International Journal of Engineering Education Vol. 34, No. 2(A), pp. 414–429, 2018
Printed in Great Britain
0949-149X/91 $3.00+0.00

Profile-Based Team Organization in Multi-University Capstone Engineering Design Teams*

BRETT STONE and STEVEN E. GORRELL
435P CTB, Brigham Young University, Provo, UT 84602, USA. E-mail: sgorrell@byu.edu

MICHAEL RICHEY
P.O. Box 3707, Mail Code: 0U-23, The Boeing Company, Seattle, WA 98124, USA. E-mail: Michael.c.richey@boeing.com

Engineers are increasingly required to work in teams that span time zones and cultures. Engineering education has begun to attempt to prepare students for this environment. An important topic is how to best form teams of students for distributed design projects. The goals of this research were to validate a method for organizing teams to maximize team performance and identify and validate metrics for individuals that would help in the organization of distributed teams of student engineers. A review of previous, related research is provided. A description of the proposed method of team organization is given, along with methods of data collection and a comparison of the proposed method to common methods. How students were divided into teams and each team's method of organizing sub-teams are described. Use of online surveys, tests, statistical methods, and other data gathering and analysis methods are explained. Statistical analysis of survey results and qualitative results of interviews and observations suggest that a profile-based method for organizing teams results in significantly higher team satisfaction. Some methods for measuring and/or predicting individual attributes related to teamwork, such as the significance of participation in team sports, were validated. No correlation was found between which university a student attended and a student's level of satisfaction with his or her team. Team success in distributed, multi-disciplinary student design teams can be improved by gathering information about team members and using a profile-based method to organize team members into sub-teams and leadership positions.

Keywords: virtual teams; multi-university capstone; distance learning; team performance

1. Introduction

The Boeing Company sponsors an annual multi-university design capstone course named "Aerospace Partners for the Advancement of Collaborative Engineering", or AerosPACE [1, 2]. This course involves several universities throughout the United States working together for two semesters to design, build, and fly an Unmanned Aerial Vehicle (UAV) with a specific mission. During the 2013–2014 academic year, three teams each sought to help farmers increase crop yields by designing a UAV that farmers could use to monitor the health of their fields. The universities involved were Brigham Young University (BYU), Purdue University, the Georgia Institute of Technology (GT), and Embry Riddle Aeronautical University, Prescott, Arizona campus (ERAU). All 36 undergraduate and graduate students participating in the two-semester long course were studying either mechanical or aeronautical engineering. Students applied to participate in the course, which was counted as their senior design, or capstone credit in most cases.

Engineering teams in industry are challenged by the fact that workers are increasingly asked to collaborate with remotely located teammates [3]. For example, in the automotive industry, one group of experts estimated that 70 percent of their time is spent on activities related to collaborating with remotely located suppliers [4]. Boeing has experienced this challenge recently with the development of the 787 Dreamliner. About 65 percent of the new aircraft is built by non-Boeing suppliers, requiring significant inter-site collaboration [5].

Students participating in AerosPACE program were faced with a similar situation when they were placed on one of three teams with members spread across various universities. A generic diagram of how students were distributed is shown in Fig. 1. As can be seen on hypothetical "Team 1" in the diagram, half of the students on each team were located at one university (University "B", for Team 1) and the rest of the team-members worked from at least two other universities. These two groups became known to the researchers as the "core" team members, and the "non-core" team members, referring to whether or not the person was a student from the university with the most students on the team. The reason program organizers chose to have cores was to facilitate the manufacturing portion of the project by having a larger number of people physically present in one location to work together on assembling the design. Each team then subdivided itself into Integrated Product Teams, or

* Accepted 3 November 2017.

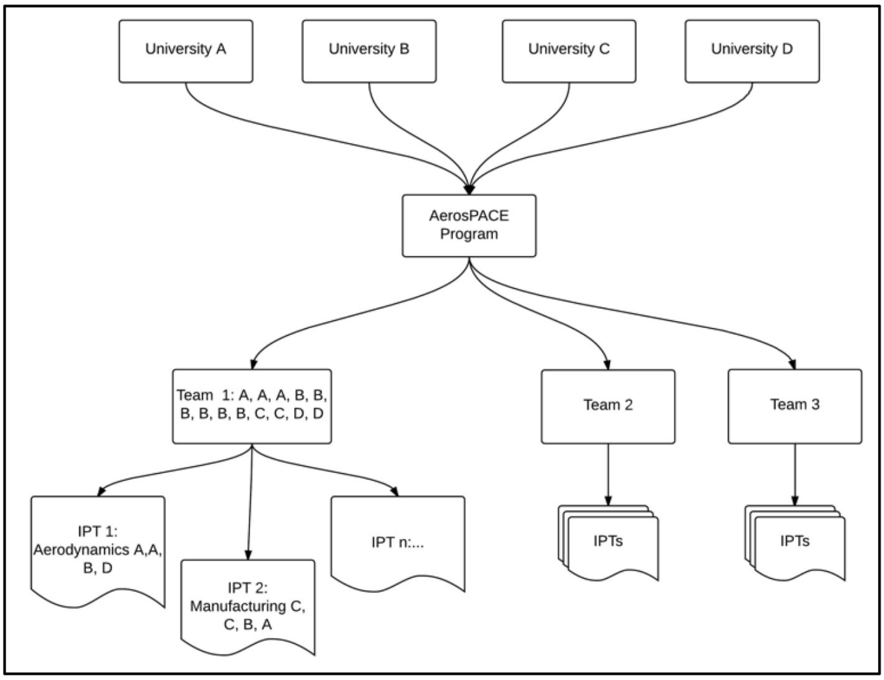

Fig. 1. Students from various universities were distributed to each team. Each team then created sub-teams, or IPTs to work on specific portions of the project.

IPTs. These IPTs focused on specific portions of the work, such as aerodynamics, manufacturing, or computer aided design (CAD).

This research focuses on the question organizers of the program faced when attempting to decide what the best method for organizing students into teams and sub-teams would be: with each student's individual skills, interests, and logistical considerations, what method can a geographically dispersed group use to optimally allocate its resources and maximize performance?

2. Background

Multiple methods for organizing teams exist. Ad-hoc methods, such as allowing students to self-select their teams, or administrators randomly assigning teams have been common, even in military settings [6]. However, as pointed out by Layton et al., these methods often lead to sub-optimal results [7].

Especially in the context of a geographically distributed, virtual team, these mostly ad-hoc methods suffer from various shortcomings. As Hackman points out, the composition of a group, such as an IPT, is the most important condition affecting the amount of knowledge and skill the group can apply to the task [8]. As well, various authors have promoted the importance of teams developing a "Shared Mental Model" of the project they are working on together [9–11]. Moreland et al., explain further that it is not only important for a team to have a shared mental model of the project they are working on, but also to have a shared mental model of the team itself and the teammates who make up the team [10]. Citing extensive evidence, they state that despite the difficulties of gathering the data needed, when a group knows who knows what and who is good at what, the organization can more optimally allocate its most important resources: its people. It is easy to see how in a virtual team, whose members are geographically far apart, building these sorts of mental models is at least as important and even more difficult.

A thought experiment, along the lines of that suggested by Moreland et al., elucidates this point further. Imagine a new team whose leader knows very little about the members of the team, or at least about certain members of the team, such as would likely be the case of a virtual design team. The potential shortcomings of organizing this team using traditional, ad-hoc sub-team organization methods could include the following, broken down by whether the shortcoming originates with those volunteering for positions or with the team leader (see Tables 1 and 2).

As can be seen by examining the ad-hoc team organization method, many potential problems have to do with a lack of knowledge regarding members of the team and low levels of trust among team members, demonstrating how the ad-hoc method can make forming a shared mental model of the team more difficult. Kramer and Tyler substantiate this idea in their work on trust in organizations [12]. They discuss how groups

Table 1. Lists some potential problems that may occur with ad-hoc team formation due to volunteers

Volunteers	
Problems	**Possible Causes**
Not volunteering for a position they are truly motivated to pursue.	Nervous or uncomfortable about volunteering for position in front of peers. Not given sufficient time to understand position, consider opportunity, or weigh options (as in a group meeting when a leader asks for volunteers before listing and describing all positions).
Not volunteering for a position they are truly qualified to hold.	Not being sufficiently aware of the responsibilities of the position – the ad-hoc description offered may be insufficient or unclear.
Volunteering for a position they are NOT truly motivated to pursue.	Desire to fit in with group, be seen favorably by others, be seen as a contributor. Misunderstanding the responsibilities and requirements of the position.
Volunteering for a position they are NOT truly qualified to hold.	Desire to fit in with group, be seen favorably by others. Misunderstanding the responsibilities and requirements of the position. Attempting to gain prominence by taking advantage of the fact that others do not realize they lack certain qualifications.
Difficulty accepting each other's roles (ex: "Why is *he* in that position?" or "How is *she* qualified to do that?" or "How is *he* more qualified than me?").	Lack of knowledge of qualifications of other individuals.

Table 2. Lists some potential problems that may occur with ad-hoc team formation due to team leaders

Team Leader	
Problems	**Possible Causes**
Not assigning a position or task to a team member who is truly motivated to pursue it. Assigning a position or task to a team member with no interest in or motivation for that position or task.	Lack of awareness of the interests, goals, or desires of the team member(s).
Assigning unqualified team members to a task.	Lack of awareness of team member skills or abilities, perhaps from not taking sufficient time to or inability to measure and consider options (as in a group meeting when a leader asks for volunteers).
"Lopsided Trust": assigning tasks only to those whom the leader already knows and trusts.	Being much more aware of the skill levels, interests, and desires of certain team members than of others (such as when they are from the leader's home organization or department). Only those who communicate most frequently or emphatically get their information heard and acknowledged by the leader.

whose members don't have time or resources provided to help teams get to know each other's qualifications and interests tend to rely on importing expectations about broad groups of people based on past experience or stereotypes. Depending on this type of information does not provide the quality of data needed to build a reliable shared mental model of a team. Woolley et al., have shown that teams with members whose skills are complementary perform better than teams with incongruent or homogenous skill-sets [13].

2.1 Measuring success

Team organization methods that are effective will help teams to be more successful. In order to measure how successful teams of students were in the 2013–2014 AerosPACE program, "success" had to be defined and a measurement method had to be identified. MacMillan et al., measured the success of teams in their experiments by examining if teams completed assigned tasks and by assigning subject-matter experts to observe and evaluate team behavior [9]. Brannick et al., describe measuring a team's success by one or two measures: "process" or "outcome" [14]. Process, according to the authors, is concerned more with interpersonal elements of teamwork while outcome has more to do with whether or not the team actually accomplished the goal or goals they set out to accomplish. Levi and Hackman argue that there are three ways to measure team success: completion of the task, the satisfaction of team members, and the learning or improvement of individuals on the team [8, 15].

To explain why the team's satisfaction is important as a measure of team success, Levi gives the following example with firefighters:

> "Obviously, completing the task or putting out the fire is an important criterion of success. However, it is also important that the crews maintain a good working relationship and the crew members do not get injured in the process. Extinguishing the fire is important, but so is preserving the ability of the team to fight future fires."

Lin et al., found that team performance ("putting out the fire") is positively correlated with team satisfaction [16]. Hertel et al., in his work on methods of characterizing virtual teams and individuals, suggests studying satisfaction ratings of team members [17]. With these sources as guiding precedents, we selected team satisfaction as the primary measure of success for this study. We recognize that there are other important ways of measuring team success and will briefly mention aspects of teams' performance related to the research.

Since teams are composed of individuals, we needed a uniform method of measuring the characteristics of individuals. Research into what areas to measure and how to measure them led to work by Dyer et al., In their respected work on team building, they propose that individual team-member motivation, or commitment to the team's goals, and having the right social and technical skills, lay the foundation for a team's success. Leadership is also cited as a crucial component of a successful team [18]. We add what perhaps Dyer et al., had taken as given—that logistical considerations, such as location should also be included when deciding who to put on a team. These areas are what we will refer to as the "fundamental areas". We attempted to measure each individual's:

- Motivation
- Technical Skill
- Social Skill
- Leadership Ability
- Logistical Considerations

These broad categories effectively encompass many important sub-areas. For example, how skilled a given candidate for a team is in Finite Element Analysis (FEA) would fall under the fundamental area of Technical Skill. How good a candidate's interpersonal communication skills are would fall under the fundamental area of Social Skill. Whether a person lives in Delaware or India and what his/her security clearance is would be Logistical Considerations. While there may be some slight overlap among these areas, in general they have proved very effective in delineating personnel.

There are many ways these areas could be measured, such as by Naikar et al.'s suggestions [19]:

- Asking an individual how s/he would rate or describe him/herself in a given area.
- Asking an individual's peer, manager, or subordinate to rate him/her in a given area.
- Testing an individual using some form of pre-validated test.
- Recording an individual's use of some sort of tool, such as a Computer Aided Design (CAD) program.
- Registering information from outside sources, such as university degrees, training certifications.

Certain of these methods of measurement are only available or useful at certain points in time for an organization. In the case of the AerosPACE program, when the three teams were initially being formed, asking individuals to rate their peers would not have been very helpful since most of them did not know each other, even if they were from the same university. Thus, in this research, gathering information about individuals for the purpose of forming teams and IPTs was limited to methods such as self-reporting, use of pre-validated tests, and registering information from outside sources. Peer evaluation did take place during the project, but this data was not used to help form teams.

One method of testing individuals which we employed was a shortened version of the Purdue Visualization of Rotations Test, originally developed by Guay [20]. The test, which has a strong reputation as a reliable instrument, has been shown to be an effective gauge for predicting student abilities in areas such as learning and using CAD software [21–23]. To reduce the survey load on students, we created a modified, shorter version of the test to administer to AerosPACE course participants called the Modified Purdue Visualization of Rotations (MPVR) test.

2.2 *Related work*

Another project called the "Hyperion UAV: An International Collaboration" involved students from universities around the world in designing and building a UAV [24]. Their project attempted to use a "follow the sun" work-flow to design, build, and fly a UAV. The Follow the Sun work-flow involves three different work locations, approximately evenly geographically spaced around the globe such that each can work an eight-hour shift, and at the end of the shift, pass the work off to the next location. As one location leaves work to go home for the night, the sun is rising and the workday just beginning in the next zone, allowing, ideally, work to continue uninterrupted.

Their research highlighted the importance of communication and common tools among the different students and universities involved and demonstrated the need for more information regarding skills and attributes of individual team members to be available to those forming design teams than simply knowing the year in university and major of each student. They explained that the need for more information when organizing teams is even greater for virtual teams of students.

Work by Kaufman et al., validated two of the primary methods used to measure individuals in this research—self and peer ratings. The researchers compared self and peer-ratings of university students in chemical engineering courses to each other and to the grades students received in the course. Their research shows that, despite faculty concerns that students would inflate their self-ratings, students tended to under-rate themselves compared to their peers. Significant positive correlation was also found between peer-ratings and course grades [25].

While many researchers have investigated methods for designing more effective, complementary teams [19, 26, 27], few have investigated how to best design geographically dispersed student engineering design teams. Researchers such as Suchan and Hayzak, state that the process of selecting team members for virtual teams is critical to team success [28]. They also state that being able to successfully identify whether or not candidates for virtual teams have traits such as sufficient levels of social skill, personal motivation, and leadership for such teams is a particular challenge for management.

3. Research objectives

Given the importance of team composition in influencing team performance, we wished to investigate its effects on virtual teams of engineering design students using the AerosPACE program. During the organization of the course, the question was posed, "How should individual students be allocated to each team, and how should each team organize its sub-teams, or IPTs?" We decided to investigate several items related to virtual team organization. The major hypothesis of this study was:

1. Teams organized using profile-based team formation methods will be more successful in at least one method of measuring success than teams utilizing more traditional organization methods, such as ad-hoc or hierarchical methods.

Several research questions were also investigated that were related to this main hypothesis. These included:

1. Will students from different universities rate their levels of satisfaction with their team differently?
2. How will "core" students differ in their satisfaction with their teams compared to non-core students?
3. What correlation between involvement in previous activities and the average peer ratings students receive from each other in the "fundamental areas"?
4. Will students who score higher on the MPVR also be ranked higher by their peers in the Technical Skill fundamental area?

4. Methods

For the quantitative portion of the data analysis, it was necessary to perform comparisons of multiple means. Since many statisticians disagree regarding exactly which test to use when comparing multiple (more than two) means, two methods of statistical analysis were used: Fisher's Protected Least Significant Difference (LSD) method, and Tukey's Honestly Significant Difference (HSD) method [29, 30]. By using both these methods, it is believed that a more complete view of the results can be obtained. When comparing only two means, unprotected Fisher's LSD t-tests were used. Simple linear regressions were also used to fit data.

Results of t-tests are reported in a format similar to Center [31]. A generalized example of how statistical results are presented is:

> "A Fisher's LSD t-test was performed with a 95 percent confidence interval. Group One's mean was X, Group Two's mean was Y, and Group Three's mean was Z. The difference between Group One and Group Two was shown to be significant, with a p-value of 0.04."

Each student participating in the course agreed via Institutional Review Board (IRB) consent form to be a research subject and complete various surveys and interviews. The primary method of this research was via online survey. In person, or web-conference personal interviews and in-person observation were used in addition to the online surveys.

Students completed various surveys at different times during the two-semester long project as part of the research.

Table 3 gives the basic timeline and descriptions of the surveys used.

The surveys used many scale based questions such as, "Think of the team that you are part of. How satisfied are you with your team? Very Dissatisfied, Dissatisfied, Neutral, Satisfied, Very Satisfied" or the example shown in Fig. 2, which asks students to rate their CAD skills.

Other questions required multiple choice/single response, multiple choice / multiple response, or text

Table 3. Outline of surveys used for this research

Survey #	Survey Name	When Administered	Description / Notes
1	Initial Survey	Beginning of Fall Semester	Recorded demographic information, self-reported interests and skill levels in the fundamental areas.
2	Modified Purdue Visualization of Rotations Test (MPVR)	Beginning of Fall Semester	A shortened, slightly modified version of the Purdue Visualization of Rotations test [20] was given to students as part of the assessment of Technical Skill.
3	IPT Survey	Early Fall Semester	Team 2 only—used to organize Team 2's IPTs with profile-based methods (see "Team 2 IPT Organization Method" subsection below).
4	Team Evaluation 1	Middle of Fall Semester	Asked students to rate satisfaction with team.
5	Peer Evaluation 1	Middle of Fall Semester	Asked students to rate teammates in fundamental areas.
6	Team and Peer Evaluation 2	End of Fall Semester	Asked students to rate satisfaction with team and to rate teammates in fundamental areas.
7	Team and Peer Evaluation 3	Middle of Winter Semester	Asked students to rate satisfaction with team and to rate teammates in fundamental areas.
8	Exit Survey/Team and Peer Evaluation 4	End of Winter Semester	Asked students to rate satisfaction with team and to rate teammates in fundamental areas and for feedback on course.

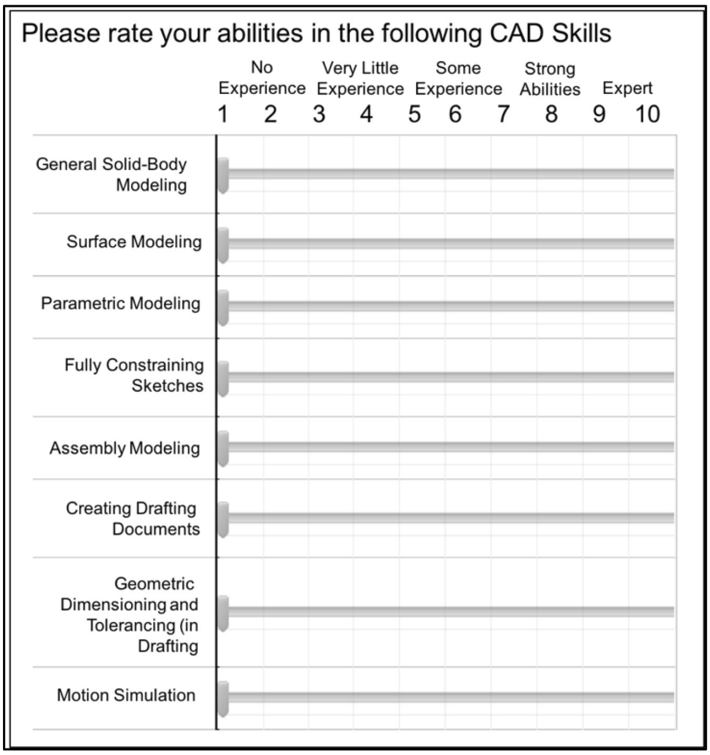

Fig. 2. An example of one question from the Initial Survey, which used a Likert-like scale to ask respondents to describe their own CAD skill level (part of the "Technical Skill" fundamental area).

response. The surveys accomplished multiple purposes. First, the Initial Survey and the MPVR allowed us to create a preliminary profile of each student according to the fundamental areas: motivation, technical skill, social skill, and leadership ability. Technical skill was, necessarily, sub-divided into various categories such as CAD, CFD, FEA, manufacturing, as well as a "general" category. Survey items that contributed to student scores in each area included items such as:

- Motivation:
  - Self-rated interest in various topics related to the course, such as aircraft design, structural design and analysis, manufacturing, materials.
  - Self-rated interest in improving skills in topics related to the course.
  - Selection of items that influenced student to apply to participate in course. (If a student selected an item such as "required for graduation" no addition was made to the student's motivation score, while selecting, "It sounded challenging" added to their score).
  - Self-rated motivation to do well in course.
- Technical Skill—CAD:
  - Score on MPVR test.
  - Self-rated CAD skill in areas such as parametric modeling, assembly modeling, Geometric dimensioning and tolerancing.
- Technical Skill—Computational Fluid Dynamics (CFD):
  - Self-rated skill in:
    - Meshing and grid generation.
    - External Flow.
    - Post-processing/Visualization.
  - Self-rated overall familiarity.
- Technical Skill—Manufacturing:
  - Self-rated experience levels in areas such as metals manufacturing, plastics manufacturing, woods manufacturing, computer aided machining.
- Social Skill:
  - Preference for working in teams.
  - Self-rating of skills such as listening, resolving conflict, tact, trustworthiness, general communication.
- Leadership:
  - Preference for acting in leadership positions in groups.
  - Self-declared rating of how respondent felt *others* would rate his/her leadership abilities.
  - Self-reported experience in leadership positions on clubs, teams, or other groups.

Each item that added to a student's score in a given area was totaled, giving a score for each area, including a "general" technical score that included all sub-areas, and can be seen in Fig. 3. Using this information, we organized the three teams so that each team had similar levels of total skill in each fundamental area and could thus be reasonably compared later. Some factors turned out to be more constraining than others. In particular, students with significant skill in using CFD tools were rare. Thus, it was necessary to make a concerted effort to equally distribute students with this skill. Fig. 3 shows the results of our efforts to evenly distribute these different attributes across the teams. Each team also had similar access to resources such as computer labs, manufacturing equipment, and coaching from experienced professors and Boeing personnel.

After completing this distribution, each team was formed and had to decide how to organize its members into sub-teams, or IPTs to work on specific portions of the project. While each team had slightly different IPTs, most shared a similar set, including aerodynamics, manufacturing, control systems, weight and balance and others.

One team, Team 2, used the profile method described in this paper to organize its IPTs. The

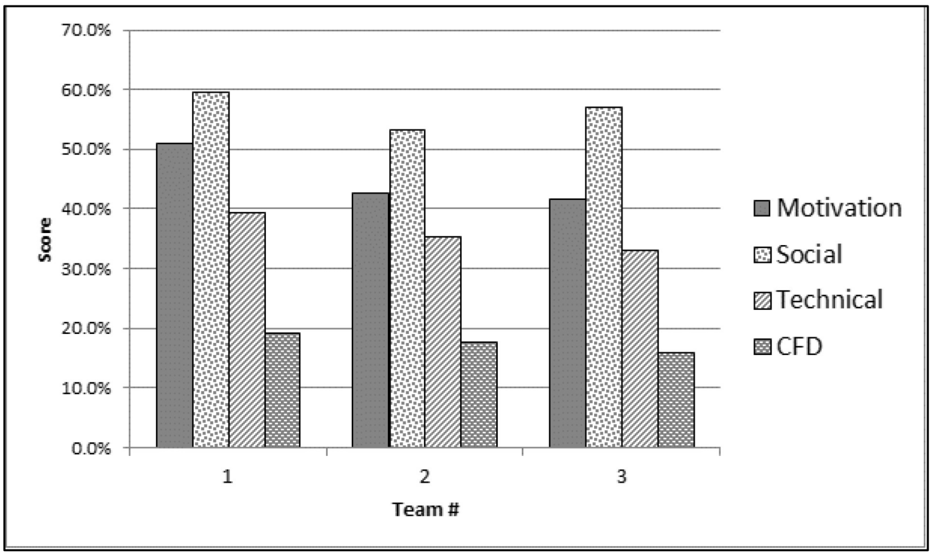

Fig. 3. Teams were organized so that each had similar levels of skill in the fundamental areas. As well, Computational Fluid Dynamics (CFD) was found to be a particularly rare, important technical skill, so we attempted to ensure that all teams had a sufficient level of that expertise.

other two teams used more traditional methods to organize their IPTs. Professors and students on Team 1 chose to use a mostly hierarchical structure, with graduate students at the top of the hierarchy selecting how to organize the team based on their judgement. Students and professors on Team 3 chose to use an extemporized method to organize its IPTs, usually by simply taking volunteers to work on each IPT as the work came up during the project.

4.1 IPT organization method

To attempt to avoid the potential pit-falls of ad-hoc teams and validate the hypothesis, we worked with Team 2's coach and student team leader to organize the team's sub-teams or IPTs by creating a profile of each team member using the following process, which is similar to Sauer and Arce's suggested method [26]:

1. Gather data from student team leaders and team coach about the tasks the team will perform, what IPTs would be created, and what each IPT would be assigned to do.
2. Gather information about each team member's interest level and skill level in each of the IPTs via an online survey (the IPT Survey), enabling a more detailed view of each team member's motivation and technical skills in specific areas. The IPT survey included a section for each IPT which contained the following descriptions and questions:
   (a) A brief description of the type of work performed by the IPT. For example, "The Aerodynamics IPT will work on the aerodynamic design of the aircraft including the wing profiles and surfaces, culminating with high fidelity Computational Fluid Dynamics (CFD) analysis of the design."
   (b) Self-rating of interest in belonging to the IPT to determine motivation for the given IPT. Clarification was made that previous experience or skill in the area was irrelevant for this question.
   (c) Self-report of previous experience with the topic (classes taken, grades received, projects, internships) to determine technical skill in the topic.
   (d) Self-rating of interest in being the IPT lead.
3. Process and present the information gathered through the survey to advise the student team lead of Team 2 (under the direction of the faculty coach) how to assign team members to each IPT by considering, in order, each of the following:
   (a) *Motivation:* To be considered for a position on an IPT, a student first had to express interest in being part of a given IPT (for example, structures or manufacturing).
   (b) *Technical Skill: S*tudents who were interested in being part of an IPT were next compared based on relevant skills and training (experience on similar projects, related courses and grades).
      – *Educational Clause:* Given that the project was part of an educational course, students who expressed strong levels of interest in a given IPT but may not have had extensive experience were considered for a position on the IPT.
   (c) *Leadership Identification:* Students who expressed interest in being the lead for each IPT were identified.
   (d) *Logistical Balancing:* After identifying students who were qualified for the different IPTs and leadership positions, consideration had to be given to having the right number of students on each IPT and ensuring that each student was involved in neither too few nor too many IPTs. It was also necessary to spread the responsibilities of leading each IPT among the students for both logistical and educational purposes. As well, the geographic location of physical items had to be considered. For example, it was necessary to ensure that at least some students on the Manufacturing IPT were at a university with the required equipment and capability.
4. Recommendations were made to the student team lead and faculty coach of Team 2.

To measure success according to how satisfied team members were with their teams, each participant was invited to complete four surveys (surveys 4, 6, 7, and 8 in Table 3) to rate their satisfaction with their team on a 1–5 Likert scale where 1 = Very Dissatisfied, 2 = Dissatisfied, 3 = Neutral, 4 = Satisfied, and 5 = Very Satisfied. Similarly, peer evaluations of teammates in the fundamental areas were performed in surveys 5, 6, 7, and 8. A web-based, Likert-scale method was also used by Ohland et al., in their related research on using peer evaluations in student teams [32]. It could be asked why we did not simply use the CATME team formation and peer evaluation tools. One reason is that at the time we were not fully aware of the system. As well, the types of questions available in the peer evaluation portion of the CATME tool are limited and did not directly match the items we wished to consider.

5. Results

Results are presented here in order of the hypotheses and research questions, starting with the main hypothesis.

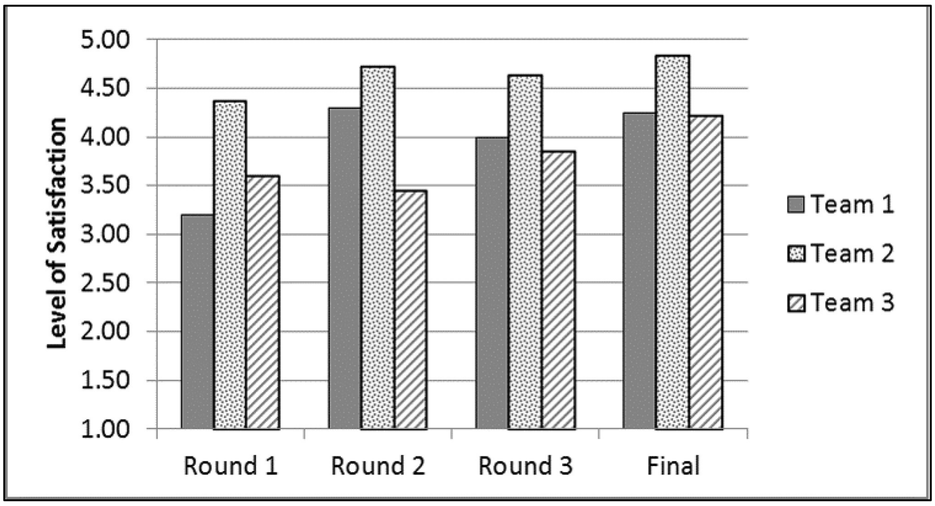

Fig. 4. Throughout the course, Team 2's average satisfaction rating was higher than the other teams.

Table 4. p-values when each team is compared to each other team using Fisher Protected LSD each pair t-tests

Team	Team	p-Value
2	3	< 0.01
2	1	0.01
1	3	0.49

5.1 Hypothesis 1

The first and main hypothesis was that teams organized using a profile-based team formation method would be more successful in at least one method of measuring success than other teams. Initial results seemed to indicate that Team 2, which was organized using a profile-based team formation method, had a higher satisfaction rating than Teams 1 or 3 (see Fig. 4 and Table 4). Fig. 4 shows the average level of satisfaction of each team throughout the course. As can be seen, Team 2's average satisfaction was consistently higher than the other two teams. When averaging all ratings from all surveys, the average satisfaction rating (out of 5) for each team was: Team 1 = 3.96, Team 2 = 4.63, Team 3 = 3.86.

Observing Fig. 4, it also can be seen that, generally, students' satisfaction with their teams increased over time. Each team did, however, have some period in the year during which their satisfaction decreased. For Team 3, the period between the first and second rounds (surveys four and six from Table 3) seems to have been particularly difficult, while for the other two teams, their ratings decreased a little later, in the third round. This may represent the teams' progression through Tuckman's classic "forming, storming, norming, performing" process [33]. Future research would be necessary to confirm this possibility. In any case, Team 2's levels of satisfactions were consistently higher than the other two teams.

Statistical analysis confirms that Team 2's higher average ratings were significant. To perform a Fisher's Protected LSD multiple comparison test, first, two outliers, which negatively affected the normality of the sample distribution and can be seen in Fig. 5, were removed. This allowed equal variance to be assumed. The new mean values for each team were: Team 1 = 4.14, Team 2 = 4.63, Team 3 = 4.01.

Then, an analysis of variance (ANOVA) with an alpha level of 0.05 rejected the null hypothesis that all teams had the same mean value ($p = 0.03$) and established the multiple comparison as "protected" [29]. Next, an Each Pair Student's t-test with an alpha level of 0.05 indicated that Team 2's advantages over Team 1 and Team 3 were statistically significant. The p-values for the comparisons between each pair (alpha = 0.05) can be seen in Table 4.

As can be seen in Table 4, the p-value related to the difference between Team 2 and Team 3's average team satisfaction ratings as well as Team 2 and Team 1 is much less than 0.05, indicating that the difference is statistically significant. A similar result can be seen with the difference between Team 2 and Team 1. However, the p-value for the difference between Team 1 and Team 3 is much greater than 0.05, indicating that the difference between those two teams' ratings is very possibly due to random chance. The more stringent All Pairs Tukey-Kramer HSD test (alpha = 0.05) also confirms the significance of the difference between Team 2 and the other teams (see Table 5).

Removal of the two outlier points merits further explanation. Besides its negative effect on normality, the point on Team 1 also represents a student who was the only student on that team from that student's university. In an interview, this student explained that being the only student from his

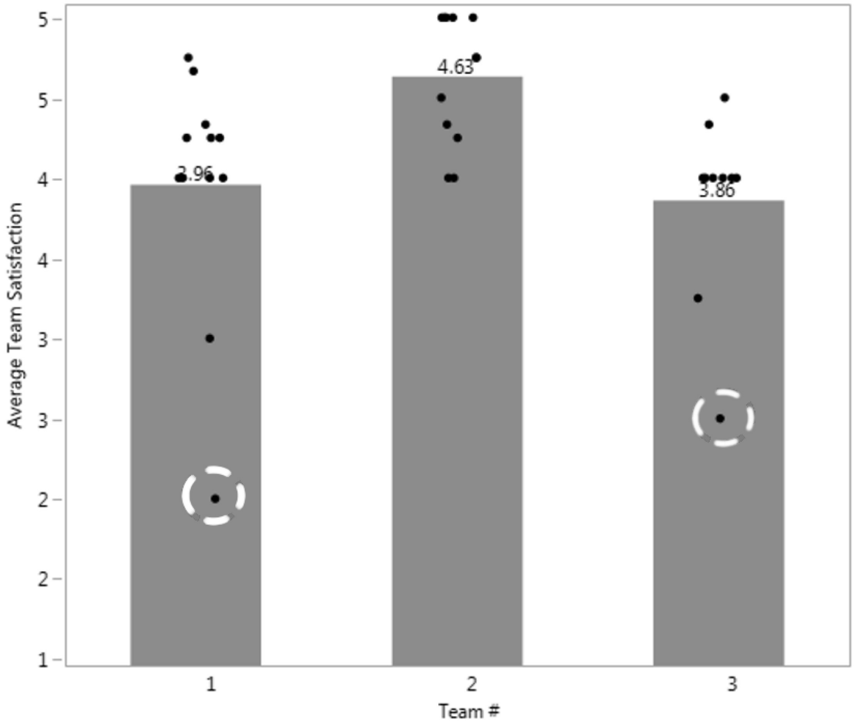

Fig. 5. Average team satisfaction for each team show with bars, and individual average team satisfaction shown with points. Two outliers which were removed are shown in dashed circles.

Table 5. p-values when each team is compared to each other using the more stringent Tukey-Kramer HSD t-test

Team	Team	p-Value
2	3	0.01
2	1	0.02
1	3	0.77

school was the reason for giving consistently low ratings. Having one team member who works alone from a remote location may or may not represent a common situation in virtual teams, but it was the only such situation in AerosPACE, and thus could be considered an anomaly.

The point on Team 3 was removed solely because of its effect on the normality of the data, but it should be noted that including either of these points in a statistical analysis only increases the significance of the differences between Team 2 and the other teams. These data points were excluded from analysis of data regarding research questions one and two as well for similar reasons.

5.2 Research question 1: satisfaction by university

Other ways of categorizing students and their potential correlation with varying levels of satisfaction were also investigated, including by what school the students were from. Research question one asked if students from one university would give different ratings than students from other universities. An ANOVA test with an alpha level of 0.05 returned a p-value of 0.79. This suggests that no significant difference exists in satisfaction levels among students from the various universities. Means for each group were: BYU = 4.35, Embry Riddle = 4.12, Georgia Tech = 4.36, Purdue = 4.23. To further demonstrate the fact that no significant difference could be identified between how any two universities rated their team satisfaction, a Fisher LSD multiple comparison was performed. Each pair of schools' p-values can be seen in Table 6. No p-value reached below 0.38, indicating again that which university students participated seemed to have very little to do with how they rated their level of satisfaction with their team.

Table 6. Showing the p-values for each school compared to each other school from the Fisher LSD each pairs t-test for illustration

University A	University B	p-Value
BYU	Purdue	0.61
BYU	Embry-Riddle	0.40
Georgia Tech	Purdue	0.59
Georgia Tech	Embry-Riddle	0.38
Georgia Tech	BYU	0.96
Purdue	Embry-Riddle	0.69

5.3 Research question 2: core vs non-core satisfaction

Research question two asked if students who were core members of their team would have higher levels of team satisfaction than students who were not core members of their team. According to the survey responses, a positive correlation exists between being a core team member and higher levels of team satisfaction. The mean satisfaction score for core students was 4.43, while the mean score for non-core students was 4.11. The difference was statistically significant ($p < 0.05$). If the points marked as outliers earlier are included, the variance between the groups becomes unequal, necessitating a slightly different t-test, but still results in a significant difference in the means (Mean Core = 4.43, Mean Non-core = 3.88, p = 0.03).

5.4 Research question 3: peer ratings and involvement in other activities

This question asked if there would be some sort of correlation between activities students had previously been involved in and the ratings they received from their peers in the fundamental areas. While most comparisons yielded no significant correlation, one set of notable correlations came from investigating student's participation in team sports during high school and college, and the average ratings they received from their peers in the areas of Social Skill and Motivation.

For the purposes of this study, a "team sport" was defined as a sport in which the team has at least moderate interdependence, as explained by Feltz et al. (2008). For example, sports like baseball and football have moderate levels of interdependence, while basketball or soccer would be considered to have high levels of interdependence. Meanwhile, track teams and swimming teams, although arguably "team" sports have lower levels of interdependence. To further explain: if a hurdler loses a race, it normally doesn't affect her teammate who throws shot-put the same way a setter missing the ball affects the outside hitter on a volleyball team.

A t-test with a 95 percent confidence interval shows that a significant difference exists between the Social Skill ratings received by students who reported having participated during high school or college in at least one team sport (15 students) and those who had not (12 students). The mean for those who had NOT participated in at least one team sport was 3.87, while the mean for those who HAD participated in at least one team sport was 4.17, and the p-value was 0.03.

If participation in one team sport is positively correlated with how one's peers rated one's Social Skill, the natural next question to ask would seem to be, "Does participating in *more* than one team sport increase the effect?" A t-test with a 95 percent confidence interval was performed to compare those who had participated in more than one team sport to those who had participated in one or none. Although a positive difference in average peer Social Skill rating does exist for those who had participated in more than one team sport (Mean for More Than One Team Sport "Yes" = 4.12, Mean for More Than One Team Sport "No" = 4.00), the difference was not statistically significant, with a p-value of 0.40.

How participation in team sports affected students' ratings in other fundamental areas was investigated. A t-test with a 95 percent confidence interval shows that although on average, students who had played at least one team sport were rated higher by their peers in the fundamental area of Motivation (Mean for those who Had NOT Participated in at least one team sport = 4.19, Mean for those who HAD participated in at least one team sport = 4.35), the difference was not statistically significant, with a p-value = 0.32. No correlation could be identified between team sport participation and the other fundamental areas.

Another interesting correlation was found between the total number of activities students had participated in and the average Social Skill ranking they received from their peers. "Activities" included team sports, non-team sports, and participation in organizations such as band, clubs, or other organized groups students had participated in during either high school or college. A linear fit of Social Skill Ratings and the total number of activities a student had been involved in resulted in a positive correlation,

$$\begin{aligned}&\textit{Average Peer Social Rating} = 3.946 + \\&0.017 * \textit{Total Number of Activities}\end{aligned} \tag{1}$$

with an $R^2 = 0.03$ and a p-value of 0.41. The variable coefficient (0.017), R^2 value, and p-value cast serious doubt on the significance of the correlation. However, after removing the only outlier from the data set, the fit became

$$\begin{aligned}&\textit{Average Peer Social Rating} = 3.860 + \\&0.043 * \textit{Total Number of Activities}\end{aligned} \tag{2}$$

with an $R^2 = 0.23$ and a p-value of 0.01. Fig.6 shows the relative improvement in the curve fit after removing the outlier. Although an R^2 value of 0.23 may seem low, as explained by [35], when studying human interaction, it is relatively rare to find R^2 values above 0.25 or 0.30.

5.5 Research question 4: MPVR score and peer ratings

Research question four asked if students who scored higher on the MPVR would be rated more highly by their peers in the Technical Skill fundamental area. Students did tend to rate peers who scored higher on the MPVR higher in Technical Skill. A linear fit of

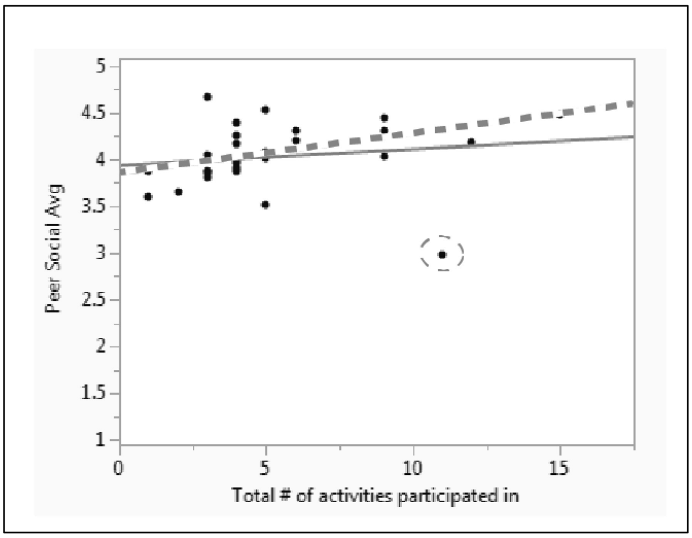

Fig. 6. After removing the outlier (circled) the original correlation (solid line) became more positive, a better fit (dashed line), and statistically significant (p-value < 0.05).

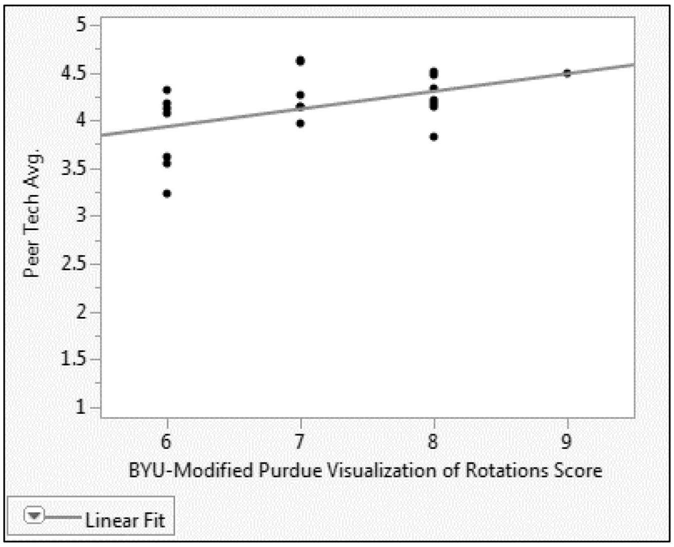

Fig. 7. A significant positive relationship was found between student scores on the Modified Purdue Visualization of Rotations Test (MPVR) and their average Peer Rated Technical Skill. Note that no students scored lower than 6/10.

scores from the MPVR to Technical Skill ratings gave a positive correlation,

$$Average\ Peer\ Technical\ Rating = 2.84 + 0.19 * MPVR\ score \quad (3)$$

with an $R^2 = 0.24$ and a p-value of 0.02. Fig.7 shows the graph of the linear fit.

6. Discussion

A discussion of the results of the study is presented here in order of the hypotheses and research questions.

6.1 Hypothesis 1

These results support the idea that the profile-based team formation method helped Team 2 achieve higher levels of success than teams that organized their IPTs using more common methods. Students who were part of the team that used a profile-based team formation method to organize their IPTs ranked their satisfaction with their team higher than students on teams who used either hierarchical or ad-hoc IPT organization methods. The difference between Team 2 and the other teams was statistically significant according to both a Protected Fisher's LSD and Tukey-Kramer HSD.

Another method of measuring success, besides the satisfaction of team members with their team, is technical. In this case, technical success could be defined by whether the UAV designed and built by the team flew and met qualifications in the time allotted. We do not attempt a detailed analysis of this measure of success in this research, but it is worth noting that of the three teams, Team 2 was the first team to fly their final UAV. These results support the idea that the time and effort needed to gather the data and go through the process of organizing a team with a profile-based method are worth it.

Examining Fig. 4, it is worth noting that while Team 2 maintained a higher average satisfaction rating throughout the project, the differences between Team 2 and the other teams were greatest at the beginning, with Team 1 and 3's last ratings being approximately equal to Team 2's first rating. One could speculate that using a profile-based method of IPT formation helped Team 2 to form a shared mental model of the team and the work to be done more quickly than the other teams, or in some other way jump-start their process of improving their team satisfaction.

6.2 Research questions

6.2.1 Research question 1: Satisfaction by University

The results indicated that no significant difference could be found in satisfaction ratings between students from any given pair of universities involved in AerosPACE course. Since what university students were from did not influence how they rated their satisfaction, a substantial potential confounding factor has been neutralized.

Another interesting observation has to do with the fact that the majority of students in AerosPACE course were aeronautical engineering majors, and a minority (all from BYU) were mechanical engineering majors. Implied in the fact that no difference between universities could be found is also the fact that no difference between majors could be found. This finding should also be verified by more research.

These results are important because they eliminate two potential sources of measurement variance. Educators and others charged with organizing and coaching distributed engineering teams can more confidently predict the reason for students' levels of satisfaction with their team.

6.2.2 Research question 2: Core vs non-core satisfaction

Having teams with core groups of students may have been desirable for manufacturing purposes, but it was suspected that it also may have hurt team success and detracted from having a true distributed team experience. Observations by faculty and researchers agreed that being core or non-core seemed to affect how "integral" to the team students felt and contributed to team disagreements.

In an interview with one student who was a non-core member of his team, the student talked about how the nature of having a majority of teammates at one location contributed to feeling less satisfied with the team as a whole. Since the schedules of the core students did not match well with the schedules of many non-core students, the core students began meeting on their own and making decisions that affected the entire team. The result was that some of those who were unable to attend felt estranged.

The most extreme case was the student who was the only teammate from his university. In interviews with this student, he indicated that, especially at the beginning of the project, he found it very difficult to coordinate and feel like he was a part of the team. He also felt a lot of pressure to represent his university well, since he was the only representative of his school. That pressure, he explained, was not all bad, since he felt it helped him perform at a higher level. In his opinion, that was, however, the only benefit to being alone on his team. While his teammates at other universities were able to easily do things like check each other's work and confirm meeting times, he had to coordinate everything via email or other electronic means, adding to his communication overhead. His favorite communication tools to overcome his challenge quickly became web-conferencing tools like Skype and Web-Ex, which allowed him and whoever he was communicating with to not only see and hear each other, but also share each other's screens. Even after discovering these tools, he said being the lone team member was still very difficult.

In conclusion, core and non-core team members from our sample reported statistically significantly different experiences as part of a virtual engineering design team depending on whether or not they were part of the core group. This should be born in mind when organizing similar programs in the future.

6.2.3 Research question 3: Peer ratings and involvement in other activities

The statistically significant positive correlations between participating in at least one team sport and the ratings students received from their peers in Social Skill aligns with findings by other researchers. For example, Artinger and Barcelona, in separate studies demonstrate the correlation between participation in sports and effectiveness in other activities [36, 37]. Artinger et al., surveyed more than 300 university students involved in recreational athletic programs. They found that on a 1 (disagree) to 5 (agree) scale, the average rating for the statement "Participation in recreational sports improves my ability to work within a team" (reverse coded to address social desirability response) was 4.14/5.00. Barcelona showed that, among the college students he studied, participation in sports programs was a significant predictor of gains in team functioning in areas outside of sports.

The difference in Social Skill ratings between students who had participated in more than one team sport compared to students who had only participated in one or none was not statistically significant. It seems as though the benefit, with regards to social skill improvement, may level off after participating in one highly interdependent team sport. Acknowledging that the results of this study can only be directly correlated with the sample of this study and that correlation is not causation, this result may indicate a "threshold" for the effect participating in team sports has on the Social Skill levels of engineering design team members. Perhaps this result could imply that the benefit to Social Skill for engineering students of participating in a team sport is getting to know how to play as a team member, and that playing multiple team sports is redundant to that end. Knowing that playing a team sport correlates positively with Social Skill could make it possible for team coaches or leaders to improve their team's performance, either by searching out team members who have played a team sport, or by encouraging current team members to participate in a team sport.

6.2.4 Research question 4: MPVR score and peer ratings

Students who scored higher on the MPVR were statistically more likely to be rated higher by their peers in Technical Skill. This implies that it may be possible to predict, based on the results of the MPVR or a similar test, how technically adept a team member will be, or at least how technically adept they will be perceived by their peers to be. Being able to more accurately predict an individual's Technical Skill would be valuable for engineering design teams by reducing the time and effort needed to form teams of individuals with complementary skills.

6.3 General observations

Ross emphasizes that, whenever possible, an in-person kick-off meeting should be held to start a project for a virtual team [38]. We found that many students requested a kick-off meeting when asked what changes could improve AerosPACE course program in the future as part of the final survey. This could, we believe, help to improve the initial levels of team satisfaction (see Fig. 3) by improving teammates' opportunities to quickly form a robust shared mental model of the team.

While this investigation studied teams of engineering students, we feel that the lessons learned can be applied to teams of engineers working in industry. The AerosPACE program was intended to mimic many of the situations engineering teams in industry now face, such as working with colleagues at different locations, with different schedules, backgrounds, and ideas, all while attempting to deliver a functional, technically complex product (relative to the skill of the team members).

While these findings focus on teams of students drawn from multiple universities, they should also be generally applicable to design teams consisting of students from the same university. This assertion is based on the fact that in this study, which university a student was from had no discernible impact on his/her team satisfaction rating, and that many universities and engineering programs are large enough that many students will not be familiar with each other before participating in a capstone or similar project. These observations would seem to suggest that using a profile-based method of team formation in a single-university setting would produce similar results.

7. Conclusions

By using a profile-based method of organizing a team of engineering design students for a multi-university design-build-fly project, team success, in at least one method of measurement, increases in a statistically significant manner compared to other methods of team organization (hierarchical and ad-hoc). Various methods for predicting some of those team member skills, such as Technical Skill and Social Skill were explored and validated. It was found that students who had participated in at least one team sport were rated higher by their peers in Social Skill than those who had not. Students who scored higher on the MPVR were also shown to be rated more highly in Technical Skill than their peers. A potential confounding factor related to determining what influences a team member's level of satisfaction with his/her team was neutralized; it was shown that what university a student is from had no statistical relation to that student's satisfaction with his or her team.

Acknowledgements—The authors would like to thank the Boeing Company, the National Science Foundation, Dr. Sheli Sillito, Dr. Greg Jensen, Dr. Ed Red, and David French for their support and assistance with this research.

References

1. L. Cannon, C. C. Cunningham, A. L. Inouye, B. Stone and F. Zender, Looking back: A Student Review and History of AerosPACE—a Multi-Looking back: A Student Review and History of AerosPACE—a Multi-University, Multi-Disciplinary, Distributed, Industry-University Capstone Project, in *122nd ASEE Annual Conference & Exposition*, 2015.
2. S. Gorrell, C. G. Jensen, B. Stone, E. Red, M. Richey, F. Zender, M. Wright, D. French, S. Hayashibara, C. Johnson and S. John, Aerospace Partners for the Advancement of Collaborative Engineering, in *121st ASEE Annual Conference & Exposition*, 2014.
3. Worldatwork, *Survey on Workplace Flexibility 2013*, 2013.
4. A. May and C. Carter, A case study of virtual team working in the European automotive industry, *Int. J. Ind. Ergon.*, **27**(3), 2001, pp. 171–186.
5. K. Peterson, *A Wing and a Prayer: Outsourcing at Boeing*, *Reuters*, no. January, 2011, pp. 1–12.
6. J. MacMillan and M. Paley, Designing the best team for the task: Optimal organizational structures for military missions, *New Trends Coop. Act. Syst. Dyn. Complex Settings.*, 2002.
7. R. A. Layton, M. L. Loughry, M. W. Ohland and G. D. Ricco, Design and validation of a web-based system for assigning members to teams using instructor-specified criteria, *Adv. Eng. Educ.*, **2**(1), 2010, pp. 1–28.
8. J. R. Hackman, The design of work teams, *Handb. Organ. Behav.*, 1987, pp. 315–342.
9. J. MacMillan, E. Entin and D. Serfaty, Communication overhead: The hidden cost of team cognition, in *Team Cognition: Understanding the factors that drive process and performance*, American Psychological Association, 2004, pp. 61–82.
10. R. L. Moreland, L. Argote and R. Krishnan, Socially Shared Cognition at Work, in *What's Social About Social Cognition?*, J. L. Nye and A. M. Brower, Eds. Thousand Oaks, California: Sage Publications, 1996, pp. 57–87.
11. J. Espinosa, R. Kraut and S. Slaughter, Shared Mental Models, Familiarity and Coordination: A Mulit-Method Study of Distributed Software Teams, in *International Conference for Information Systems, ICIS*, 2002.
12. R. M. Kramer and T. R. Tyler, *Trust in Organizations*, Sage Publications, 1996.
13. A. W. Woolley, J. R. Hackman, T. E. Jerde, C. F. Chabris, S. L. Bennett and S. M. Kosslyn, Using brain-based measures to compose teams: how individual capabilities and team collaboration strategies jointly shape performance, *Soc. Neurosci.*, **2**(2), 2007, pp. 96–105.
14. M. T. Brannick, E. Salas and C. Prince, *Team Performance Assessment and Measurement—Theory, Methods, and Applications*, Mahwah, New Jersey: Taylor & Francis, 2009.
15. D. Levi, *Group Dynamics for Teams*, 4th ed. Thousand Oaks, California: Sage Publications, 2014.
16. C. Lin, C. Standing and Y. C. Liu, A model to develop effective virtual teams, *Decis. Support Syst.*, **45**(4), 2008, pp. 1031–1045.
17. G. Hertel, U. Konradt and K. Voss, Competencies for virtual teamwork: Development and validation of a web-based selection tool for members of distributed teams, *Eur. J. Work Organ. Psychol.*, **15**(4), 2006, pp. 477–504.
18. W. G. J. Dyer, J. H. Dyer and W. G. Dyer, *Team Building*, 5th ed. San Francisco: Jossey-Bass, 2013.
19. N. Naikar, B. Pearce, D. Drumm and P. M. Sanderson, Designing Teams for First-of-a-Kind, Complex Systems Using the Initial Phases of Cognitive Work Analysis: Case Study, *Hum. Factors J. Hum. Factors Ergon. Soc.*, **45**(2), 2003, pp. 202–217.
20. R. B. Guay, *Purdue Spatial Visualization Tests—Visualization of Rotations*, Purdue Research Foundation, 1976.
21. B. J. Gimmestad, Gender differences in spatial visualization and predictors of success in an engineering design course, in *Proceedings of the National Conference on Women in Mathematics and the Sciences*, (801), 1990, pp. 133–136.
22. A. Hamlin, N. Boersma and S. Sorby, Do Spatial Abilities Impact the Learning of 3-D Solid Modeling Software?, *Am. Soc. Eng. Educ.*, 2006.
23. S. Sorby, Developing 3-D spatial visualization skills, *Eng. Des. Graph. J.*, **63**(2), 2009, pp. 21–32.
24. J. Koster, A. Velazco and C. Munz, HYPERION UAV: An International Collaboration, in *50th AIAA Aerospace Sciences Meeting including the New Horizons Forum and Aerospace Exposition*, 2012, no. January, pp. 1–23.
25. D. B. Kaufman, R. M. Felder, and H. Fuller, Accounting for individual effort in cooperative learning teams, *J. Eng. Educ.*, no. April, 2000, pp. 133–140.
26. S. Sauer and P. Arce, Team Member Selection: A Functional-Based Approach, in *Proceedings of the 2004 ASEE Annual Conference & Exposition*, 2004.
27. A. W. Woolley, M. E. Gerbasi, C. F. Chabris, S. M. Kosslyn and J. R. Hackman, Bringing in the Experts: How Team Composition and Collaborative Planning Jointly Shape Analytic Effectiveness, *Small Gr. Res.*, **39**(3), 2008, pp. 352–371.
28. J. Suchan and G. Hayzak, The communication characteristics of virtual teams: a case study, *IEEE Trans. Prof. Commun.*, **44**(3), 2001, pp. 174–186.
29. G. E. Dallal, Multiple Comparisons, in *The Little Handbook of Statistical Practice*, 1.1., Amazon.
30. A. D. Taylor, Multiple-Comparison Procedures, *Department of Zoology, University of Hawai'i at Manoa*, 2009. [Online]. Available: http://www2.hawaii.edu/~taylor/z631/multcomp.pdf.
31. U. of W. P. W. Center, Reporting Results of Common Statistical Tests in APA Format, *www.psych.uw.edu*, 2010. [Online]. Available: http://web.psych.washington.edu/writingcenter/writingguides/pdf/stats.pdf.
32. M. W. Ohland, M. L. Loughry, D. J. Woehr, L. G. Bullard, C. J. Finelli, R. A. Layton, H. R. Pomeranz and D. G. Shmucker, The comprehensive assessment of team member effectiveness: Development of a behaviorally anchored rating scale for self and peer evaluation, *Acad. Manag. Learn. Educ.*, **11**(4), 2012, pp. 609–630.
33. B. Tuckman, Developmental Sequence In Small Groups, *Psychol. Bull.*, **63**(6), 1965, pp. 384–399.
34. D. L. Feltz, S. E. Short and P. Joseph, Efficacy Beliefs of Teams, in *Self-Efficacy in Sport*, Illustrate., Human Kinetics, 2008.
35. P. S. U. Eberly College of Science, The coefficient of determination r2, *Stat 501—Regression Methods*, 2015. [Online]. Available: https://onlinecourses.science.psu.edu/stat501/node/29. [Accessed: 10-Jan-2015].
36. L. Artinger, L. Clapham, C. Hunt and M. Meigs, The social benefits of intramural sports, *Naspa J.*, **43**(1), 2006, pp. 69–86.
37. B. Barcelona, Student involvement in campus recreational sports activities and gains in team-functioning, *Cyber J. Appl. Leis. Recreat.*, no. March 2002, 2002, pp. 1–13.
38. J. A. Ross, *Trust Makes the Team Go Round*, 2006.

Brett Stone earned his bachelor's degree in mechanical engineering at Brigham Young University—Idaho and his PhD in mechanical engineering studying design and engineering design teams at Brigham Young University. He currently works for Orbital ATK in their Space Systems Division and enjoys spending time with his wife and their children, the outdoors, sports, and reading.

Steve Gorrell joined the Brigham Young University Mechanical Engineering Department in 2007 following an eighteen year career as an Aerospace Engineer at the Air Force Research Laboratory Propulsion Directorate. He teaches courses on fluid mechanics, gas turbine engine design, turbulence, computational fluid dynamics, and senior capstone design. His research interests include experimental and computational fluid dynamics, turbomachinery, computational science and engineering, and engineering education.

Michael Richey, PhD is the Boeing Chief Learning Scientist and Associate Technical Fellow currently assigned to support technology and innovation research at The Boeing Company. Michael is responsible for leading a team conducting engineering education research projects that focus on improving the learning experience for students, incumbent engineers and technicians. His research encompasses, Sociotechnical Systems, Learning Curves, and Engineering Education Research.