas for improvement. We also asked them to consider areas beyond technical competence. For course faculty and mentors, we asked them to provide a grade at each review period to describe how the student team members got along with other members of the group as a whole and whether or not they contributed to the success of the team in the best manner possible. Finally, the students were instructed that the usefulness of the feedback they provided to other students would be considered in determining their grades under this evaluation. The idea was to make them think more constructively in evaluating their peers.

These changes seemed to make some substantial improvements in the peer evaluations. Still, while most students could easily come up with good justification for high grades, they had difficulty finding useful things to say about ways to improve. Even when they gave lower grades, their inability to come up with tangible suggestions usually led to lower variances (better low-end grades) than previously observed.

This year the power of the faculty was yet again increased to combat excessive peer evaluation grades. Students were not permitted to give teammates more than +/- 20 points without first justifying it to the course instructor. Thus far, this technique has eliminated excessive point awards and peer evaluation gaming

Residual Anomalies

The method that we have adopted for evaluating student accomplishment of the multidisciplinary team member outcome works well under many team conditions. There are, however, some teams that pose problems. The problem is not that the assessment method does not effectively evaluate how well team members are working together. Quite the contrary, when the variance of the peer evaluation scores is high, it faithfully identifies teams where the students have adopted counterproductive strategies for the team. The problem is that the students have adopted a personal objective to maximize their own grades at the cost of the team grade, and likely the team performance. These problems rarely crop up on teams that are doing well. They have most commonly been found on teams that are performing poorly from the start. As a result, the overall team evaluations are below the expectations of some members of the team. They are frequently experiencing failure, or poor grades, for the first time in their academic careers.

It starts with one team member who is unhappy with their grades, both from faculty and other team members, early in the project. They are typically quite savvy regarding interactions with other team members, and in spite of the fact that they have not been performing at a high level, still want to receive high grades. They typically see the team failure as a result of the students in the “other” majors, who are not carrying their share of the load. These students might approach the other members of the team in their own major and form a “tribe”. They agree, at the behest of one of their members to create a bimodal distribution of peer grades. They will only give good grades to the other members of their tribe. This prevents any normalization measure short of changing the variance from reducing the effect of improving their grade. Initially, the other “tribes” are taken unawares by the collusion of the members of the first tribe. Inevitably, though, they figure out what has happened and retaliate with their own collusion.

Faculty members and mentors who are watching carefully can attempt to stem such behavior and confront the students early in the process. It is particularly troublesome if the faculty mentors in the offending student’s discipline back up the student, although that is frequently the initial tendency. The requirement to back up low grades with constructive comments on areas needing improvement and the authority of the faculty member to take the student’s use of the rating scheme counterproductively in determining that student’s grade for contributions to the team are effective deterrents to this behavior in most cases. In the worst case, we have been forced to tell the offending students that their peer evaluations have been discarded because of insufficient evidence to support their claims regarding students in other majors.

In the worst of these cases, we observed one particularly egregious offender switching tribes at one point, switching back, and then recruiting other members of competing tribes. Needless to say, this particular student became a poster child for how not to succeed on multidisciplinary team. To say the least, it will be interesting to follow his progress in future.

Summary

There are institutions that still do not require students to actually perform on multidisciplinary teams. That is understandable for programs with hundreds of students in each year, but we believe that it is difficult to assess the accomplishment of this outcome without watching the students do it. It is too easy to discuss differences between disciplines, ask a few questions on an exam, and based on those results say that students will perform effectively in a real, multidisciplinary environment.

Peer evaluations provide useful insights into the functioning of any team. When correctly used to determine a student’s individual grade on a team project, they provide motivation to the student to be a good team member and carry their share of the load. This technique also encourages the team to fairly distribute requirements among team members and assign tasks based on team member abilities and preparation. The variance an individual team member's scores and the variance in individual grades provides an effective measure of how well the team is functioning.

Multidisciplinary teams are harder to evaluate, and peer scores must be scrutinized for anomalies. If students give higher priority to maximizing their grade than they do to maximizing team success, there are a variety of strategies that they can pursue to inflate their own grade at the expense of other students' grades and, more importantly, the team's overall success. With appropriate attention, all but the most determined students can be deterred from this behavior. The resulting peer evaluations provide an effective way of assessing how well students have developed the ability to function on multidisciplinary teams.

Bibliography:

1. Criteria for Accrediting Engineering Programs. Published by The Accreditation Board for Engineering and Technology (ABET), Baltimore, Maryland. Last accessed on January 5, 2005; http://www.abet.org/Linked Documents-UPDATE/Criteria and PP/E001 09-10EAC Criteria 12-01-08.pdf (criteria approved November 1, 2008)
2. Brown, R.W., "Autorating: getting individual marks from team marks and enhancing teamwork," *Frontiers in Education Conference, 1995. Proceedings., 1995* , vol.2, no., pp.3c2.15-3c2.18 vol.2, 1-4 Nov 1995.
3. Kaufman, D.B., R.M. Felder, and H. Fuller, (1999) Peer ratings in cooperative learning teams. In proc. *ASEE Annual Conference*. ASEE, Charlotte, June 1999.
4. Ohland, M.W. and R.A. Layton, (2000) Comparing the Reliability of Two Peer Evaluation Instruments. In proc. *ASEE Annual Conference*. ASEE, St. Louis, June 2000.
5. Layton, R.A., and M.W. Ohland, "Peer Ratings Revisited: Focus on Teamwork, Not Ability," Proceedings of the American Society of Engineering Education Annual Conference, Albuquerque, NM, June 2001, session 2230.
6. Ohland, M.W., M.L. Loughry, R.L. Carter, and A.G. Yuhasz, "Designing a Peer Evaluation Instrument that is Simple, Reliable, and Valid" Proc. Amer. Soc. Eng. Ed., Salt Lake City, Utah, June 2004
7. Ohland, M.W., M.L. Loughry, R.L. Carter, L.F. Bullard, R.M. Felder, C.J. Finelli, R.A. Layton, and D.G.
8. Schmucker, "Developing a Peer Evaluation Instrument that is Simple, Reliable, and Valid," Proc. Amer. Soc. Eng. Ed., Portland, Oregon, June 2005.
9. Ohland, M.W., H. R. Pomeranz, and Harlan W. Feinstein, "The Comprehensive Assessment of Team Member Effectiveness: A New Peer Evaluation Instrument," Proceedings of the American Society of Engineering Education Annual Conference, Chicago, IL, June 2006.
10. Kaufman, D.B., R.M. Felder, and H. Fuller, "Accounting for Individual Effort in Cooperative Learning Teams," *J. Engr. Education*, 89(2), 133–140 (2000).