2018-02-01

Improving and Predicting the Effectiveness of Dispersed, Multi-Disciplinary Design Teams

Matthew Oliver Wald
Brigham Young University

Improving and Predicting the Effectiveness of Dispersed, Multidisciplinary Design Teams

Matthew Oliver Wald

A thesis submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Master of Science

Steven E. Gorrell, Chair
John L. Salmon
Sheli D. Sillito Walker

Department of Mechanical Engineering

Brigham Young University

ABSTRACT

Improving and Predicting the Effectiveness of Dispersed, Multidisciplinary Design Teams

Matthew Oliver Wald
Department of Mechanical Engineering, BYU
Master of Science

The use of dispersed (virtual) teams is growing rapidly in the engineering profession. To help prepare students for work in this type of industry, university engineering courses are requiring students to work in teams. Industry leaders and university faculty are interested in improving and measuring the performance of these distributed teams. Surveys, interviews, and observations from the AerosPACE Partners for the Advancement of Collaborative Engineering (AerosPACE) capstone design course are examined to demonstrate how different collaboration tools can be used to best enhance a distributed design team's effectiveness. Collaboration tools to which distributed design teams should give extra consideration at different stages of the product development process are identified and presented in a model. Teams that follow this model will be more effective in their communication patterns.

This study also consists of examining whether peer ratings can accurately predict team effectiveness (as defined by task and relational effectiveness) within a dispersed multidisciplinary, design team. The hypotheses predict that peer ratings will not be unidimensional over time, and will have a positive, significant relationship with team effectiveness. A longitudinal study was conducted on data gathered form the same capstone design course. Confirmatory factor analysis (CFA) was first used to test unidimensionality of peer ratings and structural equation modeling (SEM) was used to model the data and determine any predictive relationships. Model fit statistics are reported to confirm adequate fit for each model. Results showed that while peer ratings are unidimensional at individual time points, they don't behave equally over time and should be considered separately. The structural equation models yielded mixed results, with some parts of peer ratings significantly predicting relational effectiveness and with yet failing to predict task effectiveness. As such, by examining peer assessments, supervisors and faculty will be able to determine and predict relational effectiveness of teams working at different locations, but should use other methods to predict task effectiveness.

Keywords: Virtual teams, Dispersed teams, Collaborative learning, Multidisciplinary design, Electronic communication, Visual communication, Systems Engineering, Peer assessment, Longitudinal study

ACKNOWLEDGMENTS

I would like to acknowledge the help of Steven Gorrell who has coached me in my studies and research, John Salmon and Sheli Sillito who both provided a unique perspective to my work, Ross Larsen who was critical in the statistical analysis, Brett Stone for giving timely advice, the Boeing Company, specifically Michael Vander Wel and Fabian Zender, Brigham Young University, and the students and faculty from all the years of AerosPACE who made this research possible.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1. INTRODUCTION

Learning within teams has become increasingly popular in higher educational settings [1]. ABET accreditation requires applied science, computing, engineering, and engineering technology programs to have some sort of team experience [2] and studies have shown that there is a positive relationship between team learning for engineering students and course performance [3].

Furthermore, with the growing ease and availability of web- and teleconferencing, businesses and universities have also begun exploring and using computer mediated communication for lectures and communication. This form of communication is generally referred to as virtual communication, in which information is conveyed across large distances. As a whole, an increasing number of institutions and companies have begun relying more heavily on this type of communication [4]. This growth of technology allows for the formation of virtual teams, which rely on technology-mediated or virtual communication technologies to work at a distance [5]. The growth of these teams, is largely attributed to a confluence between technology and organizational developments [6] and, in parallel with virtual communication, will likely continue to grow. Dispersed teams have been shown to have some benefits over traditional face-to-face teams, such as making more effective decisions or generating more unique and high quality ideas [5]. In addition, team members are able to work in parallel on projects, within a multi-user paradigm, allowing projects to be completed more quickly.

In contrast to these benefits, there are some detractors for working virtually. For example, there is some concern that since virtual teams hardly (if ever) meet face-to-face, then they are unable to develop strong relational bonds that form the foundation of group identity, cohesiveness, and trust [7–9]. Driskell, Radtke, and Salas have commented that virtual team interaction that is mediated by technology most often leads to less intimacy and difficulty in establishing relationships among team members [10]. Thompson and Coovert found that in a comparison between face-to-face teams and virtual teams, virtual teams reported greater confusion and less understanding of team discussions [11].

One example of a team that relies heavily on virtual tools to communicate is the Aerospace Parters for the Advancement of Collaborative Engineering (AerosPACE) capstone project, sponsored by Boeing [12]. Teams of 10-12 students are tasked with working together to design and build an unmanned aircraft system (UAS) that will complete a chosen mission. Each team is given some latitude to select the specifics of the mission (e.g. what to survey, flight time, and configuration of the drone), while staying within some guidelines (e.g. hand launched, fixed wing, within a predetermined budget, and within the time of the course). Each team is comprised of students from multiple universities and while they're able to meet briefly at the beginning of the project for a quick face-to-face meeting, the rest of the project is done through virtual communication; in addition, each team is required to manufacture at least one part of the UAS at each location, creating a genuine dispersed team environment. The project lasts for about eight months and culminates in a final fly-off event where teams gather together at a single location to showcase their completed UAS to university faculty and Boeing.

To help and evaluate the success of virtual teams such as AerosPACE, we have researched various struggles a dislocated team may face. For example, teams are often faced with trying to decide which method of collaboration between team members is best for their given situation [13,14]. Various methods of communication have been developed and continue to be refined. Each of these mediums has its own characteristics and qualities. Maruping and Agarwal cite media synchronicity theory to emphasize that virtual collaboration effectiveness depends largely on using the correct communication medium for the given task [15]. Levi agrees, stating that a communication tool's effectiveness depends on the fit between the requirements and the characteristics of the tool [16]. It follows that knowing and understanding the characteristics of the various types of tools available for collaboration is essential to effective team communication.

Another dilemma faced when working in a virtual team, is predicting and assessing the effectiveness of a virtual team. One method frequently used to determine team performance is peer assessment. Peer assessment (or peer ratings, peer evaluations) is typically defined as an assessment methodology that allows team members to provide input into the assessment procedure through evaluating each others' performance in out-of-class interactions and activities [17]. Using peer ratings has long been known as useful in predicting success in the military [18], and further research has extensively applied peer evaluations to educational courses, such as speech communications and business [19–21]. Studies have also been conducted within engineering teams. Ohland et al. found no difference in peer ratings based on either gender or race among freshman electrical engineering students [22]. Northrup & Northrup conducted a study on multidisciplinary teams and found that there isn't enough evidence to assume that a dysfunctional team will have a lower average team peer rating [23]. While these studies have added to the understanding of using peer ratings with engineering teams (some multidisciplinary) there is still a need of further research in examining peer ratings among *dispersed* engineering teams.

The purpose of this thesis is to investigate both of these questions, adding to the body of knowledge on how to improve the performance of dispersed design teams. The body of this thesis will be organized in the following manner:

- Chapter 2: Background/Literature Review
- Chapter 3: Collaboration Task-Technology Fit for Student Distributed Engineering Design Teams
- Chapter 4: Using Peer Assessments to Predict Team Effectiveness Among Dispersed, Multi-Disciplinary Capstone Design Teams: A Longitudinal Study
- Chapter 5: Conclusions

Chapter 2 will go into more depth on the research already conducted on virtual team collaboration and performance, giving the reader a base understanding of the current knowledge available on the subject matter. Included will be an overview of the statistical procedures used. Chapter 3 will address the first question by examining and evaluating various types of tools virtual design teams can use. Based on several years of observing AerosPACE and relevant literature, a model has been developed for using collaboration tools during the various stages of the design process. Chapter 4 will answer the second question, by using statistical methods to analyze the relationship between peer assessments and team performance. Both chapters 3 and 4 are also stand alone papers, submitted to journals for publication, and as such will have some information that overlaps with chapter 2. Finally, chapter 5 will conclude with the main points and knowledge gained from this research.

CHAPTER 2. BACKGROUND / LITERATURE REVIEW

In this chapter, an overview of the various types of literature needed to understand the current research is presented. A look at collaboration among virtual teams and the tools available is given, followed by a review at the uses and practices of peer assessments with team performance. Due to the complex analysis performed on the relationship between peer assessments and team performance, a review of the statistical methods used is given.

2.1 Virtual Team Collaboration Model

An in-depth look into the uses and needs of virtual collaboration within engineering settings is presented along with a review of the various types of communication mediums. This sets the ground work for understanding the struggles a virtual team faces, along with understand the different types of tools available to help overcome the various obstacles.

2.1.1 Virtual Collaboration in Engineering

While it is not hard to accept the important connection between effective collaboration and effective engineering design work [24, 25], the context in which engineering design teams collaborate has changed considerably in recent years. Virtual teamwork, where at least one team-members interactions with the rest of the team is mediated by time, distance, or technology [10], has been on the rise over the last several years. The number of workers using virtual means to collaborate has increased and will likely continue growing at around 30 percent per year [26]. Other researchers agree, adding that most large companies use virtual teams in at least some way [27]. In a survey of hundreds of private and public organizations, WorldatWork found that in 2013 more than one third of organizations in the manufacturing, consulting, professional, scientific, and technical fields offered positions for employees to work remotely. Roughly half of organizations in those fields also offered positions which required virtual collaboration at least once a week [28].

Engineering industries have also increasingly used virtual teams as their operations have become more global. In a 2003 study of companies in the engineering, procurement, and construction industry, over half the companies surveyed used virtual teaming [29]. Nearly every company surveyed believed use of virtual teams would increase considerably over the next five years.

The increased presence of this virtual, geographically dispersed type of collaboration has become apparent to engineering educators who argue that teaching engineering students how to be successful in these types of situations is crucial for overall career success in modern industry [30, 31]. Lang et al., reflecting after surveying 15 aerospace and defense companies, note that many items important to industry (such as problem solving abilities and ability to work in a team setting) are not specifically addressed in traditional engineering education and pose a question of how engineering education ought to train young engineers in virtual team collaboration [32]. Dym et al. state that "the purpose of engineering education is to graduate engineers who can design," and that design is an inherently team based, social process [14]. Given this situation in which modern engineering teams in industry find themselves, Dym and his colleagues encourage engineering educators to embrace the concept of teaching engineering courses across geographic space. The findings of these researchers are consistent with the conclusions of the American Society of Mechanical Engineerings (ASME) "Vision 2030" report which finds that industry and academia are, in some areas, fundamentally misaligned [32]. For example, industry and academy differ greatly when assessing the strengths of recent graduates. The greatest differences came in the areas of recent graduates abilities to embrace new technologies, communicate effectively, and understand how things are made and work.

Koster et al. gave senior engineering students from four universities located around the world experience in this type of collaboration through the Hyperion project [13]. Their project attempted to use a follow the sun work-flow to design, build, and fly a UAS. The follow the sun work-flow involves three different work locations, evenly spaced around the globe such that each could 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. The authors commented on the fact that while communication is essential for this type of collaboration, students are generally only trained in how to communicate in local, face-to-face settings. The fact that these students were often working when their teammates were asleep, minimized opportunities to use "same time, different place" communication alternatives [16] like phone calls. These challenges were part of the project's pedagogical design as the educators wished to instruct the students by allowing them to gain experience.

Doerry et al. created a curricular model named the Global Engineering College to help engineering students from different countries take courses together, including design courses such as robotics [31]. In their program, students from multiple disciplines, such as mechanical, electrical, and civil engineering, worked together from the Northern Arizona University campus with "outsourcing consultants" who were students studying similar disciplines in universities in Wroclaw, Poland and Dresden, Germany. Similar to the experience of Koster et al., these students found it difficult to communicate with each other and had to adapt by learning how to effectively use novel collaboration tools designed specifically for this course. This was not an easy process, and the researchers found that students did not naturally attune themselves to effective collaboration. Even when software tools custom-built to allow students greater communication and project management were presented, they were mostly under-utilized. Students often defaulted to communicating over email. Eventually, the researchers found that media-rich, synchronous communication is important in establishing trust, commitment, and excitement among geographically dispersed team members, and that high amounts of structure may be necessary to kick-start projects of this type.

For several years now, AerosPACE has attempted to bridge not only geographic distance and disciplinary boundaries, but also involve industry more directly in preparing students for the realities of virtual engineering collaboration. In fact, when students were asked in a survey at the end of the 2014-2015 year of AerosPACE what skills they felt they had gained from the program, "Collaboration", "Teamwork", and "Virtual Teaming" were the most frequently mentioned items. This is especially encouraging given this statement from the ASME "Vision 2030" report, that, in addition to technical knowledge, successful mechanical engineers will need to, "have excellence in communication, management, global team collaboration, creativity, and problem-solving skills" [32].

2.1.2 Mediums of Communication

One of the great challenges of virtual collaboration is effective communication. Various methods of communication are available, each with its own characteristics and qualities. Maruping and Agarwal cite media synchronicity theory to emphasize that virtual collaboration effectiveness depends largely on using the correct communication medium for the task, which is corroborated by Levi [15, 16]. It follows that knowing and understanding the characteristics of the various types of tools available for collaboration is essential to effective team communication. In the case of face-to-face conversation, multiple forms of communication, such as words, voice inflection, facial expressions, and body language are all transmitted and received simultaneously. Most of these are missing in the case of a text message.

Communication mediums have various characteristics by which they can be measured. Perhaps the most commonly cited characteristic in the literature, developed by Daft and Lengel, is "richness", or the ability to transmit a given amount of information in a given amount of time [33]. An example of a rich communication medium would be face-to-face communication, while an example of a low richness communication medium is a simple text message. Maruping and Agarwal suggest five criteria for determining media richness (immediacy of feedback, symbol variety, parallelism, rehearsability, and reprocessability), while Driskell et al. give six (co-presence, visibility, audibility, cotemporality, simultaneity, and sequentiality) [10,15]. Other researchers have similarly suggested their own sets of criteria [34–36].

The definitions of these characteristics given by these researchers overlap with each other in many areas and can generally be summarized into the following categories:

- Media Richness
- Symbol Type
- Time to Response
- Permanence
- Parallelism
- Accessibility

Media Richness can be defined in the same way as Daft and Lengel described above. Symbol Type is defined as the classes of "symbols" used to transmit the message. For example, Dym et al argue that various "languages" are needed for design to successfully take place, such as verbal or textual statements, graphical representations, and mathematical or analytical models [14,37]. In addition to these suggested by Dym et al. types such as audio, video, and body language, are also important. For example, a raised eyebrow during an in-person conversation may symbolize doubt or concern more succinctly than a textual statement in an instant messaging application. It may be tempting to assume that a richer communication medium is always desirable; however, in certain situations, such as group brainstorming, too rich of a medium has actually been shown to hinder group effectiveness [16].

Time to response refers to two closely related characteristics: the ability of the medium to enable a response to a message in a certain amount of time (instantaneously or slower), and the socially dictated time within which a response is acceptable. As an example, it takes time to type a response to an email, click send, and then possibly wait for network latency. Depending on context, however, it can often be socially acceptable to not respond to an email for as long a couple days.

Permanence explains how easily the contents and sequence of an exchange are recorded and reviewed later. While the contents of an email and its subsequent replies are automatically preserved in order without any extra effort by the communicators, the same is not true of many other mediums, such as when making a telephone call or having a face-to-face conversation. Special solutions or tools to record various types of communication exist, but for this definition we consider only whether tools have built-in characteristics of automatic recording and ordering of messages as a standard feature for all users.

Parallelism describes whether a communication medium allows the user to carry on multiple conversations simultaneously. For example, when speaking with someone in person, one is unlikely, based on social acceptability and convenience, to carry on more than one conversation at a time. However, when sending text messages, it is common to be involved in multiple conversations simultaneously.

Accessibility addresses the fact that some communication tools require either special skills or special software to use them effectively. For example, to successfully use video conferencing over the internet, all participants must have the required software. They must also all have the necessary hardware, such as a webcam, and the knowledge to use the software and hardware tools. Another important aspect of Accessibility is access to resources such as high-speed networks and permissions, including firewall access. Access is also important in other, less technical mediums of communication, such as speaking face-to-face. Having to travel significant distances to communicate face-to-face affects the accessibility of this medium in today's engineering environment.

These different communication classifications will provide the foundation for determining which virtual communication tools should be used for the different stages in the design life-cycle.

2.2 Predicting and Measuring Team Effectiveness

A more in-depth look will now take place of how peer assessments may be used to predict and measure team effectiveness of a dispersed multidisciplinary design team. Due to the complex nature of the statistical analysis by adding the element of time, a review into the statistical measure and process of analyzing the data gathered will also be presented.

2.2.1 Peer Assessment

Peer assessment has been defined over the years as the process of having the members of a group judge the extent to which each of their fellow group members exhibit specified traits, behaviors, or achievements [38]. While there are many different ways to administer peer evaluations, most tend to closely follow one of two methods. The first involves a set of questions which ask team members to rate their peers on topics such as attendance at meetings, quality of work, or easiness to work with [39,40]. This method has been used primarily because of the high reliability of the ratings [41].

The second method for peer assessment, as stated by Oakley et al. [42], is a single rating on the contribution of each team member to the final project. The question may vary, but often asks students to rate their peers from "Excellent" (defined as someone who consistently went above and beyond the call of duty) to "No show" (defined as someone who had no participation at all) [43]. When using this type of peer evaluation, team members will often be asked a series of questions to get them thinking about their peers performance before being asked to give the overall rating [42]. This form of rating has been used considerably throughout the years for its ease of administration.

In the case of AerosPACE, the average team size is close to eleven students per team. If the multiple set questionnaire were used, students would be asked to consider over 100 different questions when evaluating their peers. This might potentially cause rater fatigue, or simply be more than the students are willing to answer. The single question rating would seem to be ideal for this setting, but due to a study conducted by Paswan et al. [44], in which it was argued that peer assessments are multidimensional, a hybrid of the two peer evaluations was pursued.

Gueldenzoph & May suggest that the general criteria used for peer evaluations fit into five categories: commitment to the group, ability to deal constructively with conflicts that arise, active participation in the decision-making process, accountability for assigned tasks, and assumption of initiative or leadership role [17]. Beebe et al. suggest that peer ratings fit into two main categories of task and relational competencies [39]. Dyer et al. suggest that successful teams should have adequate interpersonal skills, technical expertise, and motivation [45]. For the AerosPACE course, the peer evaluation instrument was selected based on the recommendations of Dyer et al., since the elements from other researchers peer assessments readily fit into one of these three categories (e.g. [17, 39, 41]). Each team member was asked to rate his/her teammates based on their overall motivation, technical skills, and social skills. Validation was performed to ensure that the peer assessment tool chosen would indeed be a valid measurement tool, which will be presented later.

2.2.2 Team Effectiveness

There are several different ways to measure success in team performance, but perhaps the most influential and widely cited definition comes from J. R. Hackman [46]. He states that team effectiveness is comprised of three main components: first, the degree to which the group's productive output meets the standards of quantity, quality, and timeliness; second, the degree to which the process of carrying out the work enhances the capability of members to work together interdependently in the future; third, the degree to which the group experience contributes to the growth and personal well-being of team members. While all three of these are important factors, this study focuses on the first two criteria (later referred to as task effectiveness and relational effectiveness).

A large amount of team research has been put into measuring a team's operational, quantitative, and qualitative objectives [47] or in other words a team's effectiveness on a given task. Levi et al. argue that "from a management perspective, the obvious definition of team success is successful performance on a task" [48]. This is usually a popular method of measurement due to its objective nature [49]. In academia, task effectiveness is often measured by performance on assignments, presentations, or projects. Other criteria such as overall GPA or grades assigned in a class are also commonly used [3, 50, 51].

Relational effectiveness of a team has also been a popular measure of team outcomes [52, 53]. Since the main criteria for this metric is the ability of the team to continue working together in the future [16, 54], items such as trust, commitment, and satisfaction play a key role in defining success. When studying this area, Whitman et al. examined the difference in satisfaction of a computer-mediated communication design team (dispersed team) verses a face-to-face design team, and found that dispersed teams are generally less satisfied than their face-to-face counterparts, due to ineffective team members and difficulty with using communications technology [51]. Furthermore, Hinds & Mortensen have found that distributed teams have greater difficulty in developing trust [55]. Researchers have also studied the emergence of trust within teams and concluded that it is closely tied to team effectiveness [56, 57]. It is important to note that while relational effectiveness is important, it should not be confused for *task* effectiveness, nor used as a catch-all for a team's overall success. As Hackman suggested, if a team is satisfied and yet unsuccessful in producing a quality project, then in most instances the group would be considered ineffectual [46].

In the AerosPACE program, task effectiveness was defined by looking at team scores on regularly given presentations and whether or not the team succeeded in flying their UAS at the end of the project. Relational effectiveness was defined as the satisfaction of each student with their team over time and the amount of trust within each team at the end of the project. By including both task and relational measures it gives an objective view of success by looking at how well each team performed on the course project, while also incorporating a subjective measure of the ability of each team to continue working well together.

2.2.3 Statistical Analysis

The process for the statistical analysis was taken in large part from Wang & Wang's book, *Structural Equation Modeling: Applications Using Mplus* [58]. This process involves conducting a confirmatory factor analysis (CFA), followed by a structural equation model (SEM). Mplus [59] was used as the statistical software package in order to examine the data. This analysis was used to determine the relationships between peer assessments and team effectivenss as given in the fourth chapter.

In each case, a model is created from the data gathered. The concept of the model is similar to a line of best fit, commonly used for simpler data sets. In each case, model fit statistics are used to essentially determine whether or not the model adequately represents the data. These fit statistics are defined as follows: the chi-square test (with a non-significant p-value at the α = 0.05 level indicating a good model fit), the root mean square error of approximation (RMSEA) value (with values less than 0.05 indicating a good fit), the comparative fit index (CFI) and Tucker-Lewis index (TLI)(with values greater than 0.90 indicating good model fit), weighted root mean square residual (WRMR)(with values less than 1.0 signifying good model fit) used for models with categorical data, and standardized root mean square residual (SRMR)(with a value less than 0.10 indicating an acceptable model fit) used for models with only continous data. These values for indicating good model fit were taken from Wang & Wang [58]. It's important to note that although the chi-square test is reported for each statistical model tested, it isn't given much credence due to its sample-size dependency and the large number of respondents in the current study [60].

Confirmatory Factor Analysis

CFA is used to determine if differing questions are measuring the same factor or not. The procedure followed is given by Muthen [59] and further information on this procedure is given by Brown [61] in his book *Confirmatory Factor Analysis for Applied Research*. The steps taken are summarized as follows:

1. Test for unidimensionality of indicators at each time.
2. Determine the shape of the curve for each indicator.
3. CFA of all time-points together:
   (a) Covariance structure analysis without measurement parameter invariance.
   (b) Covariance structure analysis with invariant slope coefficients.
   (c) Mean and covariance structure analysis with invariant measurement intercepts and loadings.

A test for unidimensionality is performed first at each time-point, meaning that the questions used in this study are tested to see if they measure the same thing at individual time-points. For example, two questions that ask, “How many questions did the student get right?” and “What score did the student get?” are two different questions that measure the same thing. With only three indicators, a one-factor CFA model would have just enough variables to solve the set of equations and thus be unable to provide any additional information for model fit statistics. To solve this problem, the residual variances are constrained to be equal, giving the needed degrees of freedom to calculate model fit statistics for this step.

Following the test for unidimensionality at separate time points, the shape of the model for each indicator over time is examined to determine whether the individual questions are static, linear, or quadratic. In order to continue with the CFA, each question must have the same behavior over time, meaning either quadratic, linear, or static. Each model is first assumed to to be quadratic and tested to see if the model fit the data well and if the intercept, slope, and quadratic coefficients are significant. The significance is determined by looking at the p-value assigned to each coefficient, with a value lower than 0.05 meaning the coefficient is significant. If the model fails or if the coefficients are not significant, then the highest order term that is non-significant is removed and the model retested.

Once the shape for each indicator is found, a covariance analysis without measurement parameter invariance (often referred to as configural invariance analysis) is performed. This step involves loading the indicators onto a latent variable, but not constraining any of the indicators across time. Model fit statistics are calculated to determine adequate fit. Following this model, the covariance structure analysis with invariant slope coefficients (commonly referred to as metric invariance analysis) is performed. For metric invariance analysis, the same model is used as the configural model, except the slope coefficients are constrained to be equal for each indicator. Model fit statistics are again examined, along with a comparison to the configural model. The two statistics examined are chi-square difference test (with a non-significant p-value meaning the two models are invariant) and the change in CFI statistic (with a ΔCFI ≤ 0.01 meaning the two models are invariant). Following, the mean and covariance structure analysis with invariant slope and intercept coefficients (commonly known as scalar invariance analysis) is performed. This model is similar to the configural model, except the intercept and slope coefficients are constrained to be equal for each indicator. Model fit statistics are then examined and the model is compared to the configural model. In each case, model invariance (there is no variance between the three models created) would mean that the three indicators behaved equally across time and can be treated as three questions measuring the same concept. Any variance between models would have shown a need to analyze the indicators separately.

In summary, these three procedures help to indicate whether or not the factors in question are measuring the same underlying variable.

Structural Equation Modeling

Once the CFA is performed, the SEM's are created using Mplus. The model shapes were determined within the CFA steps. The first step is to examine the model fit statistics and determine whether the model adequately fits the data. If the model fits, then correlation scores and p-values are examined between the various relationships.

Relationships are specified within the code and based on literature and assumptions. Supposing two linear regressions are calculated, then certain relationships may exist and others are simply nonsensical. For example, the slope of one linear regression may predict the slope of another linear regression, but not the intercept. The intercept of a linear regression, however, may predict the slope *and* intercept of another linear regression. The relevance of these relationships are based on the p-value, with a value less than 0.05 indicating a significant relationship.

The correlations between these relationships also have varying correlation factors, also known as weights. If the intercept of one linear regression is shown to predict the slope of another linear regression, then with each unit increase of the intercept, the slope would increase by the correlation factor. If one slope is shown to predict another, then with each unit increase of the slope, the predicted slope would again increase by the corresponding correlation factor. The predictive relationships could also be negative, meaning that an increase in one would predict a decrease in another.

CHAPTER 3. COLLABORATION TASK-TECHNOLOGY FIT FOR STUDENT DISTRIBUTED ENGINEERING DESIGN TEAMS

3.1 Introduction

Before the advent of distributed design and manufacturing, collaboration in engineering design projects was a relatively straightforward effort. Given their close physical proximity, when two or more people needed to coordinate their design efforts, they sought each other out and spoke in person. Today, with increased dispersed design and manufacturing capabilities [62–64] teams that develop and manufacture those designs have to adapt and use new collaboration techniques. Such changes suggest approaches to teaching engineering design at universities should also adapt.

One area in which teaching of engineering design should adapt is in the teaching of communication tools. The Aerospace Partners for the Advancement of Collaborative Engineering (AerosPACE), a program sponsored by Boeing, is one example of how industry and academia are attempting to adapt and better prepare engineering students for the changing environment [12, 65–67]. In AerosPACE, students from various universities majoring in various technical disciplines are combined into teams with experienced professors as coaches to design, build, and fly Unmanned Aircraft Systems (UAS) that performs a specific mission. The program has grown from 19 students from four universities in 2012, to 72 students from 8 different universities over the course of four years, involving more than 150 total students.

A question that is often asked in this type of situation is, "At a given stage of the project's development, what type of collaboration between team members is optimal?" [13, 14]. In this research, we have addressed this question by examining and evaluating various collaboration tools and their uses. While some papers have looked at task-technology fit in the context of general teams [15], this paper fills a niche by applying task-technology fit in the context of dispersed design and manufacturing teams, composed of engineering students at multiple locations. We call upon four years of AerosPACE experience, combined with current literature, to provide information on which collaboration tools should be used at the various stages of student product development.

3.2 Background

The background section will give a review of pertinent literature. An extensive review of how virtual collaboration has been employed in engineering, along with various mediums of communication and how they are defined.

3.2.1 Virtual Collaboration in Engineering

While it is not hard to accept the important connection between effective collaboration and effective engineering design work [24, 25], the context in which engineering design teams collaborate has changed considerably in recent years. Virtual teamwork, where at least one team-member's interactions with the rest of the team is mediated by time, distance, or technology [10], has been on the rise over the last several years. Golden and Raghuram cite various sources showing that the number of workers using virtual means to collaborate has increased and will likely continue growing at around 30 percent per year [26]. Other researchers agree, adding that most large companies use virtual teams in at least some way [27]. In a survey of hundreds of private and public organizations, WorldatWork found that in 2013 more than one third of organizations in the manufacturing, consulting, professional, scientific, and technical fields offered positions for employees to work remotely full-time. Roughly half of organizations in those fields also offered positions which required virtual collaboration at least once a week [28].

Engineering industries have also increasingly used virtual teams as their operations have become more global. In a 2003 study of companies in the engineering, procurement, and construction industry, over half the companies surveyed used virtual teaming [29]. Nearly every company surveyed believed use of virtual teams would increase considerably over the next five years.

In the commercial aerospace industry, Boeing's 787 offers an example. A large majority, 65 percent, of the new Dreamliner is supplied to Boeing by dozens of other companies located across the globe [63]. Engineers from supplier companies and Boeing are required to work together at unprecedented levels across great distances to generate designs, manufacture, and assemble the aircraft.

The importance of this virtual, geographically dispersed type of collaboration has become apparent to engineering educators who argue that teaching engineering students how to be successful in these types of situations is crucial for overall career success in modern industry [30, 31]. Lang et al. survey 15 aerospace and defense companies and note that many items important to industry are not specifically addressed in traditional engineering education and pose a question of how engineering education ought to train young engineers in virtual team collaboration [68]. Dym et al. state that "the purpose of engineering education is to graduate engineers who can design," and that design is an inherently team based, social process [14]. Given this situation in which modern engineering teams in industry find themselves, Dym and his colleagues encourage engineering educators to embrace the concept of teaching engineering courses across geographic space. The findings of these researchers are consistent with the conclusions of the American Society of Mechanical Engineering's (ASME) "Vision 2030" report which finds that industry and academia are, in some areas, fundamentally misaligned [32].

Koster et al. gave senior engineering students from four universities located around the world experience in this type of collaboration through the Hyperion project [13]. Their project attempted to use a "follow the sun" work-flow to design, build, and fly a UAS. The Follow the Sun work-flow involves three different work locations, evenly 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. The authors commented on the fact that while communication is essential for this type of collaboration, students are generally only trained in how to communicate in local, face-to-face settings. The fact that these students were often working when their teammates were asleep, minimized opportunities to use "same time, different place" communication alternatives like phone calls [16]. These challenges were part of the project's pedagogical design as the educators wished to instruct the students through experience.

Doerry et al. created a curricular model named the Global Engineering College to help engineering students from different countries take courses together, including design courses such as robotics [31]. In their program, students from multiple disciplines, such as mechanical, electrical, and civil engineering, worked together from the Northern Arizona University campus with "outsourcing consultants" who were students studying similar disciplines in universities in Wroclaw, Poland and Dresden, Germany. Similar to the experience of Koster et al., these students found it difficult to communicate with each other and had to learn how to effectively use novel collaboration tools. This was not an easy process, and the researchers found that students did not naturally attune themselves to effective collaboration. Even when software tools custom-built for the project were presented to students, they were mostly under-utilized. Students often defaulted to communicating over email. Eventually, the researchers found that media-rich, synchronous communication is important in establishing trust, commitment, and excitement among geographically dispersed team members, and that high amounts of structure may be necessary to kick-start projects of this type.

These examples show a gap in the current body of knowledge on how to effectively use collaboration tools in distributed, student design and manufacturing teams. For several years now, AerosPACE has attempted to bridge this gap by not only studying communication among teams separated by distance and disciplinary boundaries, but by also involving industry directly in preparing students for the realities of virtual engineering collaboration. By working directly with industry, we have noticed an increase in how prepared alumni of the AerosPACE program feel about entering the work force. When students were asked in a survey at the end of the 2014-2015 year of AerosPACE what skills they felt they had gained from the program, "Collaboration/Teamwork", and "Virtual Teaming" were the two most frequently mentioned items (Figure 3.1). The fifth most mentioned item, behind "Manufacturing" and "Project Management (Leadership)", is "Communication". This is especially encouraging given the statement from the ASME "Vision 2030" report, that, in addition to technical knowledge, successful mechanical engineers will need to, "...have excellence in communication, management, global team collaboration, creativity, and problem-solving skills" [32]. This effort by AerosPACE to prepare students for work in industry is certainly being seen in the skills and attitudes of the students that go through the program.

3.2.2 Mediums of Communication

Various methods of communication for virtual teams are available, each with its own characteristics and qualities. Maruping and Agarwal, along with Levi, emphasize that virtual collabora-

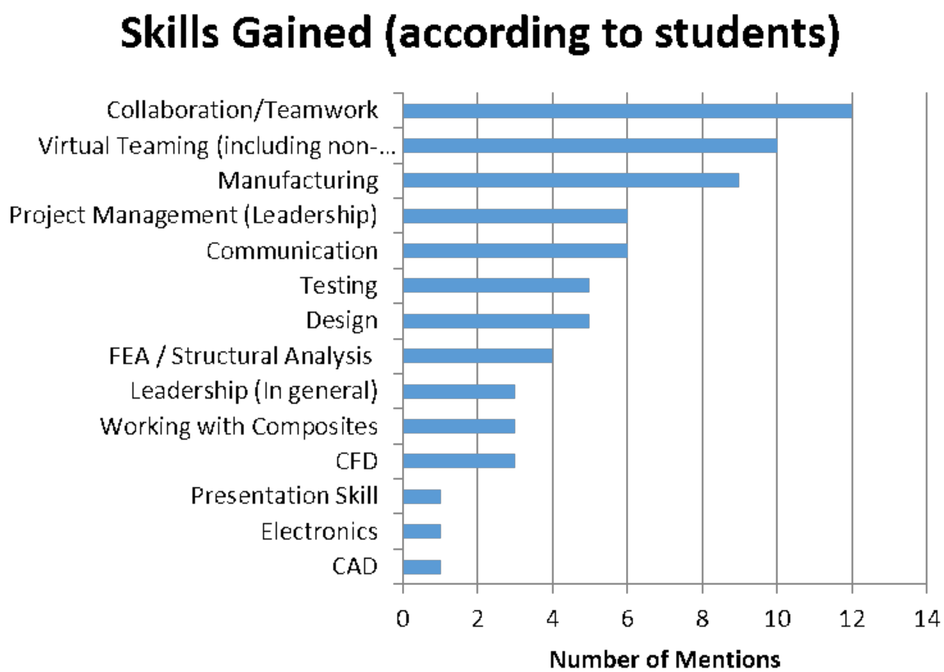

Figure 3.1: Students indicated which skills they felt AerosPACE had helped them gain.

tion effectiveness depends largely on using the correct communication medium for the task [15,16]. It follows that knowing and understanding the characteristics of the various types of tools available for collaboration is essential to effective team communication. In the case of face-to-face conversation, multiple forms of communication, such as words, voice inflection, facial expressions, and body language (along with others) are all transmitted and received simultaneously. Most of these, however, are missing in the case of a text message.

Communication mediums have various characteristics by which they can be measured. Perhaps the most commonly cited characteristic in the literature, developed by Daft and Lengel, is "richness", or the ability to transmit a given amount of information in a given amount of time [33]. An example of a rich communication medium would be face-to-face communication, while an example of a low richness communication medium is a simple text message. Maruping and Agarwal suggest five criteria for determining media richness (immediacy of feedback, symbol variety, parallelism, rehearsability, and reprocessability), while Driskell et al. give six (co-presence, visibility, audibility, cotemporality, simultaneity, and sequentiality) [10,15]. Other researchers have similarly suggested their own sets of criteria [34–36].

The definitions of these characteristics given by these researchers overlap with each other in many areas. Comparing the definitions offered by these authors for each of their characteristics and considering our own experience, we suggest the following set of metrics:

- Media Richness
- Symbol Type
- Time to Response
- Permanence
- Parallelism
- Accessibility

Media Richness we define in the same way as Daft and Lengel described above. Symbol type is defined as the classes of "symbols" used to transmit the message. For example, Dym et al. argue that various "languages" are needed for design to successfully take place, such as verbal or textual statements, graphical representations, and mathematical or analytical models [14, 37]. We propose that in addition to these suggested by Dym et al., that types such as audio, video, and body language, are also important. For example, a raised eyebrow during an in-person conversation may symbolize doubt or concern more succinctly than a textual statement in an instant messaging application. It may be tempting to assume that a richer communication medium is always desirable; however, in certain situations such as group brainstorming, a less rich medium is beneficial to the communication efforts of a team ([16] see pg. 236).

Time to response refers to two closely related characteristics: the ability of the medium to enable a response to a message in a certain amount of time (instantaneously or slower), and the socially dictated time within which a response is acceptable. As an example, it takes time to type a response to an email, click send, and then possibly wait for network latency. Depending on context, however, it can often be socially acceptable to not respond to an email for an extended period.

Permanence explains how easily the contents and sequence of an exchange are recorded and later reviewed. While the contents of an email and its subsequent replies are automatically preserved in order without any extra effort by the communicators, the same is not true of many other mediums, such as when making a telephone call or having a face-to-face conversation. Special solutions or tools to record various types of communication exist, but for this definition we consider only whether tools have built-in characteristics of automatic recording and ordering of messages as a standard feature for all users.

Parallelism describes whether a communication medium allows the user to carry on multiple conversations simultaneously. For example, when speaking with someone in person, one is unlikely, based on social acceptability and convenience, to carry on more than one conversation at a time. However, when sending text messages, it is common to be involved in multiple conversations simultaneously.

Accessibility addresses the fact that some communication tools require either special skills or special software to use them effectively. For example, to successfully use video conferencing over the internet, all participants must have the required software. They must also all have the necessary hardware, such as a webcam, and the knowledge to use the software and hardware tools. Another important aspect of accessibility is access to resources such as high-speed networks and permissions, including firewall access. Access is also important in other, less technical mediums of communication, such as speaking face-to-face. Having to travel significant distances to communicate face-to-face affects the accessibility of this medium in today's engineering environment.

Considering each of these characteristics, a clearer comparison can be drawn among the various communication mediums available to virtual teams. We adapt the lists of communication tools from Maruping and Agarwal [15], Driskell et al. [10], Daft and Lengel [33], French et al. [36], and Levi and Rinzel [34] for our use:

- Face to Face
- Telephone (one to one)
- Teleconference (many to many)
- Text / Instant Messaging
- Web Conferencing
- Video Conferencing
- Email

- Shared Database
- Social Media

Dennis et al. present various communication tools and how each is rated on a scale of low to high, slow to fast, or few to many [69]. Maruping and Agarwal also present some information on the ratings of various communication tools, while also including media richness [15]. Based on these findings, a comparison of these tools has been summarized and can be seen in 3.1.

Although most of these communication and collaboration tools are well known and easily distinguishable, some of them deserve slightly more description to avoid ambiguity. Teleconferencing, web conferencing, and video conferencing are all similar in some ways, but distinct in others. In this paper, we define teleconferencing as a telephone call for more than two people. Web conferencing includes the same services as teleconferencing and adds internet based tools that allow participants to share screens, view slides, or chat. Thus, web conferencing uses more symbol types than teleconferencing. Video conferencing tools, such as Skype, can include all the previously mentioned capabilities as well as the ability to see a live video feed of each participant. Previously, video conferencing involved purchasing expensive equipment and software, but has now evolved to being available through web based applications and using relatively inexpensive equipment.

3.3 Methods

This study includes 150 students, consisting of 128 males and 22 females enrolled in several different universities and colleges across the United States. Participating institutions were Brigham Young University, Georgia Tech University, Purdue University, Tuskegee University, Embry Riddle Aeronautical University, Clemson University, Washington State University - Pullman, Washington State University - Everett, and Everett Community College. The average age was 22.9 years old (SD = 2.49) and included mechanical engineering, aerospace engineering, manufacturing engineering, electrical engineering, and computer engineering majors. The project lasted for two semesters (late August through April) and teams worked together for the duration of the project. Surveys were administered four times throughout the project at the midpoint and end of each semester (approximately 6, 14, 23, and 30 weeks). Students were assured that responses were

Communication Tool	Media Richness	Symbol Type	Response Time	Permanence	Parallelism	Accessibility
Face-to-Face	Highest	Multiple	Low	Low	One	Low or High
Phone Call	Medium	Audio	Low	Low	One	High
Teleconference	Medium	Audio	Low	Low	One	Medium
Text / Instant message	Low	Text	Medium	Medium	Multiple	High
Web Conferencing	Medium-High	Multiple	Low	Low	One	Medium
Video Conferencing	High	Multiple	Low	Low	One	Low-Medium
Email	Low	Multiple	Low	Medium	Multiple	Medium
Shared Database	Medium	Muliple	Medium	High	Medium	Medium
Social Media	Low-Medium	Multiple	Medium-High	High	Multiple	Medium

Table 3.1: List of communication tools and their ratings based on communication metrics.

kept anonymous, wouldn't be shared with other students, and wouldn't affect students' grades. This was done to ensure honest and accurate results [70]. Responses from surveys were used to gather information about the communication patterns and experiences of students.

3.4 Proposed Model Of Collaboration

Dieter and Schmidt have proposed a product development process that includes three stages for a product design process: Conceptual Design (which includes getting customer requirements, defining the objectives, and generating concepts), Embodiment Design (which includes determining overall product architecture and individual design of parts), and Detailed Design (which includes the detailed design of all parts, part integration, and creation of final schematics) [71]. NASA has also developed a toolbox with their own suggested workflow as follows [72]:

- Conceptual Trade Studies
- Concept Definition
- Design and Development
- Fabrication, Integration, Test, and Evaluation
- Operations

These two processes are similar in stepping through concept generation and design, with NASA adding in the manufacturing, evaluation and utilization sections of the product life-cycle. We combine the NASA model with Dieter and Schmidt's process, and simplify them to create three basic phases: Early, Middle, and Late. As a general overview, the early stage encapsulates the conceptual trade studies and concept definition. The middle stage incorporates the detailed design and development of the product. The late stage includes the fabrication, integration, test, and evaluation of the product. In the following sections, each stage will be explained in greater detail.

We will provide, based on our own experience and literature, a suitable model for which collaboration tools should be used during each of the three mentioned stages. This model is intended as a guide for distributed student design and manufacturing teams. The synthesized information will allow for students and professors to better understand which collaboration tools should be used during a product design team's life-cycle.

3.4.1 Early Stages

The early stage of the design process in AerosPACE is defined by teams creating a mission definition (based on customer requirements), generating multiple concepts, and evaluating the concepts until one has been selected for detailed design. The main goal for communication during this stage is to create a cohesive relationship between team members that will enable successful dialogue in the future and to convey created concepts and evaluations to team members.

While there are specific communication tools that help encourage a good working relationship, we have learned in our experience with AerosPACE that whenever possible, in-person meetings, such as a program kickoff, should be held. In the 2013-2014 AerosPACE year, no in-person kickoff meeting was held, but students did meet each other in person at the end of the academic year. Afterwards and during course evaluations, students indicated that many of the issues or problems they faced throughout the year, such as personality conflicts, could have potentially been minimized or eliminated if they had met in person at the beginning of the year. For example, one student, when asked whether he felt a kickoff meeting would have helped with some of the interpersonal challenges said, "I really think it would. I think once you establish a person with a voice and with a face, you actually get to know everyone a little better and kind of where everyone's coming from."

At the beginning of the following year (2014-2015), a kick-off meeting was held, at which all students from all teams met in one location. They brainstormed, conducted team-building activities, began work on responding to the program Request for Proposals (RFP), and socialized during dinner. Later, when asked what portions of AerosPACE they felt had gone well, the second most mentioned item was the kickoff meeting. Students offered comments such as, "...the kickoff event was very important for the health of the team throughout the semester," and "The Kick Off meeting was a good start to the program."

Other researchers agree [27]. Siebdrat et al. stress the importance of a kick-off meeting to help virtual teams develop a shared understanding of the project and encourage social cohesion [73]. This shared understanding of the roles, skills, and responsibilities of each team member along with a general knowledge of the project is often referred to as a shared mental model and is critical in the formation and overall success of a team [74]. Lovelace et al. state that one reason dispersed teams struggle with forming a shared mental model is virtual teams are less likely in the early stages of development to have developed the norms of openness and debate required for task conflict to be effective [75]. Hackman also agrees that even well-structured virtual teams need to have everyone physically present for a launch meeting [76]. In some instances, a kick-off meeting might not be practical, such as in the case of a short project life-cycle or for a team that has already had extensive experience working together. In these instances, a kick-off meeting may still be beneficial for establishing team roles and deciding upon a mission definition, but perhaps not practical.

After an in-person meeting has been held, it's important to use communication tools that are high in media richness, allowing for the growth of the relationship among team members [77]. Tools such as video conferencing become valuable due to high media richness [15] and can enable participants to develop trust and cohesion through a richer interaction [16]. This is especially critical as teams define the mission specifications, since this is when team members begin forming the norms that are critical to effective communication and making one of their first critical decisions. Teleconferencing and web conferencing are also viable options, but should defer to video conferencing when possible, as their level of media richness is lower.

Video conferencing is beneficial during this stage due to the low level of parallelism and low time to respond. As stated by Malhotra et al., virtual meetings are most successful when team members are engaged in the meeting and not distracted [78]. Communication tools with higher levels of parallelism allows team members to multi-task while in a meeting, causing dissonance or a lack of participation [79]. A slightly higher order skill that we have also found useful, is for meeting leaders to request verbal confirmation from specific participants to confirm reception of messages and engagement [16, 78]. As Dennis and Valacich suggest, feedback (or time to response) becomes important when the goal of communication is to achieve convergence [80]. By using video conferencing, with its low time to response, student teams that have never worked together previously can build a cohesive relationship and better come to a consensus about the mission requirements.

There are times when video conferencing or web conferencing may not be a viable option due to accessibility. Although these tools offer the highest levels of media richness, they are also prone to technical difficulties. Levi et al. emphasize the need to learn how to use these tools effectively to reduce such difficulties and misunderstandings [16]. In AerosPACE, we have experienced first-hand the saying, "Technology always adds to a meeting, and it's usually about 15-20 minutes". Thus, ample time should be given for instructing team members on the correct way to use each communication tool and allowances given to team members when software doesn't behave as expected. One student, interviewed during the 2013-2014 program, stated that he had participated in several web-conference meetings where he could only hear about 30 percent of the conversation because those speaking were sitting too far away from the microphone. Test-runs should be performed with the chosen communication tool to help reduce these difficulties and allowances given to team members when software doesn't behave as expected. A slightly higher order skill that we have found useful in video conference meetings is for leaders to request verbal confirmation from specific participants to confirm reception of messages during meetings [16].

After the mission definition has been selected and design teams begin to form numerous preliminary designs, the focus should move to communication tools with high symbol variety and low permanency. This transition will help to facilitate an exchange of technical data required during the conceptual design and reduce the amount of relational conflict [43]. For design teams, research has shown that teams which sketch and generate more ideas during the conceptual stages are more likely to be successful [81–83]. Web conferencing then becomes a useful communication tool, as it allows for sketches and concepts quickly drawn on a computer to be shared with team members

Recommended tools for early stages

In summary, for the early stages of a virtual team, we recommend that extra consideration be given to tools with high media richness, multiple symbol types, low time to response and low parallelism. To achieve these communication types, we recommend the following communication tools:

- Face-to-Face (Kickoff) meetings
- Video Conferencing
- Web Conferencing

These communication tools should be taught and employed by virtual design teams to develop a good working relationship, form a shared mental model of the task and team skills, and generate the most ideas possible.

3.4.2 Middle Stages

As team members understand the task that needs to be accomplished, they narrow the design and select a specific concept. The middle stages begin after the concept has been narrowed to one specific design, and span product architecture development, configuration design, and detailed design. Work begins in earnest with specific integrated product teams (IPTs) prototyping sub-systems and incorporating them into the overall detailed design. Detailed design brings individual portions together, most system-level decisions are finalized, and a decision is made by team management to release the design for production [71].

Communication goals for this stage are to effectively communicate design schematics and details, coordinate component interfaces among multiple IPTs, and document design decisions. Assuming that sufficient levels of trust among team members were established during the early stages of the project, some forms of communication are less necessary during the middle stages. Researchers such as Golden & Raghuram and Doerry et al. found that once trust is high, mediums with high richness (such as face-to-face or video conferencing) are not as necessary and less expensive or more convenient mediums can be used more effectively in this stage [26, 31]. Thus, tools with less media richness should be exchanged for communication mediums that provide high permanency, high time to respond, and multiple symbol types.

Given the level of detail and number of decisions the team makes in the middle stages, it becomes important that discussions and decisions be automatically documented in a manner that facilitates easy review. The importance of being able to easily capture design rational has been highlighted by researchers such as Bracewell et al. [84] and Klein [85]. Hepworth et al. found that virtual teams that use a documented shared list of tasks that all members can access and edit simultaneously are able to reduce confusion and increase performance compared to virtual teams without such a tool [86].

For these reasons, all team members should be made familiar with tools that allow a shared database like Google Drive, MS Sharepoint, or other similar cloud storage systems. Hackman states that an information system is critical to the group's ability to plan and execute a performance strategy [54]. The permanency of a shared database allows team members to document design decisions, reference other team members' work, and stay up to date on the progress of others.

In AerosPACE, we have explored different options for such an online collaborative platform. During the 2013-2014 and 2014-2015 program years, students mentioned their dissatisfaction with the chosen tool, in large part because of the poor file organization capabilities. Students expressed the desire to use a tool such as Google Drive, which would allow them to organize, share, and search files as they wished. However, because of security protocol, access to Google Drive was restricted for team coaches from Boeing. To remedy this situation, MS Sharepoint, which can be federated via security protocols, was implemented in the 2015-2016 program year allowing teams to create and share items such as Gantt charts and task lists as well as organizing files and folders. An example of one team's use of the system can be seen in Figure 3.2. This screenshot shows how any team member can, in one central location, access schedule information, find files, and view the latest information posted by teammates and faculty on the Newsfeed. In the upper half of the screenshot is a Gantt chart showing the schedule for all members of the team. The lower left shows where the files are stored and organized, with recent posts in a forum-like format in the lower right side. Other apps can be added or removed as the team chooses. In addition to MS Sharepoint, edX was also used as an online organizer for all classroom material (e.g. lecture recordings, slides, handouts). Although edX allowed for seemingly better organization and easier access, students neglected edX, preferring instead one all-encompassing tool that would house their classroom needs. This highlighted the fact that ease of access was critical in selecting an online collaborative platform.

Time to response also becomes an important factor during this stage. In the middle stages, the work a design team performs is often technically complex. Scheduling challenges imposed by working across time zones and varied individual schedules, also adds to the difficulty of com-

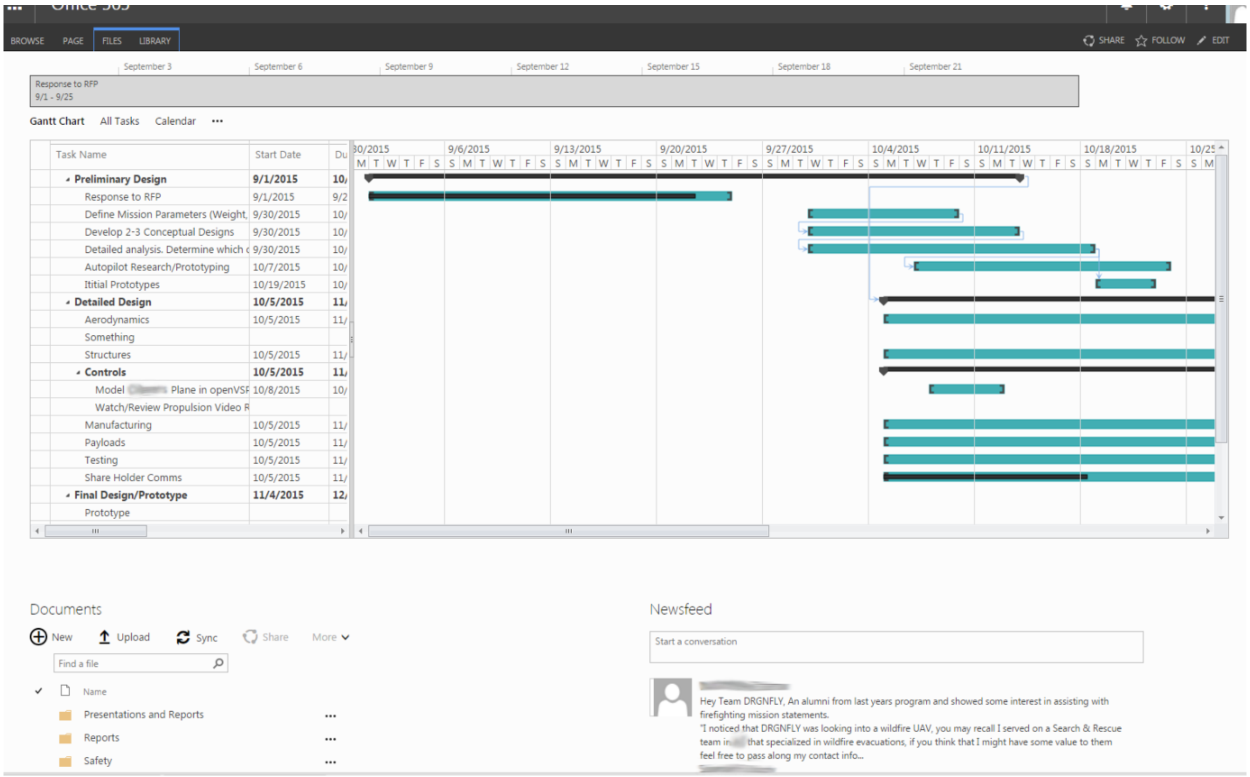

Figure 3.2: A screenshot of a team page showing a Gantt chart, file-folder organization system, and team discussion thread.

municating simultaneously. For these reasons, AerosPACE students indicated in the survey that they preferred tools that allowed for a longer time to response, such as texting, email, or shared databases for the transfer of information. These types of tools allow students to receive a message, think about the implications of that message, and then respond appropriately when most convenient. In one instance, a student received a question from a faculty coach, thought about what the question entailed, took the time necessary to perform the required calculations, and answered the question. The use of having a high time to response is also backed up by Maruping and Agarwal, who propose that teams that use communication tools with a high time to response for managing process conflict will be more effective than teams that use communication tools with other functionalities [15].

Technologies that allow for multiple symbol types are important during this stage of the project. Databases, as mentioned before, and email are useful for design teams as they allow for almost any symbol type to be transmitted. While these technologies are typically low in media richness, they focus on the substance of group tasks. These features allow for multiple team members to simultaneously participate in the task discussion and decision making process and generally gives an improved decision quality [87,88]. By thus using email or shared databases, design teams will be better equipped to share the specifics of their designs, make high quality decisions, and document design decisions.

Recommended tools for middle stages

In summary, for the middle stages of a virtual team, we recommend using tools that are high in permanence, high in symbol variety, and high in time to respond. As such, the following communication tools are recommended:

- Web Conferencing
- Shared Database
- Email

These tools will help a team to document design decisions, collaborate between IPTs about sub-system interfaces, and effectively share the technical data generated from their design work.

3.4.3 Late Stages

The late stages include early full prototyping, testing, and final manufacturing. This portion of the project is characterized by a distinct shift from digital to physical work. Small changes are sometimes made to the design, but for the most part the team is now focused on manufacturing and full-scale testing. Physical parts are shipped to and from different locations for assembly and testing, concluding with a final assembly and showcase of the finished product. Communication tasks during this stage are reviewing design decisions, communicating fulfillment of part production, and any coordination of design changes and shipment of parts.

During the late stages, communication tools that provide medium or high permanence and low time to response allow team members to recall exact values and specifications of the design. For most teams in AerosPACE, databases were used to reference manufacturing plans, design specifications, and test results. The high permanence of databases makes them ideal for this stage, as team members may feel confident that specifications and results are consistent across all IPTs. To notify other teammates of design changes, email was generally used. These results generally agree with the findings of other researchers, such as Maruping and Agarwal, who hypothesize that virtual design teams during the later stages of development should use communication mediums that are low in time to response and high in permanence [15].

Although analysis and testing are relatively easy to document in a database for remotely located teammates to view, physical work is not. To overcome this challenge, AerosPACE teams have developed interesting techniques. For example, students in one team would use their phones to take photographs of their work and post them on social media with a short caption, such as "Brand spankin' new carbon fiber ribs!" as shown in Figure 3.3. The accessibility and permanence of social media allow remote teammates to observe their work, identify potential errors, and offer suggestions, with minimal effort in recording. The entire team is thus aware of the team's progress and can adjust their plan accordingly.

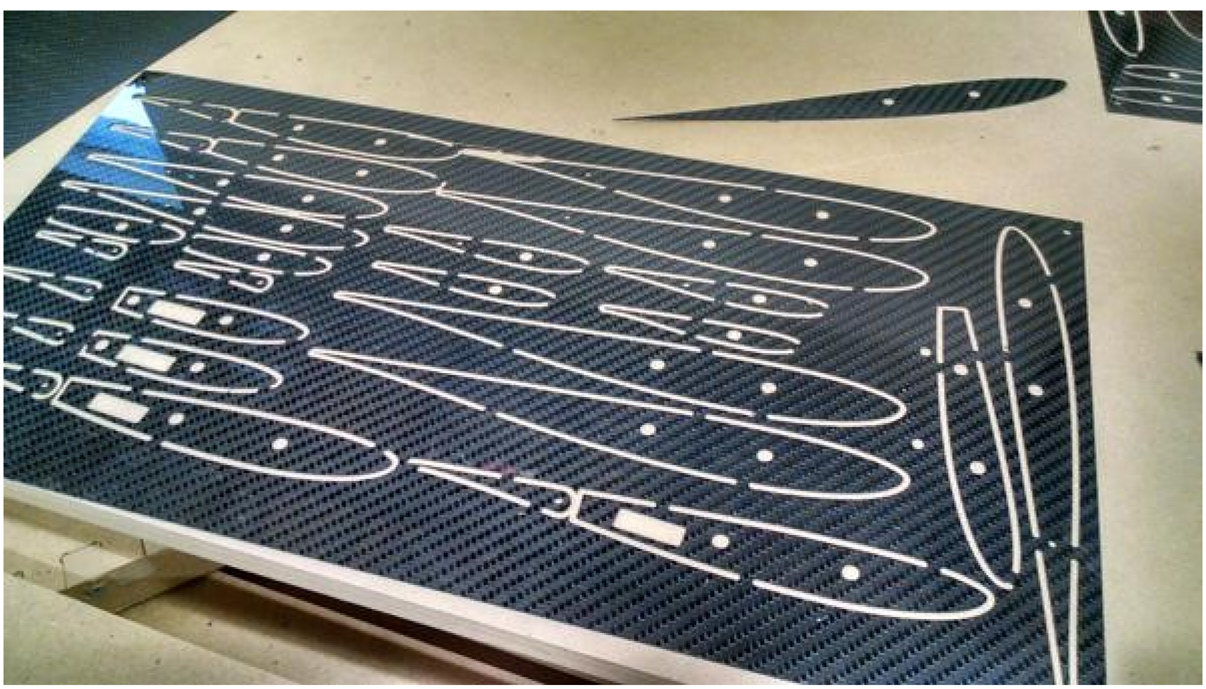

Figure 3.3: Photo of carbon fiber ribs sent over social media to show manufacturing progress.

Communication tools that are high in parallelism and accessibility are necessary to coordinate the efforts among multiple manufacturing groups in a timely manner. Easily accessible communication, such as texting, allows double-checking of numbers and dimensions before the final manufacturing and shipment of parts [89]. The accessibility of texting is high and the ability to easily look back at messages sent and their order demonstrates its high permanency. Furthermore, the high parallelism of texting allows for coordination efforts among multiple IPTs to occur at the same time with little effort. Figure 3.4 shows two examples from AerosPACE of messages exchanged during the late stages of the product development process. The conversation on the left shows two teammates quickly verifying a design decision. The conversation on the right shows one teammate commending the efforts of another, and then later verifying a critical dimension. Both are good examples of how texting allows for more efficient coordination due to the high accessibility and parallelism of the tool.

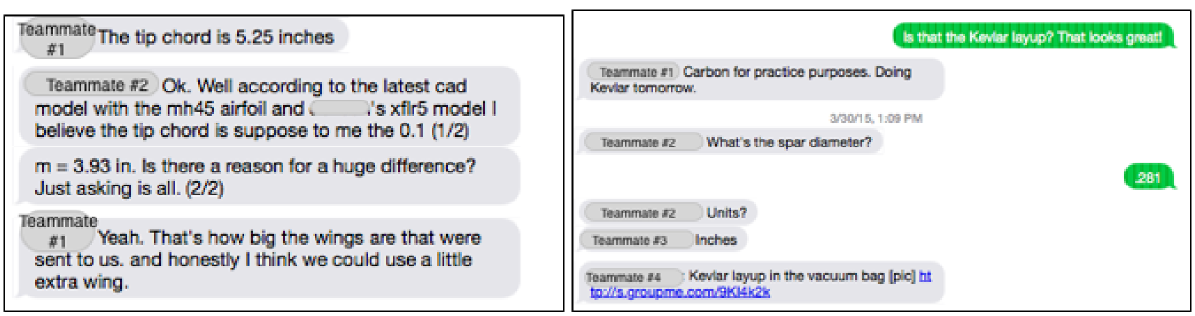

Figure 3.4: Examples of group text messages among teammates during the late stages of the product development process.

Finally, our experience has shown that in the final stages of the project, it is highly effective to allow at least some teammates from different locations to work on-site with their teammates. This idea is supported by researchers such as Hinds and Bailey, and Grinter et al. [90, 91]. For example, after one unsuccessful flight attempt which ended in a crash, one team with students from multiple universities used GroupMe to see who would be willing to help fix the UAS. Since most of the team was on-site, they were able to regroup and rebuild the plane within 24 hours of the crash. This feat was accomplished in part by the colocation of the team and the use of texting.

Recommended tools for late stages

In summary, for the late stages of a virtual team, we recommend that extra consideration be giving to communication tools that are low in time to response, high in permanence, high in parallelism, and high in accessibility. As such, we recommend using the following communication tools:

- Shared Database
- Email
- Text Messaging
- Face-to-Face Meeting

These communication tools are essential for allowing a team to quickly access and verify design parameters during manufacturing and integration of parts in the completed model. Face-to-face meetings allow for teams to quickly fix or modify any parts that don't adequately fulfill the need of the team.

3.5 Conclusion

Virtual design teams face significant challenges, not only in learning all thats necessary to complete their projects, but in learning more about effective engineering communication. This research identifies, through a review of the literature and the experience of the authors with several years of multi-university, multi-disciplinary capstone projects, which remote collaboration tools should be used at different stages of product development and production. We evaluate collaboration tools based on several different criteria, including richness, symbol type, time to response, permanence, parallelism and accessibility. Engineering students should be instructed on how to use and when to use the various communication tools available when designing and manufacturing a product.

CHAPTER 4. USING PEER ASSESSMENTS TO PREDICT TEAM EFFECTIVENESS AMONG DISPERSED, MULTI-DISCIPLINARY CAPSTONE DESIGN TEAMS: A LONGITUDINAL STUDY

4.1 Introduction

Learning within teams has become increasingly popular in higher educational settings [1]. The Accreditation Board for Engineering and Technology (ABET) requires applied science, computing, engineering, and engineering technology programs to have some sort of team experience in order to be officially accredited [2] and studies have shown that there is a positive relationship between team learning for engineering students and course performance [3]. An important question, however, is how to predict the performance of teams within engineering education. One method that has been gaining support is peer assessment.

Peer assessment (or peer ratings, peer evaluations) is typically defined as an assessment methodology that allows team members to provide input into the assessment procedure through evaluating each others' performance in out-of-class interactions and activities [17]. Using peer ratings has long been known as useful in predicting success in the military [18], and other researchers have extensively applied peer evaluations to educational courses, such as speech communication and business [19–21]. Studies have also been conducted within engineering teams. Ohland et al. found no difference in peer ratings based on either gender or race, of those evaluated, among freshman electrical engineering students [22]. Northrup & Northrup conducted a study on multi-disciplinary teams and found that there isn't enough evidence to assume that a dysfunctional team will have a lower average team peer rating [23]. While these studies have added to the understanding of using peer ratings with engineering teams (some multidisciplinary) there is still a need of further research in examining peer ratings among *dispersed* engineering teams.

Dispersed teams (or virtual teams) are defined as teams that rely on technology-mediated or virtual communication technologies to work at a distance from their fellow team members [5]. The growth of these teams, is largely attributed to a confluence between technology and organizational developments [6] and will likely continue to grow. One example in the commercial aerospace industry, is Boeing's 787. Sixty-five percent of the new Dreamliner is supplied to Boeing by dozens of companies located around the globe [63]. Engineers from supplier companies and Boeing are thus required to work together at unprecedented levels. Dispersed teams have been shown to have some benefits over traditional face-to-face teams, such as making more effective decisions or generating more unique and high quality ideas [5]. In contrast to the benefits, there are some disadvantages. For example, Thompson and Coovert found that in a comparison between face-to-face teams and virtual teams, virtual teams reported greater confusion and less understanding of team discussions [11].

Furthermore, the need for understanding how peer assessments are similar or different for a dispersed team than a team that is collocated, is ever increasing. For example, while peer ratings have been shown to work well with students that interact face-to-face, how will it perform when peers only communicate through phone calls, video conferencing, or a chat room? Since virtual teams hardly (if ever) meet face-to-face, some researchers have argued that these types of teams struggle to develop strong relational bonds that form the foundation of group identity, cohesiveness, and trust [7–10]. This could make it difficult for dispersed team members to accurately assess the performance of their peers. Others have argued the opposite, stating that while initial levels of trust suffer in virtual teams, trust grows over time to a level comparable with face-to-face teams (e.g. [92]). Furthermore, Berry suggests that peer evaluation is more accurate with dispersed teams since they are less confounded by personality or other non-task behaviors and actions as with collocated teams; team members can then be judged on what they actually accomplish rather than what they appear to be doing [93].

The most common use of peer assessments is at a single point in time, such as at the end of a team life-cycle or when a project has been completed. Channon et al. performed peer assessments over time, but looked at the assessments as separate data points, rather than a pattern over time [94]. A longitudinal study was performed by Tasa et al., but peer assessments were only conducted once, toward the end of the project [95]. There appears to be a need to understand how peer assessments behave *over time*.

There is some conflicting evidence as to whether peer assessments are unidimensional or multidimensional. Chalupa et al. studied 92 students at a Midwestern university and concluded that peer assessments are most aptly used as a unidimensional measure [41]. Other researchers have corroborated this idea (e.g. [94,96,97]). Paswan et al., on the other hand, found in their study of 54 undergraduate students that peer evaluation is better suited as multidimensional [44]. This point should be examined a bit closer, especially when peer assessments are used over a length of time.

The purpose of this paper is to understand the relationship between peer assessments and team effectiveness over a period of time. The Aerospace Partners for the Advancement of Collaborative Engineering (AerosPACE) [67,98] senior capstone design course was studied. AerosPACE is a multi-disciplinary design course sponsored by Boeing. Teams of 10-12 students are tasked with working together to design and build an unmanned aircraft system (UAS) that will complete a chosen surveillance mission. Each team is given some latitude to select the specifics of the mission (e.g. what to survey, flight time, and configuration of the drone), while staying within some guidelines (e.g. hand launched, fixed wing, within a predetermined budget, and within the time of the course). Each team is comprised of students from multiple universities and required to manufacture at least one part at each location, creating a genuine dispersed team environment. The project (which lasts for two semesters) culminates in a final fly-off event where teams gather together at one location to showcase and fly their completed UAS in front of the university faculty and Boeing.

There is a need to further research and contribute to this space, due to the lack of research on peer assessments within dispersed teams over time. Our unique contribution to this field of study is that it conducts a longitudinal study on the relationship between peer ratings and team effectiveness over time, within a dispersed, multi-disciplinary design team.

4.2 Background

This section will give a background to the pertinent literature for the current study. An in-depth look will show how peer assessments have been used in the past, specifically with an exploration into the use of these ratings in longitudinal studies. Following the review of peer assessments, background information on team effectiveness will be given.

4.2.1 Peer Assessment

Peer assessment has been defined over the years as the process of having the members of a group judge the extent to which each of their fellow group members exhibit specified traits, behaviors, or achievements [38]. While there are many different ways to administer peer evaluations, most tend to closely follow one of two methods. The first involves a set of questions which ask team members to rate their peers on topics such as attendance at meetings, quality of work, or easiness to work with [39, 40]. This method has been used primarily because of the high reliability of the ratings [41].

The second method for peer assessment, as stated by Oakley et al., is a single rating on the contribution of each team member to the final project [42]. The question may vary, but often asks students to rate their peers from ”Excellent” (defined as someone who consistently went above and beyond the call of duty) to ”No show” (defined as someone who had no participation at all) [43]. When using this type of peer evaluation, team members will often be asked a series of questions to get them thinking about their peers performance before being asked to give the overall rating [42]. This form of rating has been used considerably throughout the years for its ease of administration.

In the case of AerosPACE, the average team size is close to eleven students per team. If the multiple set questionnaire were used, students would be asked to consider over 100 different questions when evaluating their peers. This could potentially cause rater fatigue, or simply be more than the students are willing to answer. The single question rating would seem to be ideal for this setting, but due to a study conducted by Paswan et al., in which it was argued that peer assessments are multidimensional [44], it was determined that a multi-question assessment would be most viable, but that a test should be performed to verify whether the questions are indeed multidimensional or unidimensional.

Gueldenzoph & May suggest that the general criteria used for peer evaluations fit into five categories: commitment to the group, ability to deal constructively with conflicts that arise, active participation in the decision-making process, accountability for assigned tasks, and assumption of initiative or leadership role [17]. Dyer et al. suggest that successful teams should have adequate interpersonal skills, technical expertise, and motivation [45]. For the AerosPACE course, the peer evaluation instrument was selected based on the recommendations of Dyer et al., since the elements from other researchers' peer assessments readily fit into one of these three categories (e.g. [17, 39, 41]). Each team member was asked to rate his/her teammates based on their overall motivation, technical skills, and social skills. Validation was performed to ensure that these three questions would be a valid measurement tool, which will be shown hereafter.

4.2.2 Team Effectiveness

There are several different ways to measure success in team performance, but perhaps the most influential and widely cited definition comes from Hackman [46]. He states that team effectiveness is comprised of three main components: first, the degree to which the group's productive output meets the standards of quantity, quality, and timeliness; second, the degree to which the process of carrying out the work enhances the capability of members to work together interdependently in the future; third, the degree to which the group experience contributes to the growth and individual learning of team members. While all three of these are important factors, this study focused on the first two items (later referred to as task effectiveness and relational effectiveness), due to the focus of literature on these two areas and the interests of the authors.

A vast amount of team research has been put into measuring a team's operational, quantitative, and qualitative objectives [47] or in other words a team's effectiveness on a given task. Levi et al. argue that "from a management perspective, the obvious definition of team success is successful performance on a task" [48]. This is usually a popular method of measurement due to its objective nature [49]. In academia, task effectiveness is often measured by performance on assignments, presentations, or projects. Other criteria such as overall GPA, grades assigned in a class, or performance on a project are also commonly used [3, 50, 51].

Relational effectiveness of a team has also been a popular measure of team outcomes [52, 53]. Since the main criteria for this metric is the ability of the team to continue working together in the future [16, 54], items such as trust, commitment, and satisfaction play a key role in defining success. When studying this area, Whitman et al. examined the difference in satisfaction of a computer-mediated communication design team (dispersed team) verses a face-to-face design team, and found that dispersed teams are generally less satisfied than their face-to-face counterparts [51]. Furthermore, Hinds & Mortensen have found that distributed teams have greater difficulty in developing trust [55]. Researchers have also studied the emergence of trust within teams and concluded that it is closely tied to team effectiveness [56, 57]. It is important to note, however, that while relational effectiveness is important, it should not be confused for *task* effectiveness, nor used as a catch-all for a team's overall success. As Hackman put it, if a team is satisfied and yet unsuccessful in producing a quality project, then in most instances the group would be considered ineffectual [46].

In the AerosPACE program, task effectiveness was defined by looking at team scores on design reviews and whether or not the team succeeded in flying their UAS at the end of the project. Relational effectiveness was defined as the satisfaction of each student with their team over time and the amount of trust within each team at the end of the project. By including both task and relational measures, it gives an objective view of success through looking at how well each team performed on the course project, while also incorporating a subjective measure of the ability of each team to continue working well together.

4.2.3 Purpose (Hypotheses)

While many studies have researched the use of peer evaluations, there is a need for more understanding on the change of peer assessments over time and how that change relates to the success of a team. Specifically, more research could be done to examine how peer assessments behave over time and whether they are unidimensional or multidimensional. Based on the research by Paswan et al. [44], the following hypothesis is proposed:

- Hypothesis 1: Each indicator used to measure peer assessment (motivation, technical skills, and social skills ratings) will be multidimensional over time.

Several studies have been conducted to examine the relationship between task performance in a higher education setting and peer ratings. Dingel et al., for example, studied 101 students in an introductory sociology course and concluded that students' evaluations of their peers do not accurately predict the students' course grades [1]. A year later, however, a similar study was published by the same authors, which found that peer evaluations can indeed predict course performance [50]. The majority of literature appears to corroborate the later study. For example, a study of 287 students enrolled in principles of management courses found that there is a significant relationship between final course score and peer evaluations [70].

Based on these and other studies, the following hypotheses are proposed:

- Hypothesis 2a: There is a significant, positive relationship between motivation ratings and task effectiveness.
- Hypothesis 2b: There is a significant, positive relationship between technical skills ratings and task effectiveness.
- Hypothesis 2c: There is a significant, positive relationship between social skills ratings and task effectiveness.

Several researchers have also studied the relationship between peer evaluation and relational effectiveness. Among nursing students studied in 2010 and 2011, Roh et al. found that there is a positive significant correlation between peer evaluation and learner satisfaction [99]. Paswan et al. have shown that in a typical undergraduate team, overall group satisfaction related positively with peer ratings [44]. Ohland et al. have shown that contributing to the team's work, having sufficient knowledge, skills, and abilities, and interacting with teammates positively correlates with wishing to continue working with the team in the future [100].

Based on these studies, the following hypotheses are proposed:

- Hypothesis 3a: There is a significant, positive relationship between motivation ratings and relational effectiveness.
- Hypothesis 3b: There is a significant, positive relationship between technical skills ratings and relational effectiveness.
- Hypothesis 3c: There is a significant, positive relationship between social skills ratings and relational effectiveness.

4.3 Methods

In the methods section, a detailed explanation of the experiment performed will be provided. Information given in this section will include types of participants included in the study, measurements used, and the statistical procedures employed for analysis of the data gathered.

4.3.1 Participants

This study included 150 students, consisting of 128 males and 22 females enrolled in several different universities and colleges across the United States. Participating institutions were Brigham Young University, Georgia Tech University, Purdue University, Tuskegee University, Embry Riddle Aeronautical University, Clemson University, Washington State University - Pullman, Washington State University - Everett, and Everett Community College. The average age of the students was 22.9 years old (SD = 2.49). The discplines included in this program were mechanical engineering, aerospace engineering, manufacturing engineering, electrical engineering, and computer engineering majors. The project lasted for two semesters (late August through April) and teams worked together for the duration of the project. Surveys were administered four times throughout the project at the midpoint and end of each semester (approximately 6, 14, 23, and 30 weeks). Students were assured that responses were kept anonymous, wouldn't be shared with other students, and wouldn't affect students' grades. This was done to ensure honest and accurate results [70].

4.3.2 Measures

The three indicators for peer assessment were measured on a Likert scale of one to five, with one corresponding to "Very Poor", two corresponding to "Poor", three corresponding to "Fair", four corresponding to "Good", and five corresponding to "Very Good". Each respondent was primed with a definition of the factors to ensure that interpretation of each category was consistent from student to student. These prompts may be found in Appendix I. Since the number of students on each team varied, the peer assessment scores for each team member was averaged for analysis [41].

To guarantee that students were rating their peers in a consistent manner, inter-rater correlation was calculated. The range for the correlation statistics was 0.72 to 0.84 across all time periods. This is consistent with the results found in the meta-analysis performed by Ohland et al. [101], which showed that for a single item peer rating, a multiple item peer rating, and a behaviorally anchored peer rating (single-item), the inter-rater correlations were 0.67, 0.74, and 0.78, respectively. The entire list of each reliability and standard deviation for the current study is shown in Table 4.1.

Table 4.1: Inter-rater Reliability and Standard Deviation

Variable	IRR	SD
Motivation		
Time 1	0.81	0.11
Time 2	0.84	0.069
Time 3	0.81	0.092
Time 4	0.84	0.069
Technical Skills		
Time 1	0.78	0.10
Time 2	0.81	0.12
Time 3	0.78	0.15
Time 4	0.80	0.10
Social Skills		
Time 1	0.73	0.16
Time 2	0.75	0.11
Time 3	0.72	0.14
Time 4	0.72	0.18

Satisfaction was gathered during each survey, similar to previous studies [44,102]. Students were asked to think of their team as a whole and then rate how satisfied they were with their team on a Likert scale of one to five, with one defined as “Very Dissatisfied”, two defined as “Dissatisfied”, three defined as “Neutral”, four defined as “Satisfied”, and five defined as “Very Satisfied”. Students’ satisfaction scores were coded as nominal. While the options always remained constant (five options from “Very Dissatisfied” to “Very Satisfied”), the gathered responses contained a different number of bins (response categories) at the different time points. For example, in one survey students gave responses for “Neutral”, “Satisfied”, and “Very Satisfied”, for three bins; in another students gave responses for “Dissatisfied”, “Neutral”, “Satisfied”, and “Very Satisfied”, giving four bins. Since the software used is unable to handle varying numbers of bins, some of the responses were collapsed into an adjacent bin. In each survey, less than 3% of the responses were changed to keep the number of bins consistent.

Team report score was measured by the grades students received during design reviews. The presentations were scored by professors using a rubric to ensure consistent ratings. Scores were continuous with a possible maximum score of one-hundred and a possible minimum score of zero. Two outliers were removed from the final model.

Success of flying the UAS at the end of the project was scored as a dichotomous (boolean) variable. One was used to represent a successful flight at the time of the final fly-off event, and zero was used as an unsuccessful attempt. It's important to note that flying a UAS on time was an important part of this score. While some teams were eventually successful in flying a UAS, if they failed to deliver their product on time, then they were marked as not having succeeded.

4.3.3 Statistical Model

Confirmatory factory analysis (CFA) was conducted to determine if the three peer rating indicators were unidimensional over time or multidimensional. The procedure followed is given by Muthen & Muthen [59] and further information on this procedure is given by Brown [61]. The steps taken are summarized as follows:

1. Test for unidimensionality of indicators at each time.
2. Determine the shape of the curve for each indicator.
3. CFA of all time-points together:
   (a) Covariance structure analysis without measurement parameter invariance.
   (b) Covariance structure analysis with invariant slope coefficients.
   (c) Mean and covariance structure analysis with invariant measurement intercepts and slope coefficients.

A model was created for each step and model fit statistics were used to determine whether each model adequately fit the data. Fit statistics and corresponding levels were used as follows: the chi-square test (with a non-significant p-value at the $\alpha = 0.05$ level indicating a good model fit), the root mean square error of approximation (RMSEA) value (with values less than 0.05 indicating a good fit), the comparative fit index (CFI) and Tucker-Lewis index (TLI)(with values greater than 0.90 indicating good model fit), weighted root mean square residual (WRMR)(with values less than 1.0 signifying good model fit) used for models with categorical or nominal data, and standardized root mean square residual (SRMR)(with a value less than 0.10 indicating an acceptable model fit) used for models with only continous data. These values for indicating good model fit were taken from Wang & Wang [58] and Worthington et al. [60]. It's important to note that although the chi-square test is reported for each model, it isn't given much credence due to its sample-size dependency and the large number of respondents in the current study [60].

A test for unidimensionality was performed first at each time-point by loading the indicators on a single latent variable. With only three indicators, a one-factor CFA model would be just identified and unable to provide model fit statistics. The residual variances were constrained to be equal, giving the needed degrees of freedom to calculate model fit statistics for this step.

Following the test for unidimensionality at separate time points, the shape of the model for each indicator over time was determined, in order to ensure a similar shape for each indicator. Each model was first assumed to be quadratic and tested to see if the model fits the data well and if the intercept, slope, and quadratic coefficients are significant. If the model fails or if the coefficients were not significant, then the highest order term is removed and the model is retested.

Once the shape for each indicator was found, a covariance analysis without measurement parameter invariance (often referred to as configural invariance analysis) was performed. This involved loading the indicators onto a latent variable, but not constraining any of the indicators across time. Model fit statistics were calculated to determine adequate fit. Following this model, the covariance structure analysis with invariant slope coefficients (commonly referred to as metric invariance analysis) was performed. For metric invariance analysis, the same model was used as the configural model, except the slope coefficients were constrained to be equal. Model fit statistics were again examined, along with a comparison to the configural model. The two statistics examined were chi-square difference test (with a non-significant p-value meaning the two models are invariant) and the change in CFI statistic (with a $\Delta\text{CFI} \leq 0.01$ meaning the two models are invariant). Following, the mean and covariance structure analysis with invariant slope and intercept coefficients (commonly known as scalar invariance analysis) was performed. This model is similar to the configural model, except the intercept and slope coefficients were constrained to be equal. Model fit statistics were then examined and the model was compared to the configural model. In each case, model invariance would mean that the three indicators loaded equally across time, while any variance between models would have shown a need to analyze the indicators separately.

Once the confirmatory factor analysis was performed, the structural equation models were created using Mplus. As seen in Figure 4.1, a model was created for task effectiveness and relational effectiveness with each peer rating indicator in turn. In this figure, x represents the independent variables (each peer rating indicator) at the various time points. The variable y_1 represents the gathered data for the dependent longitudinal variables (report score or satisfaction) at the various time points and y_2 represents whether the team flew their UAS at the end of the project (used in determining task effectiveness) or trust (used in determining relational effectiveness). The variable ξ represents the intercept and slope of the linear regression for the independent variables, and η represents the intercept and slope of the linear regression for the dependent variables.

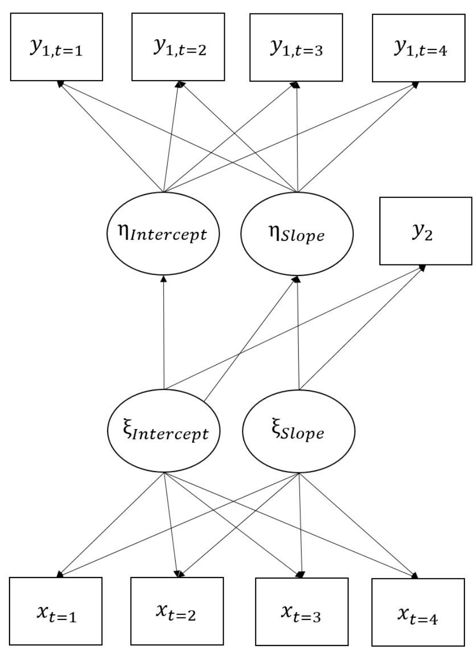

Figure 4.1: Structural Equation Model Map for Task and Relational Effectiveness.

Again, to determine whether the model adequately represents the data, model fit statistics were examined. Once an appropriate model was identified with good fit statistics, the significance between the proposed relationships was examined. A two-tailed p-value was employed with an α level of 0.05 or less indicating a significant relationship.

4.4 Results

The results of the previous statistical analysis performed on the data collected will now be presented. The confirmatory factor analysis will be presented first, followed by the model for peer ratings with task effectiveness, and finishing with the model for peer ratings with relational effectiveness.

4.4.1 Confirmatory Factor Analysis

The results of the confirmatory factor analysis are provided first. This section will be divided by first looking at the unidimensionality of the indicators for each time-point, then determining the shape of the growth curve of each model, and then examining the results of each model.

Unidimensionality of Indicators for Each Time-point

The model fit statistics for each time-point were adequate in assuming that the three factors have unidimensionality (as seen in Table 4.2). CFI, TLI, and SRMR are all adequate for each time point, with RMSEA and χ^2 showing good model fit at time-points three and four. Since the majority of the model fit statistics indicate goodness of fit for times one and two, it's reasonable to assume that the three indicators have unidimensionality at each time.

Determine the Shape of the Growth Curves

Each indicator was explored separately to determine the shape of each growth curve and was first assumed to be quadratic. Model fit statistics were examined to determine whether the model fit the data well, and the intercept, slope, and quadratic coefficients were reviewed to see if they were significant at the $\alpha = 0.05$ level.

Table 4.2: Model Fit Statistics at Each Time Step.

	CFI	TLI	RMSEA (90% C.I.)	SRMR	χ^2 (sig)
Time 1	0.98	0.94	0.25 (0.13, 0.40)	0.054	10 (0.0012)
Time 2	0.98	0.94	0.19 (0.070, 0.40)	0.054	6.3 (0.012)
Time 3	1.0	1.0	0.0 (0.0, 0.15)	0.006	0.093 (0.76)
Time 4	1.0	1.0	0.0 (0.0, 0.15)	0.004	0.086 (0.77)

A quadratic curve for motivation ratings was examined first with good model fit statistics (χ^2 = 0.004 with p-value = 0.95, CFI and TLI = 1.0, RMSEA = 0.0 with a 90% confidence interval of (0.0, 0.0), and SRMR = 0.001). The means and p-values (seen in Table 4.3) show that only the intercept value was significant at the $\alpha = 0.05$ level.

Table 4.3: Motivation Quadratic Growth Curve Coefficients.

Means	Estimate	S.E.	P-Value
Intercept	4.2	0.10	0.0
Slope	-0.004	0.070	0.95
Quadratic	-0.017	0.014	0.21

The quadratic term was next dropped from the model and the analysis again performed. The model fit the data well with good model statistics (χ^2 = 6.1 with p-value = 0.30, CFI and TLI = 1.0, RMSEA = 0.038 with a 90% confidence interval of (0.0, 0.12)). Both the intercept and slope coefficients in this model (as seen in Table 4.4) were significant at the $\alpha = 0.05$ level. The shape of motivation ratings over time was thus determined to be linear.

Table 4.4: Motivation Linear Growth Curve Coefficients.

Means	Estimate	S.E.	P-Value
Intercept	4.3	0.069	0.0
Slope	-0.086	0.026	0.001

A quadratic model for technical skills ratings was examined next. Mplus was unable to find a model and stated that the problem was due to the quadratic term of the model. Including a quadratic term in the solution caused the model to become essentially singular, preventing the software from converging. The quadratic term was removed and the model calculated again. This time the model converged and found an acceptable model with good fit statistics (χ^2 = 5.9 with p-value = 0.32, CFI and TLI = 1.0, RMSEA = 0.035 with a 90% confidence interval of (0.0, 0.12), and SRMR = 0.051). The intercept and slope terms (as seen in Table 4.5) were both significant at the α = 0.05 level. The shape of technical skills ratings over time was thus determined to be linear.

Table 4.5: Technical Skills Linear Growth Curve Coefficients.

Means	Estimate	S.E.	P-Value
Intercept	4.1	0.046	0.0
Slope	-0.047	0.020	0.022

The quadratic model for social skills ratings was examined next with good model fit statistics (χ^2 = 0.30 with p-value = 0.58, CFI and TLI = 1.0, RMSEA = 0.0 with a 90% confidence interval of (0.0, 0.18), and SRMR = 0.007). The intercept coefficient (seen in Table 4.6) was significant at the α = 0.05 level and the slope was significant at the α = 0.1 level. The quadratic term was not significant and so was removed from the model.

Table 4.6: Social Skills Quadratic Growth Curve Coefficients.

Means	Estimate	S.E.	P-Value
Intercept	4.1	0.12	0.0
Slope	-0.16	0.094	0.097
Quadratic	0.017	0.018	0.361

For the linear growth model for social skills ratings, the model fit statistics showed a good fit (χ^2 = 2.4 with p-value = 0.78, CFI and TLI = 1.0, RMSEA = 0.0 with a 90% confidence interval of (0.0, 0.075), and SRMR = 0.032). Both of the coefficients for the intercept and slope (as seen in

Table 4.7) were significant at the $\alpha = 0.05$ level. The shape of social skills ratings over time was thus determined to be linear.

Table 4.7: Social Skills Linear Growth Curve Coefficients.

Means	Estimate	S.E.	P-Value
Intercept	4.1	0.073	0.0
Slope	-0.075	0.032	0.021

Since the shape for each indicator is the same, then the confirmatory factor analysis may continue to the next step, looking at the analysis of all time-points together as a longitudinal data set.

Confirmatory Factor Analysis of all Time-points

The configural model was created, measuring the overall peer assessment score by the three indicators. The model fit statistics showed a poor model fit ($\chi^2 = 540$ with p-value = 0.0, CFI and TLI equal to 0.78 and 0.69 respectively, RMSEA = 0.26 with a 90% confidence interval of (0.24, 0.28), and SRMR = 0.07), with the exception of SRMR being less than 0.10. Model modification was avoided as it has been shown to give inconsistent results, although it can help to improve model fit [60, 103]. With the majority of the fit statistics not meeting their required values, the model was rejected. The subsequent steps in the CFA became irrelevant and the three indicators were considered separately. This supports Hypothesis 1, that the indicators used to measure peer assessments should be considered separately.

4.4.2 Peer Ratings with Task Effectiveness

Next will be an examination of the peer ratings models and how they may be used to predict task effectiveness. The model for motivation ratings will be given first, followed by technical ratings, and then social skills ratings.

Motivation Ratings

Hypothesis 2a (there is a positive, significant relationship between motivation ratings and task effectiveness) was examined first. The model fit statistics showed good model fit (χ^2 = 18.6, with df = 18, and p-value of 0.42, RMSEA = 0.015 with a 90% confidence interval of (0.00, 0.075), CFI and TLI were both calculated at 0.99, and WRMR = 0.47). The results for the predictive relationship between motivation ratings and report score along with motivation ratings and whether the team flew or not (summarized in Table 4.8) was not significant at the α = 0.05 level.

Table 4.8: Correlation Statistics for Task Effectiveness and Motivation Ratings.

Variable	Estimate	S.E.	P-Value
Report Score Intercept			
Motivation Intercept	-0.83	0.80	0.30
Report Score Slope			
Motivation Intercept	0.41	0.48	0.40
Motivation Slope	1.3	1.3	0.30
Flew?			
Motivation Intercept	-0.22	0.25	0.39
Motivation Slope	-1.7	1.6	0.28

Technical Skill Ratings

For hypothesis 2b (there is a positive, significant relationship between technical skills ratings and task effectiveness) the model fit statistics showed a good model, with similar results to the model for motivation ratings and task effectiveness (χ^2 = 21, with df = 18, and p-value = 0.27, RMSEA = 0.034 with a 90% confidence interval of (0.00, 0.083), CFI and TLI were calculated at 0.98 and 0.97 respectively, and WRMR = 0.50). The results for the predictive relationship between technical skills and report score (summarized in Table 4.9) were not significant at the α = 0.05 level. Similarly, the results for the relationship between technical skills and whether the team flew or not were not significant at the α = 0.05 level.

Table 4.9: Correlation Statistics for Task Effectiveness and Technical Skills Ratings.

Variable	Estimate	S.E.	P-Value
Report Score Intercept			
Technical Intercept	-0.35	0.86	0.69
Report Score Slope			
Technical Intercept	0.39	0.55	0.47
Technical Slope	0.99	1.2	0.42
Flew?			
Technical Intercept	-0.11	0.25	0.67
Technical Slope	-1.6	1.9	0.39

Social Skill Ratings

For hypothesis 2c (there is a positive, significant relationship between social skills ratings and task performance) the model fit statistics showed a good model fit, with similar results to the previous two models (χ^2 = 23.5, with df = 18, and p-value = 0.17, RMSEA = 0.045 with a 90% confidence interval of (0.00, 0.091), CFI and TLI were calculated at 0.96 and 0.93 respectively, and WRMR = 0.59). The results for the predictive relationship between social skills ratings and report score along with the relationship between social skills and whether the team flew or not (summarized in Table 4.10) showed no significance at the α = 0.05 level.

Table 4.10: Correlation Statistics for Task Effectiveness and Social Skills Ratings.

Variable	Estimate	S.E.	P-Value
Report Score Intercept			
Social Intercept	-1.1	0.92	0.23
Report Score Slope			
Social Intercept	0.43	0.56	0.44
Social Slope	1.0	1.3	0.46
Flew?			
Social Intercept	0.78	0.70	0.27
Social Slope	-5.2	3.8	0.17

4.4.3 Peer Ratings and Relational Effectiveness

Finally, an examination of the peer ratings models and how they may be used to predict relational effectiveness is given. The model for motivation ratings will be presented first, followed by technical ratings, and then social skills ratings.

Motivation Ratings

The model for hypothesis 3a (there is a positive, significant relationship between motivation ratings and relational effectiveness) yielded excellent model fit statistics (χ^2 = 35.6, with df = 30, and p-value = 0.22, RMSEA = 0.035 with a 90% confidence interval of (0.00, 0.074), CFI and TLI were both calculated at 0.98, and WRMR = 0.50). The results for the predictive relationship between motivation ratings and satisfaction ratings (summarized in Table 4.11) were significant for the slope of motivation ratings and the slope of satisfaction at the $\alpha = 0.05$ level. This means that for every unit increase or decrease in the slope of motivation ratings, the slope of satisfaction can be predicted to increase or decrease by a factor of 0.76. There was no significant relationship between trust and motivation ratings, although it's valuable to note the suggestion of significance between the slope of motivation ratings and trust at the $\alpha = 0.1$ level.

Table 4.11: Correlation Statistics for Relational Effectiveness and Motivation Ratings.

Variable	Estimate	S.E.	P-Value
Satisfaction Intercept			
Motivation Intercept	-0.070	0.36	0.84
Satisfaction Slope			
Motivation Intercept	0.053	0.15	0.73
Motivation Slope	0.76	0.37	0.040
Trust			
Motivation Intercept	0.16	0.25	0.52
Motivation Slope	1.3	0.75	0.080

Technical Skill Ratings

Hypothesis 3b (there is a significant, positive relationship between relational effectiveness and technical skills ratings) was examined next. The model fit statistics showed a good fit (χ^2 = 38.3, with df = 30, and p-value = 0.14, RMSEA = 0.043 with a 90% confidence interval of (0.00, 0.079), CFI and TLI were both calculated at 0.97, and WRMR = 0.51). The results for the predictive relationship between technical skills and satisfaction (summarized in Table 4.12) were not significant at the α = 0.05 level. The same is true for trust and technical skills, although again it is valuable to note the significance between the slope of technical skills ratings and trust at the α = 0.1 level.

Table 4.12: Correlation Statistics for Relational Effectiveness and Technical Skills Ratings.

Variable	Estimate	S.E.	P-Value
Satisfaction Intercept			
Technical Intercept	-0.37	0.40	0.36
Satisfaction Slope			
Technical Intercept	0.069	0.14	0.62
Technical Slope	0.43	0.48	0.38
Trust			
Technical Intercept	0.069	0.25	0.79
Technical Slope	1.7	0.98	0.093

Social Skill Ratings

Hypothesis 3c (there is a significant, positive relationship between relational effectiveness and social skills ratings) was examined next. The model fit statistics showed a good fit (χ^2 = 38.2, with df = 30, and p-value = 0.14, RMSEA = 0.043 with a 90% confidence interval of (0.00, 0.079), CFI and TLI were calculated at 0.97 and 0.96 respectively, and WRMR = 0.49). The results for the predictive relationship between social skills ratings and satisfaction (summarized in Table 4.13) were significant between the slope of social skills ratings and the slope of satisfaction at the α = 0.05 level. This means that for every unit increase or decrease in the slope of social skills ratings, the slope of satisfaction can be predicted to increase or decrease by a factor of 1.3. The predictive relationship between trust and social skills ratings was not significant at $\alpha = 0.05$ level.

Table 4.13: Correlation Statistics for Relational Effectiveness and Social Skills Ratings.

Variable	Estimate	S.E.	P-Value
Satisfaction Intercept			
Social Intercept	0.20	0.50	0.69
Satisfaction Slope			
Social Intercept	-0.19	0.28	0.49
Social Slope	1.3	0.6	0.028
Trust			
Social Intercept	0.002	0.29	0.99
Social Slope	1.6	1.0	0.13

4.5 Discussion

Confirmatory factor analysis showed that the peer evaluations used in this study are unidimensional at individual times, as is consistent with the literature, giving credence to hypothesis 1. When examined longitudinally, however, peer evaluations were found to not react similarly over time and should thus be used as separate indicators.

Based on the results of the statistical models, there is enough evidence to suggest that a correlation exists between motivation and social skills peer ratings with the satisfaction of team members, suggesting that hypotheses 2a and 2c are accurate. Specifically, the change in motivation and the change in social skills over time significantly predicts the change in satisfaction over time; as a team increases or decreases in motivation and social skill peer ratings, the satisfaction of team members will increase or decrease proportionally. With a coefficient of 1.3, social skills seems to have a greater affect on satisfaction than motivation ratings does (0.76). Peer ratings failed to predict trust at the $\alpha = 0.05$ level, but it is important to note that the p-values were suggestive of significance at the $\alpha = 0.1$ level for motivation and technical skills ratings.

Technical skills ratings failed to predict satisfaction or trust at the $\alpha = 0.05$ level, which may have to do with the level of technical abilities of each team. This suggests that Hypothesis 2b is not true. In a recent study, it was found that the performance of a team increased with the level of talent, up to a certain point. At that point, the performance of the team began to level off and in some cases decrease [104]. In the AerosPACE program, students were selected based on their past experience with building and designing airplanes, and therefore have a high level of technical abilities and knowledge. Perhaps the same phenomenon occurred and each team had sufficient levels of technical expertise; thus, any increase or decrease in technical knowledge over the course of two semesters wouldn't have much of an influence on the effectiveness of the team.

None of the peer rating indicators significantly predicted task effectiveness, suggesting that Hypotheses 3a, 3b, and 3c are not true. The main reason for the lack of significance could be that students are giving evaluations based on effort, which does not necessarily correlate with performance. Wiggins et al. concur with this idea, stating that most peer ratings seem to be tied to the effort of each student rather than their performance [18].

Another interesting explanation may be seen in the comments of some of the students at the end of the project, who commented on the presence of "free riders". Free riding can be defined as any teammate (in this case student) that is seen to be giving less effort than his/her counterparts [107]. In AerosPACE, one student complained that half of the team "did absolutely nothing". Another student suggested that a team be created where "all the rejects can be sent to when teams feel they are being detrimental to the team's success". These quotes show some of the frustration students had with team members, even though the teams of these students successfully flew at the end of the project. The question then arises, were there free riders on any of these teams? Furthermore, is there any relationship between the presence of free riders and the satisfaction of the team?

A quick exploration was taken into free riders, and whether it affected the results of the team (either task or relational). An analysis was performed to see if any students were rated as having given less effort, or being less motivated (statistically), than their peers on the team. Of the fourteen teams, five were shown to have at least one free rider throughout the year. Three of those five teams were successful in flying a drone at the end of the year (60%), while seven of the nine teams with no free riders were successful in flying at the end of the year (78%). Furthermore, there is no statistical significance in the level of satisfaction between teams with a free rider present and teams with no free rider present. A summary of these statistics may be found in 4.14.

Table 4.14: Summary of Free Rider Statistics.

	No Free Riders	Free Riders
# of Teams	5	9
# of Teams that Flew	3	7
Mean Team Satisfaction	3.90	4.04

Dingel et al. corroborate this finding, concluding that the presence of such free riders doesn't significantly lower the quality of work within a team [1]; however, other researchers have shown that equally sharing the workload is significantly related to team member satisfaction [105, 106]. While members of a team may be unsatisfied with their peers and give them poor ratings, the quality of the work wouldn't significantly suffer. This seems to show why the poor satisfaction or presence of free riders doesn't negatively affect the performance of the teams, but also creates some interesting questions as to why the presence of free riders doesn't negatively affect the level of satisfaction.

CHAPTER 5. CONCLUSION

This research identifies, through a review of the literature and the experience of the authors with several years of studying over 150 multi-disciplinary students from multiple disciplines and universities, which remote collaboration tools should be used at different stages of product development. Collaboration tools have been evaluated based on several different criteria, including richness, symbol type, time to response, permanence, parallelism, and accessibility. By following the recommended pattern, virtual design teams will improve their communication efficiency during design and manufacturing projects.

Each stage of the product development process has unique needs that should be responded to with specific tools. Table 5.1 shows a summary of which tools we recommend should be given extra consideration during each stage. Certainly, there are different circumstances that merit a divergence from the proposed pattern, but this model gives a general outline upon which teams should base their communication.

During the Early Stages, teams will benefit most by holding a "kickoff meeting" or something similar at the beginning of the team formation process. Once face-to-face meetings become impractical, web conferencing and video conferencing should be used to help develop the team relationship and generate ideas. During the Middle Stages, the team should transition to web conferencing, email, and shared databases to help give permanence to design decisions. During the Late Stages, teams should rely more heavily on databases, email, text messaging and social media to verify design values and give updates on manufacturing progress. Where possible, teams should meet face-to-face in the Late Stages to integrate the several components into the final product and troubleshoot any problems. Following these steps will ensure the strengths of the tools being used best match the needs of the stage, increasing collaboration efficiency.

As more teams in the engineering industry continue to work remotely, it is important for academic institutions to prepare engineering students to collaborate successfully in this new envi-

Table 5.1: Suggested communication tools for each stage of the design process.

Tool	Early Stages	Middle Stages	Late Stages
Face-to-Face	X		X
Phone Call			
Teleconference		X	
Text/Instant message			X
Web Conferencing	X	X	
Video Conferencing	X		
Email		X	X
Shared Database		X	X
Social Media			X

ronment. This research shows that at different stages of a project, specific tools should be given extra consideration in order to optimize collaboration. Tools that better meet the needs of teams at a given stage will enhance team effectiveness and better prepare students to work effectively in a rapidly changing environment.

In addition to the suggested collaboration model given, this study looked to contribute to the body of knowledge of peer ratings. This study hypothesized that peer ratings would be multidimensional over time. Furthermore, it was hypothesized that even through the difficulties of developing relationships in a distributed team, peer evaluations would still predict team effectiveness. Results showed that there was a significant correlation between the change in satisfaction with the change in motivation ratings and the change in social skills ratings over time, but not with technical skills ratings. None of the three measures of peer ratings predicted task effectiveness.

It is reasonable to conclude that with a distributed design team, increasing or decreasing levels of motivation and social skills ratings will accurately predict increasing or decreasing levels of satisfaction for team members. This is significant due to the growing use of virtual teams and subsequent studies that show a lack of satisfaction among virtual team members [51]. This adds to the current body of literature by showing that teams that improve their levels of motivation to the project or level of social interaction will likely have improved levels of satisfaction with the team, regardless of whether the team meets face-to-face or virtually.

5.1 Future Work

Further research on dispersed design teams should continue to examine effective collaboration patterns for design and manufacturing teams. Some areas that would especially be useful to understand, within design team literature, and for the AerosPACE faculty and staff, would be the effect that certain demographics play on the effectiveness of the team. Results could be the same (did the team fly or not, what score did they receive on the presentations, and how satisfied are they with the team), but the inputs would change to the ratio of female team members to male, or the ratios of black, white, hispanic, asian, and other on the performance of the team. It would be best, in addition, for the researcher that continues along this line of inquiry to study deeply how to measure team satisfaction and relational effectiveness. Many papers found in this thesis would be a good starting point (start with Hackman's book *Groups That Work (And Those That Don't)* [46]). By better understanding if there's a correlation between these demographics and team members, this would certainly help future AerosPACE teams and faculty in being successful.

Furthermore, there has been some interesting qualitative data that suggests the presence of free riders or social loafing. Some research has shown that free riders don't affect the outcome of a project, while others have stated that it could negatively impact the satisfaction of a team. An interesting study, would be to delve further into the literature on free riders (specifically how to measure free riders), and determine whether the presence or perception of free riders is affected by the team meeting mostly virtually. A study conducted to find free riders and the affect it has on the team's satisfaction, would be beneficial for mitigating those affects in the future.

Finally, with the AerosPACE program being so unique, it provides an invaluable resource for studying dispersed engineering design teams. Future research of the AerosPACE program should indeed continue, but should be more focused, such as those described above; by only including one or two variables of study to research in conjunction with the AerosPACE program, researchers will have better success in understanding those concepts and how to apply the results in creating a better experience for future students.

REFERENCES

[1] Dingel, M. J., Wei, W., and Huq, A., 2013. "Cooperative learning and peer evaluation : The effect of free riders on team performance and the relationship between course performance and peer evaluation." *Journal of the Scholarship of teaching and learning,* **13**(1), pp. 45–56. 1, 35, 40, 57

[2] Stump, G. S., Hilpert, J. C., Husman, J., Chung, W. T., and Kim, W., 2011. "Collaborative learning in engineering students: Gender and achievement." *Journal of Engineering Education,* **100**(3), pp. 475–497. 1, 35

[3] Chin-Min Hsiung, 2012. "The Effectiveness of Cooperative Learning." *Journal of Engineering Education,* **101**(1), pp. 119–137. 1, 11, 35, 39

[4] OlsonBuchanan, J. B., Rechner, P. L., Sanchez, R. J., and Schmidtke, J. M., 2007. "Utilizing virtual teams in a management principles course." *Education + Training,* **49**(5), pp. 408–423. 1

[5] Martins, L. L., Gilson, L. L., and Maynard, M. T., 2004. "Virtual Teams: What Do We Know and Where Do We Go From Here?." *Journal of Management,* **30**(6), pp. 805–835. 1, 35, 36

[6] Solomon, C. M., 2001. "Managing Virtual Teams." *Workforce (10928332),* **80**(6), p. 60. 1, 36

[7] Jarvenpaa, S. L., and Leidner, D. E., 1999. "Communication and Virtual Trust in Global Virtual Teams." *Organization Science,* **10**(6), pp. 791–815. 1, 36

[8] Rocco, E., 1998. "Trust breaks down in electronic contexts but can be repaired by some initial face-to-face contact." *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '98*(496), pp. 496–502. 1, 36

[9] Warkentin, M. E., Sayeed, L., and Hightower, R., 1997. "Virtual Teams versus Face-to-Face Teams: An Exploratory Study of a Web-based Conference System." *Decision Sciences,* **28**(4), pp. 975–996. 1, 36

[10] Driskell, J., Radtke, P., and Salas, E., 2003. "Virtual teams: effects of technological mediation on team performance.." *Group Dynamics: Theory, Research, and Practice,* **7**(4), pp. 297–323. 1, 4, 7, 16, 19, 21, 36

[11] Thompson, L. F., and Coovert, M. D., 2003. "Teamwork online: The effects of computer conferencing on perceived confusion, satisfaction and postdiscussion accuracy.." *Group Dynamics: Theory, Research, and Practice,* **7**(2), pp. 135–151. 2, 36

[12] Zender, F., Wright, M., Richey, M., Gorrell, S., McPherson, K., and Madni, A. M., 2014. "Aerospace Partners for the Advancement of Collaborative Engineering." *121st ASEE Annual Conference & Exposition*, pp. 1–9. 2, 15

[13] Koster, J., Velazco, A., and Munz, C., 2012. "HYPERION UAV: An International Collaboration." *50th AIAA Aerospace Sciences Meeting including the New Horizons Forum and Aerospace Exposition*(January), pp. 1–23. 2, 5, 15, 17

[14] Dym, C. L., Agogino, A. M., Eris, O., Frey, D. D., and Leifer, L. J., 2005. "Engineering Design Thinking, Teaching, and Learning." *Journal of Engineering Education*(January), pp. 103–120. 2, 5, 8, 15, 17, 20

[15] Maruping, L. M., and Agarwal, R., 2004. "Managing team interpersonal processes through technology: a task-technology fit perspective.." *The Journal of applied psychology,* **89**(6), pp. 975–990. 2, 7, 15, 19, 21, 22, 26, 30, 32

[16] Levi, D., 2014. *Group Dynamics for Teams*., 4th editio ed. Sage Publications, Inc. 2, 6, 7, 8, 11, 17, 19, 20, 26, 27, 39

[17] Gueldenzoph, L. E., and May, G. L., 2002. "Collaborative Peer Evaluation: Best Practices for Group member Assessments." *Business Communication Quarterly,* **65**(1), pp. 9–20. 2, 10, 35, 38

[18] Wiggins, N., Blackburn, M., and Hackman, J. R., 1969. "Prediction of First-Year Graduate Success in Psychology." pp. 81–85. 3, 35, 56

[19] Keaten, J. A., and Richardson, M. E., 1993. "A field investigation of peer assessment as part of the student group grading process." *Annual Meeting of the Western Speech Communication Association Convention*. 3, 35

[20] May, G. L., and Gueldenzoph, L. E., 2006. "The effect of social style on peer evaluation ratings in project teams." *Journal of Business Communication,* **43**(1), pp. 4–20. 3, 35

[21] Carvalho, A., 2013. "Students' perceptions of fairness in peer assessment: evidence from a problem-based learning course." *Teaching in Higher Education,* **18**(5), pp. 491–505. 3, 35

[22] Ohland, M. W., and Finelli, C., 2001. "Peer evaluation in a mandatory cooperative education environment." *Proceedings of the 2001 American Society for Engineering Education Annual Conference & Exposition*. 3, 35

[23] Northrup, S. G., and Northrup, D. A., 2006. "Multidisciplinary teamwork assessment: Individual contributions and interdisciplinary interaction." *Proceedings - Frontiers in Education Conference, FIE*, pp. 15–20. 3, 35

[24] Hirsch, P. L., Shwom, B. L., Yarnoff, C., Anderson, J. C., Kelso, D. M., Olson, G. B., and Colgate, J. E., 2001. "Engineering design and communication: the case for interdisciplinary collaboration." *International Journal of Engineering Education,* **17**(4/5), pp. 343–348. 4, 16

[25] Bucciarelli, L. L., 1988. "An ethnographic perspective on engineering design." *Design Studies,* **9**(3), pp. 159–168. 4, 16

[26] Golden, T. D., and Raghuram, S., 2010. "Teleworker knowledge sharing and the role of altered relational and technological interactions." *Journal of Organizational Behavior,* **31**(8), pp. 1061–1085. 4, 16, 28

[27] Hertel, G., Geister, S., and Konradt, U., 2005. "Managing virtual teams: A review of current empirical research." *Human Resource Management Review,* **15**(1), pp. 69–95. 4, 16, 26

[28] Moore, B., Rhodes, M., and Stanley, R., 2011. Telework 2011. 4, 16

[29] Chinowsky, P. S., and Rojas, E. M., 2003. "Virtual Teams: Guide to Successful Implementation." *Journal of Management in Engineering,* **19**(3), pp. 98–106. 5, 16

[30] Rugarcia, A., Felder, R. M., Woods, D. R., and Stice, J. E., 2000. "the Future of Engineering Education I. a Vision for a New Century." *Chemical Engineering Education,* **34**(1), pp. 16–25. 5, 17

[31] Doerry, E., Klempous, R., Nikodem, J., and Paetzold, W., 2004. "Virtual student exchange: Lessons learned in virtual international teaming in interdisciplinary design education." *Ithet 2004: Proceedings of the Fifth International Conference on Information Technology Based Higher Education and Training*, pp. 650–655\r676. 5, 6, 17, 28

[32] Kirkpatrick, A. T., Danielson, S., Warrington, R. O., Smith, R. N., Thole, K. A., Wepfer, W. J., and Perry, T., 2011. "Vision 2030 Creating the Future of Mechanical Engineering Education." In *Proceedings of the ASEE International Conference and Exposition*, no. December, pp. AC 2011–2804. 5, 6, 17, 18

[33] Daft, R. L., and Lengel, R. H., 1986. "Organizational Information Requirements , Media Richness and Structural Design." *Management Science,* **32**(5), pp. 554–571. 7, 19, 21

[34] Levi, D., and Rinzel, L., 1998. "Employee Attitudes Toward Various Communications Technologies When Used For Communicating About Organizational Change." In *Human Factors in organizational design and management*, P. Vink, E. Koningsveld, and S. Dhondt, eds., 6 ed. Elsevier, Amsterdam, pp. 483–488. 7, 19, 21

[35] Maznevski, M. L., and Chudoba, K. M., 2000. "Bridging Space Over Time: Global Virtual Team Dynamics and Effectiveness." *Organization Science,* **11**(5), pp. 473–492. 7, 19

[36] French, D. J., Red, W. E., Walker, S. S., Madsen, P., Jensen, C. G., Robinson, D., Ruchty, J., Ward, S., Sanchez, C., and Phelps, G., 2012. e - Handbook Summary Report. 7, 19, 21

[37] Dym, C. L., 1999. "Learning Engineering: Design, Languages, and Experiences." *1Journal of Engineering Education*(April), pp. 145–148. 8, 20

[38] Kane, J. S., and Lawler, E. E., 1978. "Methods of peer assessment.." *Psychological Bulletin,* **85**(3), pp. 555–586. 9, 38

[39] Beebe, S. A., Barge, J. K., and McCormick, C., 1995. "The Competent Group Communicator : Assessing Essential Competencies of Small Group Problem Solving." In *The Annual Meeting of the Speech Communication Association*, pp. 2–39. 9, 10, 38

[40] Johnson, K. F., 1993. "Team peer evaluations: A student-generated quantitative measurement of group membership performance." *Annual Meeting of the Association for Education in Journalism and Mass Communication*. 9, 38

[41] Chalupa, M. R., Chen, C. S., and Sormunen-Jones, C., 2000. "Reliability and Validity of the Group Member Rating Form." *Delta Pi Epsilon Journal,* **42**(4), pp. 235–245. 9, 10, 37, 38, 42

[42] Oakley, B., Felder, R. M., and Brent, R., 2004. "Turning Student Groups into Effective Teams." *Journal of Student Centered Learning,* **2**(1), pp. 9–34. 9, 38

[43] Finelli, C. J., Bergom, I., and Mesa, V., 2011. "Student teams in the engineering classroom and beyond: setting up students for success." *CRLT Occasional Papers*(29). 9, 38

[44] Paswan, A. K., and Gollakota, K., 2004. "Dimensions of Peer Evaluation, Overall Satisfaction, and Overall Evaluation: An Investigation in a Group Task Environment." *Journal of Education for Business,* **79**(4), pp. 225–231. 10, 37, 38, 40, 41, 43

[45] Dyer, W. G., Dyer, J. H., and Dyer, W. G., 2013. *Team Building: Proven Strategies for Improving Team Performance*., Vol. 5. 10, 38

[46] Hackman, J. R., 1990. *Groups That Work (And Those That Don't): Creating conditions for effective teamwork*. Jossey-Bass. 10, 11, 39, 40, 60

[47] Horwitz, S. K., 2005. "The Compositional Impact of Team Diversity on Performance: Theoretical Considerations." *Human Resource Development Review,* **4**(2), pp. 219–245. 10, 39

[48] Levi, D., and Slem, C., 1995. "Team Work in Research and Development Organizations - the Characteristics of Successful Teams." *International Journal of Industrial Ergonomics,* **16**(1), pp. 29–42. 10, 39

[49] Peeters, M. A. G., Van Tuijl, H. F. J. M., Rutte, C. G., and Reymen, I. M. M. J., 2006. "Personality and Team Performance: A Meta-Analysis." *European Journal of Personality,* **20**(December 2004), pp. 377–396. 11, 39

[50] Dingel, M., and Wei, W., 2014. "Influences on peer evaluation in a group project: an exploration of leadership, demographics and course performance." *Assessment & Evaluation in Higher Education,* **39**(6), pp. 729–742. 11, 39, 40

[51] Whitman, L. E., Malzahn, D. E., Chaparro, B. S., Russel, M., Langrall, R., and Mohler, B. A., 2005. "A Comparison of Group Processes , Performance , and Satisfaction in Face-to-Face Versus Computer - Mediated Engineering Student Design Teams." *Journal of Engineering Education,* **94**(July), pp. 327–333. 11, 39, 59

[52] Halfhill, T., Sundstrom, E., Lahner, J., Calderone, W., and Nielsen, T. M., 2005. "Group Personality Composition and Group Effectiveness: An Integrative Review of Empirical Research." *Small Group Research,* **36**(1), pp. 83–105. 11, 39

[53] Canelon, J., Ryan, T., Iriberri, A., and Eryilmaz, E., 2015. “Conflicts on Team Satisfaction and Face Loss and the Moderating Role of Face Work Behaviors in Online Discussions.” *Academy of Educational Leadership Journal,* **19**(3), pp. 45–62. 11, 39

[54] Hackman, J., 1987. “The design of work teams.” *Handbook of Organizational Behavior*, pp. 315–342. 11, 29, 39

[55] Hinds, P. J., and Mortensen, M., 2005. “Understanding Conflict in Geographically Distributed Teams: The Moderating Effects of Shared Identity, Shared Context, and Spontaneous Communication.” *Organization Science,* **16**(3), pp. 290–307. 11, 39

[56] Curseu, P. L., and Schruijer, S. G. L., 2010. “Does conflict shatter trust or does trust obliterate conflict? Revisiting the relationships between team diversity, conflict, and trust..” *Group Dynamics: Theory, Research, and Practice,* **14**(1), pp. 66–79. 11, 39

[57] Germain, M.-L., and McGuire, D., 2014. “The Role of Swift Trust in Virtual Teams and Implications for Human Resource Development.” *Advances in Developing Human Resources,* **16**(3), pp. 356–370. 11, 39

[58] Wang, J., and Wang, X., 2012. *No Title No Title*., 1 ed. Higher Education Press. 11, 12, 45

[59] Muthén, L. K., and Muthén, B., 2010. Mplus Short Courses Topic 4 Growth Modeling With Latent Variables Using Mplus : Advanced Growth Models, Survival Analysis, And Missing Data Tech. rep. 11, 12, 44

[60] Worthington, R. L., and Whittaker, T. a., 2006. “Scale development research: A content analysis and recommendations for best practices.” *The Counseling Psychologist,* **34**(6), pp. 806–838. 12, 45, 50

[61] Brown, T. A., 2015. *Confirmatory Factor Analysis for Applied Research*. The Guilford Press. 12, 44

[62] May, A., and Carter, C., 2001. “A case study of virtual team working in the European automotive industry.” *International Journal of Industrial Ergonomics,* **27**(3), mar, pp. 171–186. 15

[63] Peterson, K., 2011. “A Wing and a Prayer: Outsourcing at Boeing.” *Reuters*(January), pp. 1–12. 15, 16, 36

[64] Zolin, R., Hinds, P. J., Fruchter, R., and Levitt, R. E., 2004. “Interpersonal trust in cross-functional, geographically distributed work: A longitudinal study.” *Information and Organization,* **14**(1), pp. 1–26. 15

[65] Becar, J. S., and Gorrell, S. E., 2015. “A Collaborative Conceptual Aircraft Design Environment for the Design of Small-Scale UAVs in a Multi-University Setting.” In *53rd AIAA AerosPACE Sciences Meeting*, no. January, AIAA SciTech, pp. 1–16. 15

[66] Cannon, L., Cunningham, C. C., Inouye, A. L., Stone, B., and Zender, F., 2014. “Looking Back: A Student Review and History of AerosPACE - a Multi-University, Multi-Disciplinary, Distributed, Industry-University Capstone Project.” *122nd ASEE Annual Conference & Exposition*. 15

[67] Richey, M., Zender, F., and Schrage, D., 2012. "An innovative approach to an integrated design and manufacturing multi-site cloud-based capstone project." *ASEE Annual Conference* 15, 37

[68] Lang, J. D., Cruse, S., McVey, F. D., and McMasters, J., 1999. "Industry Expectations of New Engineers: A Survey to Assist Curriculum Designers." *Journal of Engineering Education,* **88**(1), pp. 43–51. 17

[69] Dennis, A. R., Fuller, R. M., and Valacich, J. S., 2008. "Media, Tasks, and Communication Processes: A Theory of Media Synchronicity." pp. 575–600. 22

[70] Watson, W. E., BarNir, A., and Pavur, R., 2010. "Elements Influencing Peer Evaluation: An Examination of Individual Characteristics, Academic Performance, and Collaborative Processes." *Journal of Applied Social Psychology,* **40**(12), pp. 2995–3019. 24, 40, 42

[71] Dieter, G., and Schmidt, L., 2009. *Engineering Design*. McGraw-Hill Higher Education, New York City. 24, 28

[72] Administration, N. A., and Space, 1994. System Engineering "Toolbox" for Design-Oriented Engineers. 24

[73] Siebdrat, F., Hoegl, M., and Ernst, H., 2009. "How to manage virtual teams." *MIT Sloan Management Review,* **50**(4), pp. 63–71. 26

[74] Moreland, R. L., Argote, L., and Krishnan, R., 1996. *Socially Shared Cognition at Work, What's Social About Social Cognition?*. Thousand Oaks, California. 26

[75] Lovelace, K., Shapiro, D. L., and Weingart, L. R., 2001. "Maximizing cross-functional new product teams' innovativeness and constraint adherence: A conflict communications perspective." *Academy of Management Journal,* **44**(4), pp. 779–793. 26

[76] Hackman, J. R., 2009. Why teams don't work. 26

[77] Griffith, T. L., Mannix, E. A., and Neale, M. A., 2003. "Conflict and Virtual Teams." *Virtual teams that work: Creating conditions for virtual team effectiveness*, pp. 335–352. 26

[78] Malhotra, A., Majchrzak, A., and Rosen, B., 2007. "Academy of Management." *The Academy of Management,* **21**(1), pp. 60–70. 26

[79] Dixon, K. R., and Panteli, N., 2010. "From virtual teams to virtuality in teams." *Human Relations,* **63**(8), pp. 1177–1197. 26

[80] Dennis, A. R., and Valacich, J. S., 1999. "Rethinking Media Richness: Towards a Theory of Media Synchronicity." *HICSS-32: Proceedings of the 32nd Annual Hawaii International Conference on Systems Sciences - 1999,* **Track1**(c), p. 10. 26

[81] Yang, M. C., 2009. "Observations on concept generation and sketching in engineering design." *Research in Engineering Design,* **20**(1), dec, pp. 1–11. 27

[82] Chandrasegaran, S., Badam, S. K., Zhao, Z., Elmqvist, N., Kisselburgh, L., and Ramani, K., 2014. Collaborative Sketching with skWIKI: A Case Study. 27

[83] Kopp, G., 1999. "Engineering graphics in the new millennium: integrating the strengths of sketching and CAD." *Frontiers in Education Conference, 1999.*, pp. 12–13. 27

[84] Bracewell, R. H., Ahmed, S., and Wallace, K. M., 2004. "Dred and Design Folders, a Way of Capturing, Storing and Passing on, knowledge generated during design projects." *DETC'04, ASME 2004 Design Engineering Technical Conferences*(1970), pp. 1–12. 28

[85] Klein, M., 1993. "Capturing design rationale in concurrent engineering teams." *IEEE Computer Journal,* **26**(Special Issue on Computer Support for Concurrent Engineering), pp. 1–17. 28

[86] Hepworth, A., Halterman, K., Stone, B., Yarn, J., and Jensen, C. G., 2015. "An Integrated Task Management System to Reduce Semantic Conflicts in Multi-user CAD." *Concurrent Engineering: Research and Applications*(March), pp. 1–12. 29

[87] Dennis, A. R., and Valacich, J. S., 1993. Computer Brainstorms: More Heads Are Better Than One. 31

[88] Gallupe, R. B., Bastianutti, L. M., and Cooper, W. H., 1991. "Unblocking Brainstorms." *Journal of Applied Psychology,* **76**(1), pp. 137–142. 31

[89] Jehn, K. A., and Mannix, E. A., 2001. "The Dynamic Nature of Conflict : A Longitudinal Study of Intragroup Conflict and Group Performance." *Academy of Management Journal,* **44**(2), pp. 238–251. 33

[90] Hinds, P. J., and Bailey, D. E., 2003. "Out of Sight, Out of Sync: Understanding Conflict in Distributed Teams." *Organization Science,* **14**(6), pp. 615–632. 33

[91] Grinter, R. E., Herbsleb, J. D., and Perry, D. E., 1999. "The Geography of Coordination : Dealing with Distance in R&D Work." *Proceedings of the ACM Conference on Supporting Group Work*, pp. 306–315. 33

[92] Wilson, J. M., Straus, S. G., and McEvily, B., 2006. "All in due time: The development of trust in computer-mediated and face-to-face teams." *Organizational Behavior and Human Decision Processes,* **99**(1), pp. 16–33. 36

[93] Berry, G. R., 2011. "Enhancing Effectiveness on Virtual Teams: Understanding Why Traditional Team Skills Are Insufficient." *Journal of Business Communication,* **48**(2), pp. 186–206. 36

[94] Channon, S. B., Davis, R. C., Goode, N. T., and May, S. A., 2016. "What makes a 'good group'? Exploring the characteristics and performance of undergraduate student groups." *Advances in Health Sciences Education*, pp. 1–25. 36, 37

[95] Tasa, K., Taggar, S., and Seijts, G. H., 2007. "The development of collective efficacy in teams: a multilevel and longitudinal perspective.." *The Journal of applied psychology,* **92**(1), pp. 17–27. 36

[96] Baker, D. F., 2007. "Peer Assessment in Small Groups: A Comparison of Methods." *Journal of Management Education,* **32**(2), pp. 183–209. 37

[97] Haas, A. L., Haas, R. W., and Wotruba, T. R., 1998. "The Use of Self-Ratings and Peer Ratings to Evaluate Performances of Student Group Members." *Journal of Marketing Education*. 37

[98] Gorrell, S., Jensen, C. G., Stone, B., Red, E., Richey, M., Zender, F., Wright, M., French, D., Hayashibara, S., Johnson, C., and John, S., 2014. "Aerospace Partners for the Advancement of Collaborative Engineering." In *121st ASEE Annual Conference & Exposition*. 37

[99] Roh, Y. S., Lee, S. J., and Mennenga, H., 2014. "Factors influencing learner satisfaction with team-based learning among nursing students." *Nursing and Health Sciences,* **16**(4), pp. 490–497. 41

[100] Ohland, M. W., Loughry, M. L., Woehr, D. J., Bullard, L. G., Felder, R. M., Finelli, C. J., Layton, R. A., Pomeranz, H. R., and Schmucker, D. G., 2012. "The comprehensive assessment of team member effectiveness: Development of a behaviorally anchored rating scale for self- and peer evaluation." *Academy of Management Learning and Education,* **11**(4), pp. 609–630. 41

[101] Ohland, M. W., Layton, R. A., Loughry, M. L., and Yuhasz, A. G., 2005. "Effects of Behavioral Anchors on Peer Evaluation Reliability." *Journal of Engineering Education,* **94**(3), pp. 319–326. 42

[102] Aggarwal, P., and O'Brien, C. L., 2008. "Social Loafing on Group Projects: Structural Antecedents and Effect on Student Satisfaction." *Journal of Marketing Education,* **30**(3), pp. 255–264. 43

[103] Maccallum, R. C., Roznowski, M., and Necowitz, L. B., 1992. "Quantitative Methods In Pyschology Model Modifications in Covariance Structure Analysis: The Problem of Capitalization on Chance." *Psychological Bulletin,* **111**(3), pp. 490–504. 50

[104] Swaab, R. I., Schaerer, M., Anicich, E. M., Ronay, R., and Galinsky, A. D., 2014. "The Too-Much-Talent Effect: Team Interdependence Determines When More Talent Is Too Much or Not Enough." *Psychological Science,* **25**(8), pp. 1581–1591. 56

[105] Werner, J. M., and Lester, S. W., 2001. "Applying a Team Effectiveness Framework to the Performance of Student Case Teams." *Human Resource Development Quarterly,* **12**(4), pp. 385–402. 57

[106] Pang, E., and Hung, H., 2012. "Learning Effectiveness and Satisfaction in Study Groups : A Consideration of the Moderating Factors." pp. 79–94. 57

[107] Brooks, C. M., and Ammons, J. L., 2003. "Free Riding in Group Projects and the Effects of Timing , Frequency , and Specificity of Criteria in Peer Assessments." *Journal of Education for Business*(78:5), pp. 268–272. 56

APPENDIX A. SURVEY PROMPTS FOR MOTIVATION, TECHNICAL SKILLS, SOCIAL SKILLS RATINGS

A.1

A brief explanation of the categories follows:

Motivation / Commitment: How motivated do you feel this person is to participate in the AerosPACE project? How well does (s)he keep commitments? Is (s)he passionate about his/her work? Does (s)he demonstrate a desire to learn more and improve her/his skills and understanding?

Technical Skill: How knowledgeable and/or experienced is this person? What level of working knowledge does (s)he demonstrate in the areas concerned? How effective at implementing this knowledge is the person? How well does this person learn and apply new topics or concepts? How creative is this person in applying his/her knowledge and experience?

Social Skill: Does this person know how to work well with others? How well does this person communicate her/his ideas, thoughts, and opinions? Does (s)he make other team members feel valued? Can (s)he disagree without being disagreeable? Does this person come to meetings as agreed upon and and participate? Does (s)he ask insightful questions or prompt thoughtful discussion?