**Michigan
Technological
University**

Michigan Technological University
Digital Commons @ Michigan Tech

Dissertations, Master's Theses and Master's Reports

2021

ENHANCING PEER FEEDBACK STRATEGIES TO IMPROVE TEAM EFFECTIVENESS

Amber Kemppainen
amber@mtu.edu

Copyright 2021 Amber Kemppainen

Recommended Citation

Kemppainen, Amber, "ENHANCING PEER FEEDBACK STRATEGIES TO IMPROVE TEAM EFFECTIVENESS", Open Access Dissertation, Michigan Technological University, 2021.
<https://doi.org/10.37099/mtu.dc.etdr/1255>

Follow this and additional works at: <https://digitalcommons.mtu.edu/etdr>

Part of the [Educational Assessment, Evaluation, and Research Commons](#)

ENHANCING PEER FEEDBACK STRATEGIES TO IMPROVE TEAM
EFFECTIVENESS

By

Amber J. Kemppainen

A DISSERTATION

Submitted in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

In Applied Cognitive Science and Human Factors

MICHIGAN TECHNOLOGICAL UNIVERSITY

2021

© 2021 Amber J. Kemppainen

This dissertation has been approved in partial fulfillment of the requirements for the Degree of DOCTOR OF PHILOSOPHY in Applied Cognitive Science and Human Factors.

Department of Cognitive and Learning Sciences

Dissertation Advisor: *Dr. Susan Amato-Henderson*

Committee Member: *Dr. Elizabeth Veinott*

Committee Member: *Dr. Kelly Steelman*

Committee Member: *Dr. Emily Dare*

Department Chair: *Dr. Kelly Steelman*

Table of Contents

List of figures	6
List of tables	9
Acknowledgements	11
Abstract	12
1 Introduction	13
2 Literature Review	22
2.1 Collaborative Learning	22
2.2 Facilitating Collaborative Teams	25
2.2.1 Team Development Model	25
2.2.2 Team Creation Process	27
2.2.3 Challenges Affecting Teams	28
2.2.4 Evaluation of Teaming	32
2.2.5 Strategies for Mitigating Team Issues	36
2.3 Research Goals and Design	42
3 Study 1	44
3.1 Methods	44
3.1.1 Participants	44
3.1.2 Materials	44
3.1.3 Design and Procedure	45
3.2 Results	45
3.3 Discussion	46
4 Study 2	48
4.1 Methods	48
4.1.1 Participants	48
4.1.2 Surveys	48
4.1.2.1 CATME Peer Evaluation Survey	48
4.1.3 Design and Procedure	50
4.2 Results	51
4.3 Discussion	57
5 Study 3	59
5.1 Methods	62

5.1.1	Participants.....	62
5.1.2	Materials	63
5.1.2.1	Surveys.....	64
5.1.2.1.1	CATME Survey	64
5.1.2.1.2	Team Interdependence	64
5.1.2.1.3	Team Satisfaction.....	65
5.1.2.1.4	Trust and Conflict	65
5.1.2.2	Feedback Instruction Manipulations.....	66
5.1.2.2.1	Reflective Manipulation.....	67
5.1.2.2.2	Procedural Manipulation.....	69
5.1.2.3	Instructional Materials	71
5.1.2.3.1	Videos	71
5.1.2.3.2	Handouts	74
5.1.2.4	Data Coding	77
5.1.3	Design and Procedure	79
5.1.3.1	Data Collection Timeline	80
5.2	Results	81
5.2.1	CATME Quantitative Measures	81
5.2.2	Dysfunction Measures	90
5.2.3	Peer Review Comments.....	93
5.3	Discussion	95
6	Study 4	99
6.1	Methods	100
6.1.1	Participants.....	100
6.1.2	Materials	101
6.1.2.1	Survey Instrument Manipulation	101
6.1.2.2	Peer Evaluation Introduction Video.....	104
6.1.2.3	Teaming Project	105
6.1.2.4	Task Manipulation	108
6.1.2.4.1	Study Skill Reflection	109
6.1.2.4.2	CTA Teaming Reflection.....	110
6.1.2.5	Qualitative Coding	112
6.1.2.5.1	Primary Coding.....	112
6.1.2.5.2	Secondary Coding.....	115
6.1.2.5.3	Reliability.....	117
6.1.3	Design and Procedure	120
6.1.3.1	Hypotheses.....	122
6.1.3.2	Procedure	123
6.2	Results	126
6.2.1	CATME Behavioral Code Results.....	126
6.2.1.1	Primary Code Results	128
6.2.1.2	Secondary Code Results	135

6.2.2	Relationships.....	142
6.2.2.1	Self-Evaluations	143
6.2.2.2	Peer-Evaluations	150
6.3	Discussion	157
7	Project Summary.....	165
7.1	Overall Results	165
7.2	Future Work	168
8	Reference List	171
A	Identifying Teaming Issues.....	176
B	CATME Peer Evaluation Instrument.....	177
C	Team Evaluation Follow-up Questions	180
D	Assessing Team Trust and Conflict	181
E	Assessing Teamwork with Adapted Surveys.....	183
F	Peer Evaluation Instructional Video Transcript.....	190
G	Instructional Activities – Fall 2020.....	192
H	Copyright Documentation.....	195

List of figures

Figure 1.1. First-Year Engineering Pathway at Michigan Technological University	16
Figure 1.2. First-Year Engineering Studio Session Classroom Layout	17
Figure 1.3. Design-Based Research Approaches in Educational Research (Adapted from Van den Akker, 2006).....	19
Figure 2.1. Tuckman's (2010) Teaming Model (Wilson, 2010).....	26
Figure 2.2. Five Dysfunctions of a Team Model	30
Figure 2.3. SARA model for Receiving Feedback (decisionwise, 2020).....	39
Figure 4.1. Example CATME Team Report.....	52
Figure 4.2. Example CATME Team Report (numeric results only).....	53
Figure 5.1. Screen Capture of CATME Peer-to-Peer Comments Window	61
Figure 5.2. Feedback Instruction Planning	67
Figure 5.3. Giving Effective Feedback: Reflecting on your Team's Experience Handout	75
Figure 5.4. Procedural Method: Steps for Giving Effective Feedback Handout	76
Figure 5.5. Spring 2020 Study Timeline.....	81
Figure 5.6. CATME Peer Evaluation Score Results Spring 2020 ($n = 452$, $m=0.9996$, $SD = 0.06427$)	86
Figure 5.7. CATME Team Interdependence Score Week 5 Results Spring 2020 ($n = 302$, $M = 16.38$, $SD = 2.96$)	87
Figure 5.8. Change in Team Interdependence Value in Spring 2020 ($n = 284$, $m=0.222$, $SD = 2.36$).....	88
Figure 5.9. CATME Team Satisfaction Score Week 5 Results Spring 2020 ($n = 302$, $M = 13.0$, $SD = 2.09$).....	89
Figure 5.10. Change in Team Satisfaction Values in Spring 2020 ($n = 284$, $M = -.173$, $SD = 1.58$).....	89
Figure 5.11. Trust Scale Score Week 5 Results Spring 2020 ($n = 462$, $M = 33.02$, $SD = 4.65$)	90

Figure 5.12. Change in Trust Value in Spring 2020 (n = 413, M = 0.70, SD = 3.90).....	91
Figure 5.13. Conflict Scale Score Week 5 Results Spring 2020 (n = 462, M = 32.27, SD = 4.74).....	92
Figure 5.14. Change in Conflict Values in Spring 2020 (n = 413, M = 0.67, SD = 4.27).....	92
Figure 5.15. Length of Feedback Comments in ENG1102 Spring 2020 Sample.....	94
Figure 5.16. CATME Behavioral Categories in ENG1102 Spring 2020 Sample	95
Figure 6.1. Original CATME Peer Comment Prompt	103
Figure 6.2. Enhanced CATME Peer Comment Prompt.....	104
Figure 6.3. Example Heat Treatment Analysis Plot	108
Figure 6.4. Fall 2020 Study Overview	122
Figure 6.5. Fall 2020 Study Timeline	124
Figure 6.6. Zoom Breakout Room Activities.....	125
Figure 6.7. Distribution of Behavioral Category Comments by Study Group (Midterm Evaluations)	129
Figure 6.8. Contributing to Team's work by Study Group (Midterm Evaluations)	130
Figure 6.9. Expecting Quality by Study Group (Midterm Evaluations).....	131
Figure 6.10. Average Behaviors by Study Group (Midterm Evaluations)	131
Figure 6.11. Distribution of Behavioral Category Comments over Time	134
Figure 6.12. Feedback Examples by Study Group (Midterm Evaluations).....	135
Figure 6.13. Actionable Feedback by Study Group (Midterm Evaluations)	136
Figure 6.14. Impact Feedback by Study Group (Midterm Evaluations).....	138
Figure 6.15. Comments containing at least two modifiers by Study Group (Midterm Evaluations)	139
Figure 6.16. Comments containing all three modifiers by Study Group (Midterm Evaluations)	140
Figure 6.17. Distribution of Secondary Behavioral Comments over Time	142

Figure B.1. CATME Rating: Contributing to Team’s Work (Ohland et al., 2012).....	177
Figure B.2. CATME Rating: Interacting with Teammates (Ohland et al., 2012).....	177
Figure B.3. CATME Rating: Keeping the Team on Track (Ohland et al., 2012).....	178
Figure B.4. CATME Rating: Expecting Quality (Ohland et al., 2012)	178
Figure B.5. CATME Rating: Having Related Knowledge, Skills, and Abilities (Ohland et al., 2012)	179
Figure E.1. Instructions and CATME Behavior Scale Questions for Fall 2020 Peer Evaluation Surveys	185
Figure E.2. Instructions and Peer Evaluation Prompts for Control Group	186
Figure E.3. Instructions and Peer Evaluation Prompts for Enhanced and CTA Groups.	188
Figure E.4. End of Peer Evaluation Survey and Consent Information.....	189
Figure G.1. Study Skills Reflection Assignment	193
Figure G.2. CTA Teaming Assignment	194
Figure H.1. Permission Documentation for Model of Teaming (Wilson, 2010)	196
Figure H.2. Permission Documentation for Points of Leverage for Creating Conditions that Enhance Group Task Performance (Tuckman, 1990)	197
Figure H.3. Permission Documentation for SARA model for Receiving Feedback (decisionwise, 2020).....	198

List of tables

Table 2.1. Characteristics of Teams who Experience one of the Five Dysfunctions of a Team (Lencioni, 2007).....	31
Table 2.2. Points of Leverage for Creating Conditions That Enhance Group Task Performance (Hackman, 1990, pg. 13)	37
Table 2.3. The Differences between Attributes and Behaviors in Team Member Feedback	38
Table 3.1. Top 10 Team Issues Identified by First-Year Engineering Faculty and LEAP Leaders in Fall 2019.....	46
Table 5.1. Initial Codes from Spring 2020 sample	77
Table 5.2. CATME Themes and Initial Codes	78
Table 5.3. Pre-Intervention Results for Dependent Variables by Time of Day (Week 5).....	82
Table 5.4. Post Intervention results between Reflective and Procedural Manipulations (Week 8).....	83
Table 5.5. Changes in pre-post results between Reflective and Procedural manipulations (Week 8 – Week 5)	84
Table 5.6. Correlations and p-values between Teaming Measures (week 5) ⁺	85
Table 6.1. Sample Distribution	101
Table 6.2. Codebook for Peer Evaluation Analysis.....	114
Table 6.3. Secondary Code Definitions	116
Table 6.4. Inter-rater Reliability Measures on Midterm Peer Evaluation Coding.....	118
Table 6.5. Classification of Cohen’s kappa (Laerd, 2015)	119
Table 6.6. Inter-rater Reliability Measures – Independent Coding of Final Peer Evaluations (n = 17 students, 107 comments)	120
Table 6.7. Effects of Study Manipulations on Behavioral Categories (Midterm Evaluations)	133
Table 6.8. Changes in Primary Codes over Time	134

Table 6.9. Effects of Study Manipulations on Secondary Codes (FE, FI, FA) (Midterm Evaluations)	141
Table 6.10. Changes in Secondary Codes over Time	142
Table 6.11. Correlations between Qualitative Self-Evaluation Comments on Midterm Peer Evaluations (Control: n=45) ⁺	146
Table 6.12. Correlations between Qualitative Self-Evaluation Comments on Midterm Peer Evaluations (EP: n = 39) ⁺	147
Table 6.13. Correlations between Qualitative Self-Evaluation Comments on Midterm Peer Evaluations (CTA: n=47) ⁺	148
Table 6.14. Correlations between Qualitative Self-Evaluation Comments on Midterm Peer Evaluations (All groups: n=131) ⁺	149
Table 6.15. Correlations between Qualitative Peer-Evaluation Comments on Midterm Peer Evaluations (Control: n=44) ⁺	153
Table 6.16. Correlations between Qualitative Peer-Evaluation Comments on Midterm Peer Evaluations (EP: n=38) ⁺	154
Table 6.17. Correlations between Qualitative Peer-Evaluation Comments on Midterm Peer Evaluations (CTA: n=47) ⁺	155
Table 6.18. Correlations between Qualitative Peer-Evaluation Comments on Midterm Peer Evaluations (All groups: n=131) ⁺	156
Table A.1 Faculty/LEAP Leader Issue Identification Survey.....	176
Table C.1. Team Satisfaction follow-up Questions	180
Table C.2 Team Interdependence follow-up Questions.....	180
Table D.1. Team Reflection Survey.....	181
Table D.2. Teaming Reflection Scoring Interpretation.....	182

Acknowledgements

I want to thank a number of individuals and groups that made this possible. First, I would like to thank my Lord and Savior, Jesus Christ for “I can do all things through Christ who strengthens me.” (Philippians 4:13). I would like to thank my wonderful husband and my daughters for their support and understanding why mommy suddenly became addicted to coffee. Thanks to my friends and co-workers in the Engineering Fundamentals Department for their support and encouragement and help with data collection and the ENG1101 and ENG1102 students who participated in this research. I would like to thank my wonderful advisor, Dr. Susan Amato-Henderson, and my committee members, Dr. Emily Dare, Dr. Kelly Steelman, and Dr. Elizabeth Veinott, for their advice and for pushing me to make this work the best it can be. I would like to thank the faculty and staff of the Applied Cognitive Science and Human Factors Program for their help and sharing their knowledge. Lastly, I would like to thank the King-Chávez-Parks Future Faculty Fellowship Program who provided funding for me to pursue this degree and Michigan Technological University for their support through the employee education program.

Abstract

This dissertation is focused on improving the teaming experience of students by evaluating the effectiveness of current peer evaluation instructional techniques and evaluation methods in producing meaningful ratings of individual and team member performance and constructive feedback. There exists a need for revision to the current peer evaluation methods as many students are unable or unwilling to provide meaningful feedback or assessment of their own or teammate's contributions on assignments. A series of four studies were used to develop a Cognitive Task Analysis (CTA) instructional intervention, a codebook to evaluate the qualitative feedback provided in peer evaluation comments, and a modification to a popular peer evaluation tool, CATME. These interventions resulted in an increasing percentage of students who provided comments on their peer evaluations and significant increases in the utility of these comments (operationally defined as detailed, references specific team work, and providing actionable and/or impactful feedback).

1 Introduction

In 2004, I started as an instructor in the First-Year Engineering Program at Michigan Tech. My experience as a Chemical Engineer focused primarily on process optimization and I quickly learned that optimizing a process is different than optimizing a learning experience. I jumped into learning educational pedagogy and researching best practices. My classroom became a testbed for new practices and I found myself engaging in Design-Based Educational Research (DBR) long before I knew the terminology.

According to Anderson and Shattuck (2012), a quality DBR study has the following characteristics: a) is situated in a real educational context, b) focuses on the design and testing of a significant intervention, c) uses mixed methods for assessment, d) involves multiple iterations, e) involves a collaborative partnership between researchers and practitioners, and f) has design principles that evolve or lead to practical principles, patterns, or grounded theories that advance the domain and educational practice (Anderson and Shattuck, 2012). My initial work met several of these criteria, but lacked others. For example, learning to assess outcomes was something I learned along the way. I found myself wanting a stronger foundation in statistical methods and assessment for some of the challenges I wanted to address in the first-year engineering classroom and curriculum. In particular, the collaborative component of the first-year engineering program.

From the beginning of my career, the emphasis was placed on creating a balanced, supportive team environment. Much effort was put into the team making process and trying to balance various factors (e.g., student schedules, interests, hobbies,

culture) in order to make a positive, effective teaming experience. Peer evaluations were used to assess individual effort at the end of the semester. There were several issues that I noticed with these peer evaluation models. First, some teams do not show signs of dysfunction in the classroom when being observed by the instructor, but report issues in the peer evaluation. Second, there were teams where it was apparent not everyone was contributing equally, yet their peers rated everyone the same on peer evaluations. Additionally, some poor team experiences were not relayed on the peer evaluations, but were observed by the instructor. Throughout my experience, these issues seem to be consistent in first-year engineering teams, but the root cause of these issues and student willingness to identify these issues was unknown.

In 2015, I started pursuing my PhD in Applied Cognitive Science and Human Factors (ACSHF). My focus in pursuing this degree has been to develop my statistical and analytical skills to assess student learning, the use of instructional techniques, and implementation of educational technology to enhance student learning and success. I focused my dissertation work on improving the teaming experience of first-year engineering students by evaluating the effectiveness of current peer evaluation instructional techniques and evaluation methods in producing meaningful ratings of individual and team member performance and constructive feedback. This project has the potential to have a great impact on student learning in the first-year engineering program at Michigan Tech as well as the potential to impact learning and instruction on a national or international level.

Collaborative learning has shown to benefit student learning, engagement, and performance (Svinicki and McKeachie, 2014). The U. S. Department of Education considers Collaborative Problem Solving (CPS) as an essential 21st century skill, which requires both collaborative (i.e., working in a group) and problem-solving skills (U.S. Department of Education, 2017). While historically K-12 education has been primarily concerned with content delivery, collaborative problem solving has recently been integrated and emphasized in the Common Core State Standards (U.S. Department of Education, 2017). Unfortunately, despite the collaborative learning experiences now available in the majority of K-12 classrooms, building and maintaining effective teams remains a problem in post-secondary education and industry (Loughry, Ohland, & Woehr, 2014; Felder and Brent, 2016).

The importance of this issue in engineering education is outlined in ABET Student Outcome 5 as, at the end of their undergraduate program, students should have “an ability to function effectively on a team whose members together provide leadership, create a collaborative and inclusive environment, establish goals, plan tasks, and meet objectives” (www.abet.org). In order to work towards that objective, the first-year engineering program at Michigan Technological University (Michigan Tech) focuses on the use of active, collaborative learning to deliver the content of their courses (Sorby, Monte, & Hein, 2001). The first-year engineering program at Michigan Tech consists of two main pathways depending on math-readiness (see Figure 1.1 below). For those engineering students who are initially enrolled in Calculus I or higher, they are enrolled in a two-semester course sequence of ENG1101: Engineering Problem Solving and

Analysis and ENG1102: Engineering Modeling and Design. For those students who are initially enrolled in Pre-Calculus, they take a three-semester engineering sequence, where ENG1101 is spread out into ENG1001: Engineering Problem Solving in the fall semester and ENG1100: Engineering Analysis in the spring semester, followed by ENG1102 the following fall semester.

Figure 1.1. First-Year Engineering Pathway at Michigan Technological University

In 2017, the first-year program moved to a flipped model for course delivery. ENG1101 and ENG1102 were taught with 5 sections meeting together twice a week for

two hours in a Studio Session with the instructor and LEAP Leaders (near-peer student mentors). Once a week, each section met separately with their LEAP Leader for a 50-minute LEAP Session. An example classroom layout of the Studio Session is shown below in Figure 1.2. In this diagram, it shows a Studio Session layout with sections 26 – 30, which would have a maximum of 120 students total. Each section has a total of three tables, with two teams of four students assigned to each table.

Figure 1.2. First-Year Engineering Studio Session Classroom Layout

Within this flipped learning environment, some of the course material (reading, video, activities) occurred before each Studio Session. In class, student teams focused on problem-solving using computational tools in a collaborative learning environment. Each student had a laptop computer to engage with the course materials. Additionally, there were large monitors embedded in the tables for each team to use as a collaborative tool. One team member could attach their laptop to this monitor and all students could work on team assignments together.

Although the delivery methods and course content have changed since the program was established in 2000, active, collaborative learning remains a hallmark of the

program. As mentioned previously, students are grouped into teams of four to work on the class activities and projects. Success for Michigan Tech's first-year engineering program hinges on the approximately 240 student teams (each semester) learning to work collaboratively (working with others on a structured activity (EBS, 2020)) if not completely cooperatively (working in a group for a focused project (EBS, 2020)) in their classes.

To provide some context, in the first-year engineering program at Michigan Tech, the problem-solving class activities mentioned above would be considered collaborative learning. The students work may work with their classmates on these activities. Longer "mini-projects" (ENG1001/1100/1101) or semester-long design projects (ENG1102) would be considered cooperative learning. These projects require interdependence, where teammates are mutually dependent on each other to complete all of the project tasks.

For this project, I devised and completed a series of four studies over the past two academic years (2019-20 and 2020-21) which explore different independent variables, tool manipulations, or instructional manipulations to determine their effect on peer evaluation ratings and feedback comments. I used a design-based educational research (DBR) process as a framework to explore, implement, and evaluate these different manipulations. A general overview of the Design-Based Research (DBR) approach for educational research is shown below in Figure 1.3.

To begin, the research is focused on the analysis of practical problems by researchers and practitioners in collaboration (Figure 1.3, step 1). In this case, I am focused on the analysis of the teaming experience of first-year engineering students by

evaluating their peer evaluation ratings and feedback to each other. This has been a historically difficult area to assess and one which greatly impacts the student learning and experience in the first-year engineering classroom. Chapter 2 focuses on discussing the literature regarding peer collaboration and evaluation, and common team issues experienced by undergraduate students in general and first-year engineering students specifically. Additionally, I discuss the role that giving and receiving honest and constructive peer feedback has in team functionality. This literature review informs the development of solutions by existing design principles and educational innovations (Figure 1.3, step 2).

Figure 1.3. Design-Based Research Approaches in Educational Research (Adapted from Van den Akker, 2006)

Copyright 2006 from Educational Design Research by J. Van den Akker, K. Gravemeijer, S. McKenney, & N. Nieveen (Eds.). Reproduced by permission of Taylor and Francis Group, LLC, a division of Informa plc. This permission does not cover any third party copyrighted work which may appear in the material requested. User is responsible for obtaining permission for such material separately from this grant.

In Chapters 3 and 4, I review the results of two studies from Fall 2019 sections of ENG1101: Engineering Problem Solving and Analysis. These studies seek to narrow the

scope of problem to the first-year program at Michigan Technological University. In the first study (Chapter 3), I explored which teaming issues(s) were most observed in first-year engineering students by tapping into the collective knowledge and experience of first-year engineering faculty and near-peer mentors. A second study, discussed in Chapter 4, examined how the first-year engineering teams at Michigan Technological University typically function and utilize current peer evaluation tools. The information gained from this study was important for understanding students' evaluative behaviors prior to any instruction, training, or feedback manipulations. The results of these studies guided the design and hypotheses for a study in Spring 2020.

Chapters 5 and 6 demonstrate the iterative cycle of testing and refinement of my interventions in the classroom (Figure 1.3, step 3). Chapter 5 focuses on the study of Spring 2020 students in ENG1102: Engineering Modeling and Design. This study introduced two instructional manipulations (Reflective and Procedural) designed to enhance peer review feedback and measured the changes in the dependent teaming variables (CATME peer rating, team satisfaction, team interdependence, conflict, and trust) between and within these manipulations utilizing a pre/post methodology. The peer review feedback provided by the students allowed for the development of a codebook that identified recurring ideas and themes brought forward by the students in their self and peer evaluation comments. This developed a tool for others to use in the evaluation of peer evaluation comments (Figure 1.3, step 4).

Unfortunately, the changing academic landscape due to COVID-19 modified Study 3 in situ, which opened up an opportunity for a fourth study to refine my methods.

This final study was completed in Fall 2020 (Chapter 6). The study design included experimental manipulation of the type of feedback instruction, CATME survey instrument, and midterm/end-of-term timing of surveys on quantitative and qualitative peer feedback. The coding scheme developed in Study 3 was revised and refined. The overall conclusions across this series of studies are discussed in Chapter 7.

2 Literature Review

In this literature review, I focus on a few main areas. First, I define collaborative learning in general and discuss several theories of collaborative learning. This is followed by a discussion of the benefits and challenges related to collaborative learning. I then focus on the development and facilitation of collaborative teams in the educational environment. This is followed by a discussion of the specific challenges facing teams and strategies for mitigation that instructors may use in their classrooms. I finish with a discussion of the research design and goals for this dissertation.

2.1 Collaborative Learning

Collaborative and cooperative learning have been used to describe working with others in an educational environment. As a point of clarification, cooperative learning is simply a type of collaborative learning where “students work together in small groups on a structured activity” (EBS, 2020). Cooperative learning involves students working on team projects where there is positive interdependence, individual accountability, promotive interaction, development of teamwork skills, and self-assessment of team functioning (Felder and Brent, 2016). Collaborative learning, on the other hand, focuses on students working together to learn more broadly. Examples of collaborative learning could be working on a semester design project, engaging in group discussions, or working in pairs on a class assignment.

Collaborative learning is an active, constructive process that depends on rich contexts and includes diverse, social learners (Smith and MacGregor, 1992).

Collaborative learning has roots with two prominent learning theories: Social Learning Theory and Social Constructivism (Kuwabara, Einbinder, Sun, and Azizi, 2020). Social Learning Theory states that behavior must be learned before it can be performed and it is most often learned through experience or observing others (Bandura & Walters, 1977). This theory emphasizes the role of self-regulatory processes in cognition. For example, an individual selects, organizes, analyzes, and interprets the stimuli from the environment (e.g., others, surroundings) and uses this process to impact their behavior (Bandura & Walters, 1977).

Social Constructivist Learning Theory identifies peer interaction as an essential part of cognitive development (Vygotsky, 1978; Pritchard and Woolard, 2013). Vygotsky stated that the “two primary means of learning occur through social interaction and language” (Vygotsky, 1978). Additionally, Vygotsky emphasized that learning should be matched with a child’s development level as determined “through problem solving under adult guidance or in collaboration with more capable peers” (Vygotsky, 1978; Clabaugh, 2010; Wink & Putney, 2002). These collaborations are most effective in what Vygotsky called the “zone of proximal development” where collaborative instruction is used to move a student from what they know to what they don’t know (Vygotsky, 1978). While, Social Constructivist Learning Theory emphasizes the role of community in this construction of knowledge (Pritchard and Woolard, 2013), this does not counter Social Learning Theory. Bandura indicates that behavior is most learned by observing others (Bandura & Walters, 1977). In fact, without others such as peers, parents, or community members, learning would be much more difficult.

There are many benefits to collaborative learning. Broadly, collaborative learning teaches students how to work collaboratively with others, a skill that is necessary in most professions. Peer interactions to facilitate learning can often be more productive than teacher-student interactions (Svinicki and McKeachie, 2014; Boud, Cohen, and Sampson, 1999). In terms of motivation, collaborative learning allows for mutual support, stimulation, reduces absenteeism, and increases the likelihood of completing individual work (Svinicki and McKeachie, 2014). Additional benefits include increased cognitive skills (Svinicki and McKeachie, 2014), increased interpersonal skills (EDS, 2020; Svinicki and McKeachie, 2014), and increased communication skills and critical enquiry and reflection (Boud, Cohen, and Sampson, 1999).

Boud, Cohen, and Sampson (1999) found as part of their literature review that collaborative, peer learning may demonstrate additional benefits for women and underrepresented minorities as this learning method values cooperation over competition and shows a greater respect for diverse backgrounds and experiences. Increased opportunities for diversity and personal feedback and acknowledgement of individual differences were also found by the Educational Broadcasting Corporation in their 2010 workshop (EDS, 2020). Smith and MacGregor (1992) found that collaborative learning encourages the development of learning communities (Smith and MacGregor, 1992).

From an instructional standpoint, collaborative learning approaches such as peer learning allow for flexibility in class structure and timing (Boud, Cohen, and Sampson, 1999). Larger projects and project-based learning can be used to increase the scope of

learning opportunities for peer groups or teams. Additionally, collaborative learning lends itself to increased use of active learning strategies (EDS, 2020).

However, designing collaborative activities requires conscious design of social space for in and out-of-class interactions (Smith and MacGregor, 1992), which can be challenging with traditional lecture-based classrooms. Boud, Cohen, and Sampson (1999) indicated that teams comprised of vastly different ages, experiences, cultures may form management and pedagogical challenges for instructors as well. Additionally, if collaborative learning is employed, group assessment, peer feedback, and self-assessment should be built into the curriculum (Boud, Cohen, and Sampson, 1999).

2.2 Facilitating Collaborative Teams

Creating effective, collaborative teams is an essential part of the first-year engineering program at Michigan Tech. In this section, I focus on discussing a popular model of team development, common methods for team formation, common challenges facing student teams, and instructional strategies for mitigating teaming issues.

2.2.1 Team Development Model

A common teaming model was developed by Tuckman in 1965. This five-step model describes the common stages experienced by a team from initial membership to dissolution of the group and outlines the times when team issues are most prevalent. This model, shown in Figure 2.1 below, starts at the forming stage. At this stage, the group has been created and group members are learning about the purpose, roles, and members of the group (Wilson, 2010). The next stage is storming. This stage is characterized by

unproductive conflict; where individuals are testing boundaries and determining what is acceptable within the group (Wilson, 2010). Some groups never make it through this phase to reach a place where they can work together effectively and productively. The norming phase comes next and is typically where roles, relationships, ground rules, expectations, and the norms of the group are established (Wilson, 2010). This leads to performing, where the team is productive and productive conflict can be used effectively (Wilson, 2010). When the team task is complete, a sense of closure and accomplishment characterizes the adjourning phase (Wilson, 2010).

Figure 2.1. Tuckman's (2010) Teaming Model (Wilson, 2010)
See Appendix H for documentation of permission to reprint this material.

2.2.2 Team Creation Process

In Chapter 1, I alluded to the effort I placed into the team making process for first-year engineering students and the emphasis on balancing factors such as student schedules, interests, hobbies, and culture in order to make a positive, effective teaming experience. There are many suggestions on how to create effective student teams, but many times these suggestions can be contradictory. For example, there is an emphasis placed on ensuring common schedules, interests, hobbies, culture, and gender, but teams should also include diverse personalities, abilities, and learning styles (Felder & Brent, 2001).

Felder and Brent (2001) suggest that instructors should focus on considering those students who are at risk academically, especially in the early years of undergraduate work. For example, in lower-division classes, instructors should create groups where at-risk students are supported rather than isolated (e.g., in engineering groups try not to put a female student alone on a team) (Felder & Brent, 2001). Felder and Brent do recommend removing this restriction in the upper-division classes, however, as these students may be isolated in their professional work and should have some experience with this before moving into their careers (Felder & Brent, 2001).

While the importance placed on creating effective teams is found throughout the literature, the process for creating teams may differ broadly (Felder and Brent, 2001; Felder and Brent, 2016). Some instructors prefer to collect teaming criteria via paper or online forms and create teams manually (Felder and Brent, 2001; Felder and Brent, 2016). Other instructors may allow students to self-select into groups (Felder and Brent, 2001; Felder and Brent, 2016). Finally, some instructors prefer to use online systems to

collect and create the teams (Loughry, Ohland, and Woehr, 2013). In terms of an online system to gather student information, there is software available such as CATME TeamMaker which helps balance the factors you personally are most concerned with along with the importance of these factors (Loughry, Ohland, and Woehr, 2013).

For the past few years in ENG1101 and ENG1102, the course instructors made teams of 3-5 students using the CATME Team Maker Software. The teaming factors chosen by instructors included student gender, schedule, major, hobbies, athletic commitments, and prior programming experience. The TeamMaker software creates teams based on these factors using grouping strategies defined by the instructor (e.g., similar hobbies, different experience, etc.) and instructors have the opportunity to accept or adjust teams. Student teams are then notified by email through the CATME interface who their teammates are, what their contact information is, and what their combined schedule (with busy/free times highlighted) is for their semester team.

2.2.3 Challenges Affecting Teams

As groups are often put together for a single project, the time spent in the group affects the development of the group dynamics and problem-solving strategies (McGrath, 1990). In general, the more time or instances in which individuals work together on project, the more opportunities they have to overcome challenges. Some challenges facing teams and the development of team skills are team members not pulling their weight, team members who dominate discussions, and conflict which can come from unclear or differing expectations between teammates (Felder & Brent, 2001). Additional challenges were summarized by Loughry, Ohland, and Woehr (2014) as part of their literature review

such as division of work, poor communication, conflict, individual work preferences, and social loafing (Loughry, Ohland, and Woehr, 2014).

Van Tyne (2017) also point out that students fear that working on a team may negatively impact their overall individual grade, this fear impacts how individuals trust and interact with their teammates (Van Tyne, 2017). As mentioned previously, level two, conflict in a team, or “storming”, occurs in most teams (Tuckman, 1965; Wilson, 2010). Tekleab, Quigley, and Tesluk (2009) state that when teams overcome this conflict, trust develops between team members, leading to higher levels of team cohesion and greater team effectiveness (Tekleab, Quigley, & Tesluk, 2009, pg. 171). Additionally, Harley (2013) emphasizes that the building up and strengthening of working relationships requires building and maintaining trust.

Patrick Lencioni developed a model of five key dysfunctions of teams who are not demonstrating effectiveness (Lencioni, 2007). Lencioni’s model of Five Dysfunctions of a Team, shown in Figure 2.2 below, has been shown to demonstrate both construct and predictive validity with team effectiveness, in addition to having face validity (Hamlin, 2008). This model depicts a hierarchy in the issues experienced by teams in industry as well as those in academia. Starting at the bottom, if a team is unable to trust one another, they cannot engage in productive conflict, fail to commit to a plan of action, are not accountable to one another, and ultimately their work does not meet the desired outcome. Lencioni’s model (2007) builds on each step.

Figure 2.2. Five Dysfunctions of a Team Model

With respect to first-year engineering teams, the ENG course series at Michigan Tech provides students with their first experience with teams in the university environment, so it would make sense that their ENG class teams will struggle most with these first two dysfunctions. A team who is unable to establish trust will more likely have issues with the other levels as well. So, trust is the foundation on which all teams are built upon. For example, without this trust-based relationship, feedback (especially tough feedback) is difficult to both provide and accept (Harley, 2013). It is, therefore, important for instructional strategies to ensure that trust is established in teams before moving on to addressing other issues.

Table 2.1 summarizes the team issues faced by teams experiencing each dysfunction in this model according to Lencioni. Recommendations from Hamlin (2008)

on the application of this model states that self-assessment on the various dysfunctions can identify which dysfunction(s) should be worked on within the team to increase overall team effectiveness.

Table 2.1. Characteristics of Teams who Experience one of the Five Dysfunctions of a Team (Lencioni, 2007)

Dysfunction	Teams with this dysfunction...
Absence of Trust	<ul style="list-style-type: none"> ● conceal their weaknesses and mistakes from one another ● hold grudges ● fail to recognize and tap into one another's skills and experiences ● hesitate to offer help outside their own areas of responsibility ● jump to conclusions about the intentions and aptitudes of others without attempting to clarify them ● dread meetings and find reasons to avoid spending time together ● hesitate to ask for help or provide constructive feedback ● waste time and energy managing their behaviors for effect
Fear of Conflict	<ul style="list-style-type: none"> ● create environments where back-channel politics and personal attacks thrive ● fail to tap into all the opinions and perspectives of team members ● ignore controversial topics that are critical to team success
Lack of Commitment	<ul style="list-style-type: none"> ● breeds lack of confidence and fear of failure ● revisits discussion and decisions again and again ● encourages second-guessing among team members ● creates ambiguity among the team about direction and priorities
Avoidance of Accountability	<ul style="list-style-type: none"> ● places an undue burden on the team leader as the sole source of discipline ● misses deadlines and key deliverables ● creates resentment among team members who have different standards of performance ● encourages mediocrity
Inattention to Results	<ul style="list-style-type: none"> ● is easily distracted ● encourages team members to focus on their own individual goals

2.2.4 Evaluation of Teaming

One of the major concerns from instructors is the assessment of individual learning, which is made more challenging in a collaborative environment. With the possibility of social loafing, there is also the concern that some students will coast on their team member's work and not put forth the effort required to learn the course material (Karau and Williams, 1993, Brooks and Ammons, 2003; Loughry, Ohland, and Woehr, 2014). Quite often, instructors turn to the use of peer evaluations to evaluate the individual contributions of team members on group work or during group activities (Loughry, Ohland, and Woehr, 2014, Svinicki and McKeachie, 2014). It is important to distinguish the difference between peer evaluations used as a grading mechanism for a student product and peer evaluations to evaluate collaboration. For the purposes of this dissertation, the term peer evaluation is used to indicate evaluations that measure the collaboration occurring within student groups.

The goal of peer evaluation in a collaborative environment is to have the evaluators rate themselves and their teammates on how well they carry out their assigned team functions, meeting attendance and preparation, contributions to the team, and cooperation (Felder and Brent, 2001). Research has shown that the purpose of the evaluation has a large effect on how these evaluations are completed. For example, a formal evaluation used for assessment purposes versus an informal evaluation used for development is more likely to be inflated (Brutus, Donia, and Ronen, 2013) or may inhibit cooperation (Boud, Cohen, and Sampson, 1999). Additionally, evaluations

focused on evaluating effort were more unreliable than those focused on achievement (Topping, 2003).

However, peer evaluations used for development processes (e.g., at the middle of the semester), can have a potential to improve their team skills before the end of their collaborative work (Baker, 2008). Additionally, these formative assessments have shown to reduce the frequency of social loafing (Karau and Williams, 1993; Brooks and Ammons, 2003; Brutus and Donia, 2010). For example, a study by Brooks and Ammons (2003) demonstrated that when student saw the peer evaluation process as a way to reduce social loafing, they had a greater satisfaction with their team. This result occurred when students provided specific feedback earlier in the team experience (Brooks and Ammons, 2003). A study of 511 undergraduate business and graduate MBA students using developmental feedback showed immediate impacts on open communication, group task focus, group viability, and member relationships within their assign project groups (Druskat and Woehr, 1999).

There are many benefits to using peer feedback to evaluate collaborative activities. Peer evaluation allows for aggregation of different points of view which helps to partially remove bias in evaluation (Brutus and Donia, 2010). Evaluating peers demonstrates gains in team skills such as cognitive challenge and development whereas evaluating yourself enhances meta-cognition (Topping, 2003). Peer evaluation allows for true measurements of interpersonal skills, team-work effectiveness, and individual's strengths and weaknesses (Brutus and Donia, 2010). A study by Kaufman, Felder, and Fuller (2000) showed significant positive correlations between peer ratings and test

scores indicating that students are able to evaluate knowledge and performance-based skills accurately (Kaufman, Felder, and Fuller, 2000). Additionally, when peer feedback is received thoughtfully and positively, it has many positive effects on learning (Topping, 2003).

There are some issues regarding the use of peer evaluations that instructors should be aware of. A study by Kaufman, Felder, and Fuller (2000) found that male students in a sophomore-level chemical engineering course gave and received higher ratings from their teammates than female students, although these differences were not statistically significant. However, minority students tended to give higher ratings and receive lower ratings than their teammates, which was found to be statistically significant in one of the courses. This same study found that there is a greater tendency for students to give themselves a lower evaluation than it is to give themselves a higher evaluation (14% vs 6%) (Kaufman, Felder, and Fuller, 2000).

A study by the United States Air Force Academy (2010) found that engineering students on multidisciplinary teams often demonstrated the creation of subgroups of like-majored individuals and subsequently rated others with their major higher than students of a different major (Dudevior, Laffely, and Mundy, 2010). A study by Topping (2003) found that advanced students tended to underestimate themselves whereas weaker students tended to overrate themselves (Topping, 2003). Richmond, Satyamurthy, and DeFranco (2016) found that in a graduate engineering program, the opposite trend occurred where students tended to rate themselves higher than their peers on evaluations (Richmond, Satyamurthy, and DeFranco, 2016).

A study by Cheville and Duvall (2008) showed changes in how students use the peer evaluations as they gained additional experience with their team. For example, with an initial team evaluation, students rated each other primarily due to their personal attributes, but in subsequent evaluations rated each other more on the amount and quality of work completed by each team member (Cheville and Duvall, 2008). However, students were unable to distinguish between effort and results. Meaning that despite the final grade received on the project, students used the amount of perceived work as the impetus for the peer rating given and not the final grade (Cheville and Duvall, 2008). Additionally, it was noted that those students who completed the most work received the highest ratings by their peers, but also gave the lowest and most variable ratings to their teammates (Cheville and Duvall, 2008).

While the previous studies focused primarily on quantitative ratings, peer evaluations often contain comments for teammates as well. These comments can often add insight into the team functionality. However, Harley (2013) states that peer comments from evaluations tend not to give focused feedback (i.e., feedback that focuses on specific behaviors), but instead focuses on attributes (i.e., vague feedback that says nothing about the actions or behaviors of the team member) (Harley, 2013). This would indicate that the information learned by an instructor on the team functionality could be severely limited. For example, in a 2018 study by Brawner, Murch, Ferguson, and Ohland, only about a third of the students provided comments to themselves or their teammates and these comments lacked details that explained their quantitative rating of their peer (Brawner, Murch, Ferguson, and Ohland, 2018).

2.2.5 Strategies for Mitigating Team Issues

With the challenges inherent in both creating, implementing, and assessing the teaming experience, it is important to review opportunities for improvement. Hackman (1990) investigated the research surrounding the conditions necessary for creating effective teamwork and identified specific points of leverage for enhancing group performance as shown in Table 2.2. While instructors can create the teams and team projects and facilitate skill development, within the first-year engineering program, a place where instructors can most help student skills is in the third row: task-appropriate performance strategies. By helping the students develop effective strategies, instructors can help assist in other areas of the table as well.

For example, under group structure, it is suggested that group norms be created that regulate member behavior as a stepping stone toward effective teaming (Hackman, 1990). Felder and Brent (2001) also recommend having teams set ground rules, norms, or team contracts early and revisit and evaluate these initial expectations as a recommended instructional strategy for team building. This procedure is followed in the first-year engineering program at Michigan Tech. One of the earliest team assignments in the semester is to create a team contract that establishes group norms for the semester.

Another opportunity that appears to be a good source for instruction in the first-year engineering environment is under coaching and consultation: remedying implementation problems and fostering creativity in strategy development (Hackman, 1990). With the need for group assessment and the challenges demonstrated by researchers in student use of these tools (Brawner, Murch, Ferguson, and Ohland, 2018;

Cheville and Duvall, 2008; Dudevior, Laffely, and Mundy, 2010; Harley, 2013; Kaufman, Felder, and Fuller, 2000; Richmond, Satyamurthy, and DeFranco, 2016; Topping, 2003), this would indicate the need to focus on the instruction and use of peer evaluation tools in the classroom.

Table 2.2. Points of Leverage for Creating Conditions That Enhance Group Task Performance (Hackman, 1990, pg. 13)

Process Criteria of Effectiveness	Points of Leverage		
	Group Structure	Organizational Context	Coaching and Consultation
Ample effort	Motivational structure of group task	Organizational reward system	Remedying coordination of problems and building group commitment
Sufficient knowledge and skill	Group composition	Organizational education system	Remedying inappropriate “weighting” of member inputs and fostering cross-training
Task-appropriate performance strategies	Group norms that regulate member behavior and foster scanning and planning	Organizational information system	Remedying implementation problems and fostering creativity in strategy development

See Appendix H for documentation of permission to reprint this material.

Some suggestions by researchers on how instructors could improve the peer evaluation process include instructional methods, practice, and changes to the peer evaluation forms. To begin, Jamnia (2018) has outlined several tips to giving qualitative feedback that could be provided to students before they complete the peer evaluation form. First, the feedback should focus on a specific behavior or deliverable (e.g., product, report). This feedback should state the facts, offer validation (e.g., keep doing what

you're doing) or suggestions for improvement, and invite their opinion. The feedback should be provided within a few minutes (e.g., don't take too long to say what needs to be said). Feedback should be focused on the future (e.g., how can this be improved for next time). Instructors should also think about how to deliver the feedback – pick a time that works for both parties involved and a neutral location. Finally, the feedback should end with an agreement on what to do next (Jamnia, 2018).

To address the lack of detail in peer feedback comments as noted by Brawner, Murch, Ferguson, and Ohland (2018), instructors could provide these suggestions by Harley (2013) on how team members can describe the actual behaviors the team member is demonstrating. Examples of some common team feedback (attributes versus behaviors), as shown in Table 2.3 below, could be provided to the students so that their team feedback could become more focused and informative (Harley, 2013).

Table 2.3. The Differences between Attributes and Behaviors in Team Member Feedback

Attributes (Vague Feedback)	Behaviors (Focused Feedback)
Team Player	Helps other members of the team with their work when needed.
Difficult to work with	Questions every decision that's made. They argue all the time.

Adapted from Harley, S. (2013). *How to Say Anything to Anyone*. Austin, TX: Greenleaf Book Group, 117.

Additionally, as Topping (2003) indicated, when peer feedback is received thoughtfully and positively, it has many positive effects on learning (Topping, 2003). An instructor could provide instruction on a model for receiving feedback such as the S.A.R.A model, shown in Figure 2.3. SARA stands for Shock, Anger, Resistance and

Acceptance and is a useful model for visualizing the spectrum of emotions and reactions one might experience when receiving feedback (decisionwise, 2020; Flanagan, 2017; Harley 2018). The initial response to feedback (especially negative feedback) is shock and defensiveness and possibly resistance to the message. By understanding the S.A.R.A model and giving themselves time to think before reacting, this maximizes the chance that students will receive honest feedback in the future. (Harley, 2013)

Figure 2.3. SARA model for Receiving Feedback (decisionwise, 2020)
See Appendix H for documentation of permission to use this material.

In addition to training that can be provided by the instructor, several studies indicate that more instruction and additional practice with assigning ratings on a peer evaluation improves rater performance and accuracy (Kaufman, Felder, and Fuller, 2000; Topping, 2003; Bernadin and Vilanova, 2005). For example, a 2013 study of 182

undergraduate business students by Brutus, Donia and Ronan, found that repeated use of a standardized peer evaluation instrument increased the quality of the evaluation, the student's confidence in using the instrument, and the number of positive comments provided. A second study by Brutus and Donia (2010) indicated that repeated use of a peer evaluation instrument showed an increase in group effectiveness (Brutus and Donia, 2010).

Peer Evaluation software such as the CATME tool can be used repeatedly for formative, informal feedback as well as formal feedback. There is an option for rater calibration, which familiarizes team members on these teaming behaviors and the use of the CATME system to rate their peers (Ohland et al, 2012). Training is available through short YouTube videos on how to complete a peer evaluation. Additionally, CATME has the ability to identify teams having teamwork difficulties by flagging exceptional rating patterns that warrant attention by the instructor such as a team conflict or a manipulative student (Ohland and Loughry, 2006). This feature is helpful as Dudevior, Laffely, and Mundy (2010) recommend that faculty need to identify and rectify student bias behavior when it is observed (Dudevior, Laffely, and Mundy, 2010).

Finally, modifications could be made to the peer evaluation tools themselves to improve utility. For example, an examination of four different peer evaluation measures by 669 undergraduate business students indicated that students prefer that evaluations of the contributions of group work include: a) student input versus relying solely on instructor grades, b) self-evaluation as well as peer evaluation, c) clear, specified evaluation criteria, d) a place for open-ended comments, and d) confidentiality in

responses (Kilpatrick, Linville, and Stout, 2001). Additionally, this study indicated that having the opportunity to provide mid-term feedback on individual performance within a team is preferable to not having this opportunity (Kilpatrick, Linville, and Stout, 2001).

Several studies have demonstrated that increasing the visual size of the response box can result in longer responses for paper surveys (Christian and Dillman, 2004) and in web-based surveys for less motivated individuals (Smyth, Dillman, Christian, & McBride, 2009). Kalton and Schuman (1982) found in their review of the effects of questions on survey responses that increasing the length of the question significantly increased the length of the response. Additionally, other studies have indicated that providing clarifying or motivating instructions can improve the thoroughness of responses to open-ended questions (Smyth et al., 2009; Chaudhary & Israel, 2016).

Another method to change the peer evaluations is to change the structure of these evaluations. Most commonly peer evaluations use a combination of quantitative ratings and open-ended comments, but other options exist for gaining feedback on the teaming experience. A team Cognitive Task Analysis (CTA) utilizes an interview protocol that can identify the thoughts, behaviors, and decisions the individuals and team were making throughout the project instead of just the steps needed to complete the project and who was responsible for completing these steps (Klein, 2000). A team CTA can be used to improve a team's cognitive performance (e.g, decision making, problem solving) by identifying their strengths and weaknesses. However, a team CTA is less effective in assisting with motivational or emotional issues (Klein, 2000).

A team Knowledge Audit can also be used to gather information regarding team members knowledge and skills on specific team tasks (Crandall, Klein, & Hoffman, 2006). A Knowledge Audit is a structured interview that covers eight dimensions using specifically worded probes (questions) (Crandall, Klein, & Hoffman, 2006). A study by Miller, Hirshfield, and Chachra (2016) on 31 undergraduate engineering students at several midwestern universities used a similar interview strategy to gather negative comments about their students teaming experiences. Students were asked at the end of the semester about their team's decision-making process, a challenging situation their team faced (and did well/poorly), and which tasks they felt their team trusted them to complete. These interviews yielded detailed positive and negative comments about their teaming experiences (Miller, Hirshfield, and Chachra, 2016).

2.3 Research Goals and Design

The literature indicates that there are many challenges with collaborative learning (Felder & Brent, 2001; Loughry, Ohland, and Woehr, 2014) and teams in general (Lencioni, 2007). Forming collaborative groups, supporting effective team interactions, and evaluating individual work can be a difficult task for instructors (Felder & Brent, 2001). Many instructors are moving to peer evaluations to help assess students work and evaluate team effectiveness, but these evaluations may not accurately reflect the team situation (Brawner, Murch, Ferguson, and Ohland, 2018; Brutus, Donia, and Ronen, 2013; Cheville and Duvall, 2008; Dudevior, Laffely, and Mundy, 2010; Felder & Brent, 2001; Kaufman, Felder, and Fuller, 2000; Loughry, Ohland, and Woehr, 2014; Svinicki and McKeachie, 2014; Richmond, Satyamurthy, and DeFranco, 2016; Topping, 2003).

Because these challenges exist, the overall goal for my dissertation was to improve the utility of student peer evaluation feedback. As mentioned previously, I used the Design-Based Research (DBR) approach for educational research to frame my work. I utilized the suggestions from my literature review to identify possible solutions and used pilot studies (Chapter 3 and 4) to narrow the focus of this study to first-year engineering students at Michigan Tech. From this body of literature, I identified alterations in the peer evaluation instrument (CATME), changes to instructions for students regarding peer evaluation, and the introduction of a new instructional method which utilizes Knowledge Audit procedures that may result in increased utility of peer feedback. These alterations were tested over multiple semesters and classes to refine the procedure (Chapters 5 and 6). To assess this work, I created a coding template for instructors to evaluate the utility of the qualitative student comments provided via the peer evaluation instrument. My overall conclusions are discussed in Chapter 7.

3 Study 1

As mentioned previously, the success for Michigan Tech's first-year engineering program hinges on our approximately 240 student teams each semester learning to work collaboratively and cooperatively in their classes. The purpose of this first study was to solicit the opinions of instructors and peer mentors regarding the teaming issues they felt were most common in first-year engineering students at Michigan Technological University. This information was essential to determine the foundational teaming issues to focus the remaining studies.

3.1 Methods

3.1.1 Participants

The first-year engineering instructors ($n = 9$) and LEAP Leaders ($n = 51$) were surveyed for this study. A total of four instructors and 18 LEAP Leaders consented to and completed the survey. In the next section, I will discuss the survey in more detail.

3.1.2 Materials

To assess opinions regarding the most frequent teaming issues observed, I used the list of negative characteristics for the five dysfunctions of a team proposed by Lencioni (2007) and converted this into a Google Form survey for the faculty and LEAP Leaders in the first-year engineering program to complete. As mentioned previously, Lencioni's model has been shown to demonstrate both construct and predictive validity with team effectiveness, in addition to having face validity (Hamlin, 2008). See Appendix A for the complete survey used in this study.

3.1.3 Design and Procedure

This survey was intended to identify the five most common negative teaming behaviors observed in first-year engineering teams. As this is the first teaming experience at the University level for most students, I hypothesized that the primary issues were in the Absence of Trust and Fear of Conflict team dysfunction categories from Lencioni's model. Descriptive statistics were conducted to identify those primary issues to inform future studies.

I solicited first-year engineering faculty and LEAP Leaders to complete the Google Form survey during week 12 of the semester. At this time, students had been in their teams for 9-10 weeks, which allowed for 18 to 20 Studio Sessions for instructors and LEAP Leaders to observe team interactions and 8 to 9 additional LEAP Sessions for LEAP Leaders to observe team interactions. From the list of negative behaviors, actions, or consequences, participants were asked to identify those they see most frequently in first-year engineering teams by selecting the checkbox next to those items in the list. They were asked to only include those items they felt were in the top five most frequently occurring items that they have personally observed in the teams they have instructed.

3.2 Results

The results from this survey are shown in Table 3.1. The top two issues identified did belong to the first two teaming dysfunctions identified by Lencioni (2007): Lack of Trust and Fear of Conflict. A total of 70% of those surveyed stated that students "hesitate to ask for help or provide constructive feedback" from the Absence of Trust dysfunction. The second most common issue was that students "fail to tap into all the opinions and

perspectives of team members” from the Fear of Conflict dysfunction. Sixty-one percent of respondents rated this as one of the top five dysfunctions experienced by first-year engineering teams. Additionally, five of the top ten issues identified by respondents are contained within these first two dysfunctions. While the remaining five issues were from the top two dysfunctions (Lack of Accountability and Attention to Results), there were none identified from the third dysfunction - Lack of Commitment.

Table 3.1. Top 10 Team Issues Identified by First-Year Engineering Faculty and LEAP Leaders in Fall 2019

	Team Issue	Dysfunction
1	hesitate to ask for help or provide constructive feedback (70%)	Lack of Trust
2	fail to tap into all the opinions and perspectives of team members (61%)	Fear of Conflict
3	is easily distracted (57%)	Attention to Results
4	encourages mediocrity (48%)	Lack of Accountability
5	misses deadlines and key deliverables (39%)	Lack of Accountability
6	creates resentment among team members who have different standards of performance (35%)	Lack of Accountability
7	fail to recognize and tap into one another's skills and experiences (30%)	Lack of Trust
8	ignore controversial topics that are critical to team success (26%)	Fear of Conflict
9	conceal their weaknesses and mistakes from one another (26%)	Lack of Trust
10	places an undue burden on the team leader as the sole source of discipline (22%)	Lack of Accountability

3.3 Discussion

These results correspond well with the hypothesis that first-year engineering teams struggle most with the first two dysfunctions (Absence of Trust and Fear of Conflict) as behaviors from these dysfunctions were identified as the top two issues experienced by first-year engineering teams. While there were several items from the Lack of

Accountability dysfunction identified as issues, Lencioni (2007) points out that the Five Dysfunctions of a Team Model is a hierarchy that builds on each step. If teams are unable to build a foundation of trust, they cannot engage in productive conflict, which ensures that the team will have a lack of commitment. With a lack of commitment, team members develop an avoidance of accountability which then leads to inattention to results (Lencioni, 2007). So, while the largest number of issues occur with Lack of Accountability (step 4), if there are issues present in the lower tiers of the model, these foundational issues should be addressed first. Additionally, several of the behaviors characterized by the Absence of Trust and Fear of Conflict dysfunctions (e.g., items 1, 2, 8, and 9 listed above) support the need to develop and use instructional materials for student use in providing and receiving feedback to help address these dysfunctions.

4 Study 2

The first study identified that several of the top teaming issues experienced by first-year engineering students might be addressed by additional support or development of instructional materials for students to use in the peer feedback process. This second study performed in Fall 2019 in ENG1101 was used to determine the students' peer evaluation behaviors using the current peer evaluation instrument, CATME. The information gained from this study was important for understanding students' evaluative behaviors prior to any instruction, training, or feedback manipulations.

4.1 Methods

4.1.1 Participants

This study was conducted within the first-year engineering program at Michigan Tech, with the students enrolled in ENG1101: Engineering Analysis and Problem Solving in Fall 2019 ($n = 769$). For this study, complete teams of students who completed the CATME midterm peer evaluation and consented to the study were included in the sample ($n = 163$ students, 43 teams, 26 sections). Below, I will discuss the CATME peer evaluation instrument.

4.1.2 Surveys

4.1.2.1 CATME Peer Evaluation Survey

The CATME Teamwork software (www.catme.org) was used to collect structured peer evaluations at the midpoint and end of the semester. The CATME peer evaluation tool is currently being used by over 22,000 instructors in 2,400 institutions in 88 countries to

evaluate the effectiveness of academic teams (info.catme.org). Students were asked to rate themselves and their teammates on five dimensions of teaming (Contributing to Team's Work, Interacting with Teammates, Keeping the Team on Track, Expecting Quality, and Having Related Knowledge and Skills). This tool is designed as a behaviorally anchored rating scale (BARS), which provides a description of specific observable behaviors for each performance level (Ohland, et al., 2012). For example, behaviors for the Contributing to the Team's Work category might include "completes a fair share of the team's work with acceptable quality" or "keeps commitments and completes assignments on time". See Appendix B for the complete survey, with questions and observable behaviors defined for each of the dimensions of teaming. These ratings are translated into a 5-point Likert scale and two teaming factors are calculated, one that takes into account self-evaluations and one that does not.

This tool can be used repeatedly for formative, informal feedback as well as formal feedback. There is an option for rater calibration, which was used by the ENG1101 students, which familiarizes team members on these behaviors and the use of the CATME BARS system (Ohland et al, 2012). The rater calibration function walks the students through the CATME instrument by providing a description of four hypothetical teammates and then asking the student to rate each of these teammates on each of the five dimensions of teaming (Contributing to Team's Work, Interacting with Teammates, Keeping the Team on Track, Expecting Quality, and Having Related Knowledge and Skills). Using this rating scale and allowing for rater calibration has been shown to

reduce rater ambiguity and increases interrater agreement on the rating of teaming dimensions to 0.71 (3-person team) or 0.80 (4-person team) (Ohland, et al, 2012).

The last component of this tool is the ability for students to provide comments to each other to explain how they evaluated each other. First, the student has the ability to provide comments on themselves and their performance to the instructor. Next, they have the opportunity to provide comments to each of their teammates. These comments can be made identifiable or anonymous by the instructor. For ENG1101, these comments were anonymous. There is a short instructional video linked above the comment instructions on how to provide peer-to-peer comments. It is unknown how many ENG1101 students selected to watch this video.

4.1.3 Design and Procedure

This research was exploratory to assess how students use the CATME evaluation instrument. For this study, I hypothesized that there would be differences in the quantity of positive and negative comments in the written feedback to teammates. Specifically, I hypothesized that few students would provide negative or constructive feedback to their teammates.

At the beginning of the semester (weeks 2-3), ENG1101 instructors created teams for the semester. These teams collaborated on activities and projects in the class. Students in ENG1101 completed a midterm and end-of-term peer review of themselves and their teammates using the team evaluation software CATME (www.catme.org). Prior to

completing their first peer evaluation (midterm evaluation) for their teammates, students completed the rater calibration on an example team as described above.

ENG1101 students completed their midterm peer evaluation at the end of week 8 of the 14-week semester as part of a formative assessment. These evaluations were intended to provide students an opportunity to offer their teammates with anonymous feedback on their team performance at this mid-point of team work. At this point, the teams had created their team contracts and worked together on several small projects for the class. Additionally, the teams were encouraged to work on in-class activities together during the Studio and LEAP Sessions. In week 9, students completed the consent form for this study as well.

4.2 Results

This section focuses on the results of the broader exploratory study to assess how first-year engineering students use the CATME evaluation instrument. To begin, Figure 4.1 shows an example team report in CATME. Each row represents a single student's ratings on the five teaming dimensions. The first row represents rater 1, their ratings for themselves on the five dimensions and their ratings for each of their three teammates on these dimensions. Next in the row are the adjusted rating taking into account their own ratings and their adjusted rating without their ratings, which I will discuss in more detail below. The next items in the report show if the student gave the same rating to all team members within a teaming dimension or if they gave the same rating on all teaming dimensions for a teammate. Finally, there is a column for notes if the CATME system has

flagged any evaluations which may warrant additional attention by the instructor such as an overconfident or underconfident student, a potential conflict, a clique, a manipulative student, or students who are rated uniformly high or low by all teammates (Ohland and Loughry, 2006).

Rater #	Rater 1					Rater 2					Rater 3					Rater 4					Adj Factor (w/ Self)	Adj Factor (w/o Self)	Same Rating Across Team Members On Dimensions	Same Rating Across Dimensions On Teammates	Note
	C	I	K	E	H	C	I	K	E	H	C	I	K	E	H	C	I	K	E	H					
1	2	3	2	3	2	2	2	4	5	3	3	3	3	2	3	4	4	4	4	3	0.78	0.81	E		
2	5	4	5	3	5	4	3	5	5	4	5	3	3	3	4	5	5	5	5	5	1.05	1.05	E		
3	4	4	4	3	4	4	3	5	5	4	4	3	3	3	4	3	3	4	4	5	1.00	1.02	I K E		
4	5	4	5	3	5	5	4	5	5	4	4	3	3	3	4	4	5	5	4	5	1.05	1.05		2	

Figure 4.1. Example CATME Team Report

You will notice that in this example, rater 1 gave all teammates the same rating on Expecting Quality, as shown in the fourth column from the left, under heading *Rater 1* and subheading *E* (which symbolizes for Expecting Quality). Rater 3 gave all teammates the same rating for three of the teaming dimensions (Interacting with Teammates, Keeping the Team on Track, Expecting Quality). Another thing to note is that rater 4 gave rater 2 the same rating across all five of the teaming dimensions. Additionally, rater 1 and rater 3 gave themselves lower ratings than their teammates did as demonstrated by the lower adjusted score when taking into account self-ratings.

These adjusted CATME scores are calculated using Equation 1 where M_S is the average rating for the student by the student and their teammates on all teaming dimensions and M_T is the average rating for the team on all teaming dimensions.

$$CATME = \frac{M_S}{M_T} \quad (1)$$

Zooming in a little on the numerical results, as shown in Figure 4.2 below, I'll discuss how these results are calculated. If a student's average CATME ratings (rater 1, shown circled in red) on the five dimensions of teaming, $M_S = 60/20$ or 3.00, and the team's average rating on all five dimensions (M_T) (shown circled in blue) equals $307/80$ or 3.84, the CATME score for rater 1 would be 0.78 (adjusted factor with self). If the self-ratings (circled in green) were removed from the average rating for the student, M_S and average rating for the team, M_T , the adjusted CATME score w/o self can be calculated. The new value for M_S would be equal to $48/15$ or 3.20, while $M_T = 295/75$ or 3.93, and the total CATME score (adjusted factor w/o self) is now equal to $3.2/3.93 = 0.81$.

Rater #	Rater 1					Rater 2					Rater 3					Rater 4					Adj Factor (w/ Self)	Adj Factor (w/o Self)
	C	I	K	E	H	C	I	K	E	H	C	I	K	E	H	C	I	K	E	H		
1	2	3	2	3	2	2	2	4	5	3	3	3	3	3	2	3	4	4	4	3	0.78	0.81
2	5	4	5	3	5	4	3	5	5	4	5	3	3	3	4	5	5	5	5	5	1.05	1.05
3	4	4	4	3	4	4	3	5	5	4	4	3	3	3	4	3	3	4	4	5	1.00	1.02
4	5	4	5	3	5	5	4	5	5	4	4	3	3	3	4	4	5	5	4	5	1.05	1.05

Figure 4.2. Example CATME Team Report (numeric results only)

In general, the following statements regarding the CATME factor score are true regarding the student and team. First, if $M_S < M_T$, then $CATME < 1$. This would indicate that that particular individual was ranked lower than their teammates. Conversely, if $M_S > M_T$, then $CATME > 1$. This would indicate that that particular individual was ranked higher than their teammates.

To understand the behavioral profile of students using the CATME evaluation tool, I first examined how students assigned ratings within a CATME teaming

dimensions. Overall, 77% of students rated all of their teammates the same on at least one of the behavioral dimensions. In other words, students didn't appear to differentiate their teammate's contributions within a teaming dimension. For example, in Figure 4.2 Rater 3 rated all their teammates the same on Interacting with Teammates, Keeping the Team on Track, and Expecting Quality. This may warrant further investigation to determine why students are demonstrating this behavior. For example, it could be that students were unable to comprehend what the rating was intending (and thus just rated everyone the same), or that they were unable to discriminate between teammates due to a lack of observation of that dimension given the teaming tasks they had engaged in.

I also examined how students rated their teammates across all the dimensions (e.g., did they give a teammate all 4's on each teaming dimension). In general, 61% (n = 97) of students evaluated their teammates differently across teaming dimensions. Compared with only 23% of students differentiating teammate behavior within a CATME teaming dimensions, it appears that students are more willing to rate each other differently across the dimensions than differentiating between teammates within a teaming dimension. Looking back at Figure 4.2 is a good illustration of this phenomena. In this example, three of the team members (Raters 1, 2, and 3) provided the same rating on at least one dimension for all their teammates, whereas only one student provided the same rating for a teammate across all of the teaming dimensions (e.g., Rater 4 gave 5's on all dimensions for Rater 2).

It is important to note that there is a significant positive correlation between the lack of differentiation within both teaming dimensions and teammates ($r = 0.422$, $p <$

0.001). This would indicate that students who are more likely to provide the same ratings within all the teaming dimensions are also more likely to rate their teammates the same on all the teaming dimensions. The reverse would be true as well, with students who are more likely to differentiate ratings within a dimension (i.e., provided different ratings for each teammate within one dimension) are also more likely to rate a given team member differently between the dimensions.

Looking more specifically at the qualitative results (peer evaluation comments), I hypothesized that fewer students would provide negative or constructive feedback to their teammates. To test this hypothesis, I coded the comments on the midterm evaluations for the instructor, for the student, and for their teammates to specifically look at two main categories. First, I coded the comments as either blank (0), negative (1), or positive (2). Next, I coded the comments relayed to the team as either inconsistent (1) or consistent (2) with comments relayed to the instructor. In other words, are the students willing to let their teammates know of the issues they bring to the instructor's attention?

Looking at the comments provided on these midterm evaluations, only 11% of the students ($n = 17$) provided negative comments to the instructor regarding their team. Of these, approximately half ($n = 9$) also provided this feedback to their teammates. Overall, 94% of students did not report negative comments to their teammates, and 89% did not report negative comments to the instructor – perhaps suggesting that students are simply hesitant to be the one to “tattle” on their teammates. Even when negative feedback was anonymous, less than 50% of the small number of students who reported negative experiences to the instructor also provide information regarding the issue to their

teammates. Only 15% of students ($n = 24$) provided positive comments to their instructor regarding their team. Of these, approximately 75% ($n = 18$) also provided this feedback to their teammates. This demonstrates that students were more willing to provide positive feedback to their teammates than negative.

Looking at the comments a little deeper, I expanded the coding categories. Comments made to the rater and their teammates were additionally classified as team related (1 = not team related, 2 = team related) and/or constructive (1 = not constructive, 2 = constructive). For this study, “constructive” was operationally defined as comments that suggested actions or behaviors that should be added, stopped, altered, or maintained as is. Ninety-five percent of the comments provided to the teammates were team related, such as “Communicates efficiently. Will ask for help from teammates if stuck on a problem.” Comments such as “NAME does a fine job”, with no team context, were coded as not team related and accounted for (5%) of all comments.

Approximately 45% of students ($n = 69$) did not provide constructive comments for either themselves or their teammates. This again seems to indicate, at minimum, a hesitancy by team members to provide constructive feedback to their teammates. An additional 36% ($n = 56$) provided either constructive comments for themselves or for their teammates, but not both. Only 19% of students ($n = 29$) provided constructive comments to both themselves and their teammates.

4.3 Discussion

To summarize, this study yielded the following information on the peer evaluation behavior of first-year engineering students using the CATME evaluation tool. First, I hypothesized that a low percentage of students would provide negative or constructive feedback to their teammates. In general, there were students who focused on evaluating their teammates across the dimensions thoughtfully (as evidenced by different ratings across dimensions and teammates) and those who do not. This is consistent with the findings of several faculty at the University of Michigan with students in their Multidisciplinary Design Program, who concluded that students were unable to effectively utilize and/or interpret the peer feedback provided from the CATME peer evaluation tool (Adams, DeGraaff, and Hohner, 2015).

Looking at comments, only 55% of ENG1101 students provided feedback (of any kind) to themselves and/or their teammates using this evaluation method. While this may not be a “low” number, it does not bode well if approximately half of the students in the first-year engineering program are unwilling to provide feedback to their teammates. Overall, only 26% of students provided any type of comments on their team functionality to their instructor (either positive or negative).

The CATME tool has provided a way to ensure students complete training provided in the form of Rater Practice for the CATME behavioral ratings. Additionally, while there is a short instructional video created by CATME on how to provide peer-to-peer comments questions, this is not mandatory and is also located in a position on the survey where it was unlikely to be seen by the students. This demonstrates a need for

training for students with respect to providing or receiving constructive feedback on peer evaluation surveys, which could lead to reducing the top teaming issues identified in Study 1.

5 Study 3

As mentioned previously, teaming is an important skill for both industry and academia and providing the most experience with effective teams is essential to the long-term success of undergraduate engineering students. With the active, collaborative learning environment used in the first-year engineering course instruction, it is essential that teams learn how to overcome common team dysfunctions and work together collaboratively. The two studies from Fall 2019 provided key information regarding how the first-year engineering teams at Michigan Technological University function.

From the first study, the common dysfunctions that a first-year engineering team face were identified from surveying the first-year engineering faculty and near-peer mentors. As expected, one critical team issue was that students hesitate to ask for help or provide constructive feedback which is a characteristic of the Absence of Trust, the first dysfunction from Lencioni's model (2007). An intervention that provides a strategy to facilitate giving and receiving feedback could effectively address this dysfunction. Such an intervention would be consistent with the recommendations of Hackman (1990) for promoting team effectiveness.

The second study demonstrated that there is a need for instruction on giving and receiving feedback. In general, there are students who focus on evaluating their teammates across the dimensions thoughtfully and those who do not. Formative feedback in the form of anonymous comments resulted in only approximately half of the students providing any type of constructive feedback to themselves or their teammates, which is

consistent with my initial hypothesis which stated that few students would provide negative or constructive feedback to their teammates.

With the need outlined in the Fall 2019 studies for instructional strategies on giving and receiving feedback, I completed a study in the Spring of 2020 that focused on developing an instructional intervention to address this issue. As mentioned previously, the CATME software allows for students to provide themselves and each other with comments regarding their team experiences (Ohland et al., 2012). In ENG1101 in Fall 2019, and historically in the first-year program, this feedback was provided to the students anonymously (i.e., they did not know which team member provided the feedback).

However, while there is training provided in the form of Rater Practice for the Likert questions, the training provided for giving feedback does not appear to be effective. The results from study 1 indicates that with the use of the CATME survey as is, the top teaming issue identified by first-year engineering instructors and LEAP Leaders is the hesitation to provide constructive feedback. Additionally, this reluctance to provide constructive feedback was demonstrated in study 2 with only 55% of students providing constructive feedback in the evaluation. The instructional video provided by CATME is linked above the instructions for the peer-to-peer comments section as shown in Figure 5.1 below. Note that for an active team survey, the blanks behind “Comments to” would be automatically filled in with the student’s teammates.

Peer-to-Peer Comments

[<< Back](#)[Next >>](#)

[Student video on Peer to Peer Comments](#)

Please provide constructive comments about your fellow teammates as well as yourself. The purpose of these comments is to give you the opportunity to explain how you rated your peers and if there was behavior or experiences in particular that influenced you when doing your peer and self-evaluations.

NOTE: This survey is currently set up so that your teammates will read your comments but will not know that it was you that wrote them.

Comments to the instructor about your own performance

Comments to _____

Comments to _____

Comments to _____

Figure 5.1. Screen Capture of CATME Peer-to-Peer Comments Window

My goal for this study was to test two different instructional manipulations as an intervention to determine if they impacted feedback to other team members, team satisfaction, team interdependence, trust, and productive conflict in the first-year engineering classroom. I summarize the specific methods, research design, hypotheses, and results of this study in the following sections.

5.1 Methods

Spring 2020 saw many changes to the academic landscape, the primary one being the emergence and global spread of the COVID-19 virus. This necessitated some major changes to my study design (and consequently resulted in conducting one more study during Fall 2020, discussed in Chapter 6). I will provide an overview of the intended study and changes made to the study in response to this situation here. The remainder of this chapter discusses the methods, design, and results used during Spring 2020.

I intended to measure feedback, team satisfaction, and team interdependence at three points during the semester (pre-intervention, immediately post intervention, and at the end of the semester). Additionally, I planned to measure conflict and trust at both pre-intervention and end of semester times. With the emergence of COVID-19, Michigan Technological University made the switch to remote instruction over spring break (between week 8 and 9 of the semester). At that time, in order to reduce the potential of confounds with the switch to remote learning, I administered the post trust and conflict measures and did not include the post CATME measures (team satisfaction, interdependence, and feedback) in the analyses. The following sections present the methods as they occurred post-COVID, which differed from what had been planned.

5.1.1 Participants

Participants were solicited from students enrolled in ENG1102: Engineering Modeling and Design during Spring 2020 ($n = 633$). There were 29 sections of this course offered and taught by 3 different instructors, each instructor was responsible for approximately 10 sections. Each section had between 13 – 24 students, with an average of 20 students.

A total of 472 students (63% male, 28% female, 9% did not identify) consented to this study.

Similar to ENG1101 (the Fall 2019 study group), ENG1102 was taught with 5 sections meeting together twice a week for two hours in a Studio Session with the instructor and LEAP Leaders (near-peer student mentors). Once a week, each section met with their LEAP Leader for a 50-minute LEAP Session. As with ENG1101, ENG1102 was taught in a flipped learning environment, with some course material required before attending each Studio Session. Although it is best practice for students to complete these materials before the Studio Sessions, some students chose not to complete this material. Because of this inconsistency in attention to the required materials for the flipped portion of class, any information presented as relevant to this study's manipulations (such as instructions related to the surveys and giving or receiving feedback) were administered consistently to all students during the Studio Sessions. In addition, relevant instructional videos and handouts were given to the students by the faculty within the Studio Sessions.

5.1.2 Materials

My goal for this project was to test two different instructional manipulations (Reflective and Procedural) to see if they had an impact on my dependent variables (feedback towards other team members, team satisfaction, team interdependence, trust, and productive conflict) in the first-year engineering classroom. The instruments I used to measure my dependent variables are described in the next section.

5.1.2.1 Surveys

5.1.2.1.1 CATME Survey

The CATME peer evaluation (outlined in section 4.1.2.1 and Appendix B), included both quantitative and qualitative measures and was used for formative and summative evaluations of teaming. Students used quantitative ratings for individual contributions (self and peer) on the five dimensions of teaming (Contributing to the Team's Work, Interacting with Teammates, Keeping the Team on Track, Expecting Quality, and Having Related Knowledge and Skills). Students also provided anonymous, qualitative peer feedback to the instructor, themselves, and their peers. One option for the CATME tool is to include follow-on surveys within the instrument. Two follow-on surveys were included and are discussed below.

5.1.2.1.2 Team Interdependence

The team interdependence scale is a follow-on survey that can be added to the CATME peer evaluation. This five-item scale is designed to measure task interdependence: the degree to which team members share materials, information or expertise (Van der Vegt, Emans, & Van De Vliert, 2001). Four items are phrased in a positive manner ("I depend on my colleagues to do my work properly") while one item is phrased on a reverse-scale ("I have a one-person job; I rarely have to check or work with others" (Van der Vegt, Emans, & Van De Vliert, 2001). The questions are on a five-point Likert-type scale from "*Strongly Disagree*" (1) to "*Strongly Agree*" (5). The scores on this scale can range from 5 to 25 with lower scores indicating lower team interdependence and higher scores

indicating higher team interdependence. See Appendix C for the complete survey. This survey was included as part of the evaluation of teaming.

5.1.2.1.3 Team Satisfaction

The team satisfaction scale was added to the CATME peer evaluation and was designed to measure satisfaction with being a team member (Gladstein, 1984). This measure uses a three-item scale based on the work by Gladstein (1984) and adapted by Van der Vegt, Emans, & Van De Vliert, (2001). An example question is “I am satisfied with my present teammates” (Van der Vegt, Emans, & Van De Vliert, 2001). The questions are on a five-point Likert-type scale from “*Strongly Disagree*” (1) to “*Strongly Agree*” (5). The scores on this scale can range from 3 to 15 with lower scores indicating lower satisfaction with the team and higher scores indicating higher satisfaction with the team. See Appendix C for the complete survey. This survey was included as part of the evaluation of teaming.

5.1.2.1.4 Trust and Conflict

Trust and conflict ratings were measured using the trust and conflict subscales from the Five Dysfunctions of a Team: Team assessment (Lencioni, 2012). Both subscales are eight-item scales. An example question from the trust subscale is “Team members admit their mistakes,” while an example item from the conflict subscale is “Team members are passionate and unguarded in their discussion of issues” (Lencioni, 2012). Each question is rated on a five-point Likert-type scale from “*Never*” (1) to “*Always*” (5).

A total score for each subscale is the average of the ratings within the items of each subscale. On each subscale, ratings below 3.24 indicate a low trust or conflict score

between teammates, which would indicate that this team is not comfortable being open, vulnerable, or engaging in unfiltered dialogue between teammates (Lencioni, 2012).

Ratings between 3.25 and 3.74 indicate that there is growth needed in these areas, while scores 3.75 and above indicate teams with high levels of trust and productive conflict.

See Appendix D for the complete survey and scoring interpretation of these subscales.

This survey was included as part of the evaluation of teaming.

5.1.2.2 Feedback Instruction Manipulations

There were two between-subjects feedback instructional manipulations I designed for this project (Reflective and Procedural) which will be discussed in more detail below. Each ENG1102 instructor had two 5-section Studio Sessions they instructed. The instructors used the Reflective Manipulation in their first section of the week (designated 1st Class in Figure 5.2 below) and the Procedural Manipulation in their second section of the week (2nd Class). I used this design to control for possible instructor effects, as all instructors used both manipulations. Timing of the classes was not expected to be an issue as students self-selected the sections of ENG1102 they were enrolled in during Spring 2020. However, the results were evaluated to determine if time of day effects existed. Note that this design did not allow me to assess if time spent in group may have affected the results, so this should be addressed in future studies.

Figure 5.2. Feedback Instruction Planning

I met with the ENG1102 faculty several times during the semester to review the logistics for the study, set up the surveys (conflict, trust, and CATME peer evaluations), go through the materials needed for the study (video and handouts) and to describe the instructions for the intervention. I also provided them with an instructional video for each manipulation and a physical handout for each student so they would have an immediate application for the feedback instructional method. Next, I review the different manipulations.

5.1.2.2.1 Reflective Manipulation

The Reflective Manipulation was based on the Cognitive Task Analysis (CTA) Knowledge Audit Procedure: a series of probing questions (Crandall, Klein, and Hoffman, 2006; Militello and Hutton, 1998). In my study, the questions were designed to help the students think critically about their team experience before completing the peer evaluation. This method was chosen as a team Knowledge Audit has been shown to elicit information regarding team process, member knowledge and skills, and specific

examples (Crandall, Klein, and Hoffman, 2006). This elicited information could then be used by the students when they complete their peer evaluation.

For this study, the Knowledge Audit was used for the ENG1102 students to focus on their team experience to date. Instead of using the full list of probes, I used the following three probes in order to focus the students into areas where I expected the students to have examples from their teaming experience they could discuss. Note the wording in the brackets is added to emphasize the connection to the team's work or experience:

- 1) Anomalies Probe: "Can you describe an instance when you knew something was amiss [with your team]?"
- 2) Opportunities Probe: "Can you think of an example when you were working [with your team] when you noticed an opportunity to do something better?"
- 3) Self-monitoring Probe: "Can you think of a time when you realized that you would need to change the way you [or another team member] were performing to get the job done?"

The important component with using the Knowledge Audit probes is that you "stick to the script" as these probes were designed to elicit the correct information if used properly (Militello, Hutton, Pliske, Knight, Klein, and Randel, 1997). My goal in using these probes was to determine if a) did they notice any team issues (anomalies probe), b) did they think of a way to work better as a team (opportunities probe), and c) did they identify behaviors that would be important to provide feedback on for the good of the team (self-monitoring probe).

During one of the Studio Sessions, students watched an instructional video describing how to use this method with specific case studies and completed the handout discussed in section 5.1.2.3 below. After this Studio session, students were asked to complete their reflection of their own team's experience using this method and then complete their peer evaluation using CATME.

5.1.2.2.2 Procedural Manipulation

The second manipulation provided instructions for a detailed procedure for giving feedback (Procedural Manipulation). The Procedural Manipulation was based upon the feedback procedures by Shari Harley (2013). Shari Harley is the author of How to Say Anything to Anyone and the pioneer behind Candid Culture, a training and consulting firm that promotes the use of candor to improve interpersonal relationships (Harley, 2013). This method was chosen as this feedback procedure has been shown to improve the quality of feedback (Harley, 2018) and follows the suggestions from Jamnia (2018) on effective tips for giving feedback. The original steps from Shari Harley's feedback formula are listed below: (Harley, 2013)

1. Introduce the conversation by asking for a few minutes to talk.
2. State your motive.
3. Describe the observed behavior.
4. Share the impact or result of that behavior.
5. Have some dialogue, and ask the recipient for his perception of the situation.
6. Make a suggestion or request for what you'd like the person to do next time.
7. Build an agreement on next steps.

8. Say “Thank you.”

For this project, Harley’s formula was modified from an eight-step version to a four-step version as some of the steps would not happen within the CATME peer evaluation software. For example, steps 1 and 2 (introducing the conversation and stating your motive) were not needed with the CATME. The instructor assigned the peer evaluation assignment, so there was no need to introduce a conversation. Additionally, there was no need to state the motive for the conversation as the student motive was to complete the assignment.

Steps 3 and 4 were needed. The student needed to describe the behavior(s) that they noticed (both positive and negative), discuss the impact of this behavior, and to be clear and specific. Steps 5 (asking what the other person thinks), 7 (building an agreement on next steps) and 8 (saying thank you) were not possible at this point in the peer evaluation process with a static peer evaluation form and were removed from the instructions for completing the peer evaluation form. Therefore, Step 6 (making a suggestion or request) was the last step in the process for completing this CATME peer evaluation. The final process presented to students to complete their CATME peer evaluation included the following four steps.

1. Discuss attributes and behaviors of your team to paint a picture
2. Discuss the impact of this behavior
3. Be clear and specific
4. Offer an alternative

As with the Reflective Manipulation, students watched an instructional video describing how to use the Procedural Method with specific case studies and completed the handout for this manipulation (discussed in section 5.1.2.3 below). After this Studio session, students were asked to complete their reflection of their own team's experience using this method and then complete their peer evaluation using CATME.

5.1.2.3 Instructional Materials

I provided all of the material for the study (videos, surveys, and handouts) to the instructors. This section details the creation of the videos and handouts used for each manipulation.

5.1.2.3.1 Videos

To be certain that the instructions for the intervention were the same between instructors, I created an instructional video for each method that the instructors showed to their classes during their studio sessions. To make the instructional manipulations as similar as possible, the videos were created in three parts: overall importance of providing feedback, specific instructions and examples for each manipulation, and information on how to receive and respond to feedback. The first and third sections were the same in both videos.

The first part of both videos was 3 minutes and 30 seconds long. In this section, I defined the purpose of team feedback (to maintain positive teaming behaviors or to cease negative teaming behaviors) and provided examples of how to be more specific with feedback using examples from Shari Harley's book How to Say Anything to Anyone. In

the video, participants were told that students often use unfocused, unhelpful feedback in their peer reviews for their teammates. Additionally, using words such as “team player” are considered vague feedback. “Team player” describes an attribute (or quality) of a good teammate, but does not describe the specific behaviors that make a teammate a “team player” (Harley, 2013). It also included examples of the desired focused feedback on evaluations (i.e., feedback that focuses on specific behaviors) such as “helps other members of the team with their work when needed” (Harley, 2013). This feedback provides an example of specific behavior that someone who is a “team player” may exhibit. Several examples of these types of feedback were provided in the video to provide context to both the positive and negative behaviors they might comment on in their peer evaluations.

The second section of each video was specific to the manipulation. For the Reflective Manipulation, this section of the video totaled 12 minutes 21 seconds. In this time, I provided three case studies and used each one to specifically target one of the CTA Knowledge Audit probes. The first case study was a short sketch by the NC State Physics Education Research Group for Matter and Interactions Laboratories (https://www.youtube.com/watch?v=vgF_lmPqbOA) showing a group working on a lab report. A second case study was a short sketch on effective team meetings, while the third case study was a sketch on ineffective team meetings. These last two sketches were created for the Department of Engineering Fundamentals by “The Troupe”, an improvisational comedy group at Michigan Technological University.

While the case study was presented in the video, the students completed their handout on that case study (described in the next section). The video was then paused for two minutes (as directed in the video) for the students to discuss the case with their teams. After the two minutes were finished, the instructor continued the video which then modeled some of the behaviors that helped or hindered the team in the case study and how to describe them using the Reflective Manipulation in a peer evaluation. This was repeated for the next two case studies as well.

The Procedural Manipulation used the same three case studies. As the students watched the video, they completed their handout (as described in the next section). After each case study, the instructors paused the video for 2 minutes (as directed) to give the students time to complete the information on the case study and discuss it with their teams. After the two minutes were finished, the instructor continued the video which highlighted some of the behaviors that helped or hindered the team in the case study and how to describe them in a peer evaluation using the Procedural Manipulation of giving feedback. This section of the video, containing the procedural manipulation, was 13 minutes and 21 seconds, one minute longer than the corresponding section from the reflection manipulation video.

The last section of both videos covered how to respond to feedback following the tips outlined by Shari Harley in Say Anything to Anyone (Harley, 2013). In the video, students are instructed to follow the S.A.R.A model of receiving feedback and give themselves time to process the feedback they received from their peers to work through these initial stages of emotion independently (Harley, 2013). They are instructed to think

about what was said, ask others for clarification, and then when they are not upset to return to the team with questions (Harley, 2013). Although the students would not be doing this phase of peer evaluation immediately, I thought it was important to include this in the video. This section of the video was 2 minutes and 30 seconds.

The complete videos for both manipulations are linked below.

- Video Link for Reflective Manipulation:
<https://huskycast.hosted.panopto.com/Panopto/Pages/Viewer.aspx?id=77784064-ff8a-4db7-9c6f-ab6d0111be69>
- Video Link for Procedural Manipulation:
<https://huskycast.hosted.panopto.com/Panopto/Pages/Viewer.aspx?id=d15b2ab3-e9a6-4de6-b4ae-ab6d01152065>

5.1.2.3.2 Handouts

For each manipulation, I created a handout that each student could use during the instructional manipulation in the Studio Session. For the Reflective Manipulation, the handout (shown in Figure 5.3) lists the probing questions along with questions to deepen their descriptions of the case studies in the video. Students completed this in conjunction with the video. While watching the first case study situation, the students are directed to use the first probing question in the handout and asked to describe a) what happened, b) who was present, c) what behaviors were present that helped the team, and d) what behaviors that were present that hindered the team.

<p>Name: _____ Team: _____ Section: _____</p> <p>Question 1: Can you describe an instance when you knew something was amiss with your team?</p> <ul style="list-style-type: none"> • What happened? • Who was present? • Behaviors that help the team? • Behaviors that hinder the team? <p>Question 2: Can you think of an example when you were working with your team when you noticed an opportunity to do something better?</p> <ul style="list-style-type: none"> • What happened? • Who was present? • Behaviors that help the team? • Behaviors that hinder the team? <p>Question 3: Can you think of a time when you realized that you would need to change the way you (or another team member) were performing to get the job done?</p> <ul style="list-style-type: none"> • What happened? • Who was present?

Figure 5.3. Giving Effective Feedback: Reflecting on your Team's Experience Handout

Once the students identified the details of the case, it was hypothesized that they would be better able to provide information on the specific behaviors that helped or hindered the team in the CATME survey later. After each case study, the instructors paused the video (as directed) for 2 minutes to give the students time to complete the information on the case study and discuss it within their teams. This process was repeated with the second and third case study and second and third probing question respectively.

The handout for the Procedural Manipulation (Figure 5.4) shows the steps in the process to deepen their descriptions of the case studies in the video. Students in this manipulation completed this handout while watching the video. After the case study, the instructors paused the video (as directed) for 2 minutes to give the students time to complete the information on the case study and discuss it within their teams. This process was repeated with the second case study (Scenario 2) and third case study (Scenario 3).

Name: _____	Team: _____	Section: _____
Scenario 1:		
<ol style="list-style-type: none">1. Discuss the attributes and behaviors of the team to paint a picture of the situation. Be clear and specific.2. Discuss the impact of this behavior3. Offer an alternative		
Scenario 2:		
<ol style="list-style-type: none">1. Discuss the attributes and behaviors of the team to paint a picture of the situation. Be clear and specific.2. Discuss the impact of this behavior3. Offer an alternative		
Scenario 3:		
<ol style="list-style-type: none">1. Discuss the attributes and behaviors of the team to paint a picture of the situation. Be clear and specific.2. Discuss the impact of this behavior3. Offer an alternative		

Figure 5.4. Procedural Method: Steps for Giving Effective Feedback Handout

5.1.2.4 Data Coding

A major component of analyzing this study's results was the qualitative content analysis of the peer evaluation comments (Hsieh & Shannon, 2005). To effectively analyze these comments, I focused on the development of a codebook to identify recurring ideas and themes brought forward by the students in their self and peer evaluation comments. I started the coding process using an inductive, open coding approach to develop the codes used in the study. With the help of another Engineering Fundamentals Departmental Faculty member, we independently read through a sample of 266 ENG1102 comments, generated our codes, and compared our results. This process was repeated a second time with a new sample of 100 ENG1102 comments, where we compared our codes, added additional codes, and corrected descriptive text for clarity. Table 5.1 shows these initial codes and revised codes (bolded).

Table 5.1. Initial Codes from Spring 2020 sample

Code	Description
AH	Asking for help
AT	Attendance
COL	Collaboration
COM	Communication
CR	Creativity
CT	Completes Tasks
D	Dividing Tasks
EI	Encouraging Improvement
H	Helping others
IV	Investment in tasks
L	Leadership
OR	Personal Organization Skills
P	Personality
Q	Quality Work
S	Technical Skills
TI	Timeliness

The next step in open coding after generating the initial codes was to complete secondary, focused coding to develop themes and categories (Saldaña, 2016). This step moves the coding from a conventional content analysis into a more directed content analysis (Hsieh & Shannon, 2005). Reviewing the codes and the CATME behaviors, it became clear that the behaviors the students were commenting on fit within the five dimensions of teaming defined by CATME (Contributing to the Team’s Work, Interacting with Teammates, Keeping the Team on Track, Expecting Quality, and Having Related Knowledge and Skills) as shown in Table 5.2 below. It was noted that students generated the most codes around Contributing to the Team’s Work, followed by Interacting with Teammates and Having Related Knowledge, Skills, and Abilities. Very few codes were identified for the other two themes.

Table 5.2. CATME Themes and Initial Codes

CATME Dimension, Code	Behavior	Code
Contributing to the Team’s Work, C	Asking for help	AH
	Attendance	AT
	Completes Tasks	CT
	Dividing Tasks	D
	Helping others	H
	Investment in tasks	IV
	Timeliness	TI
	Quality Work	Q
Interacting with Teammates, IT	Collaboration	COL
	Communication	COM
	Personality	P
Keeping the Team on Task, K	Leadership	L
Expecting Quality, EQ	Encouraging Improvement	EI
Having Related Knowledge, Skills, and Abilities, KSA	Creativity	CR
	Personal Organization Skills	OR
	Technical Skills	S

5.1.3 Design and Procedure

An experimental manipulation of the type of feedback instruction was utilized for this study. I used a 2 (instruction type) x 2 (pre-instruction, post-instruction) mixed design with instruction type (Reflective vs. Procedural) a between subjects' variable and pre-post instruction a within subjects' variable (2 levels: pre-training and immediate post training). My dependent variables included self-report measures of team satisfaction and team interdependence (assessed using the CATME Team Satisfaction and Team Interdependence Scales respectively), trust and conflict (assessed using a Teaming Reflection Survey), and feedback measures (CATME peer score and CATME qualitative results).

In general, my expectation was that student responses on the dependent variables regarding teaming (conflict, trust, interdependence, and satisfaction) would change over time. Change could be due to the effects of the instructional manipulation or time spent with the team over the semester. For example, students become more comfortable with their teammate and may be willing to give more honest feedback. However, I do expect that the changes in these dependent variables will be larger on those items in which improvement was found in the Reflective Manipulation. Therefore, recognizing the potential confounding variable of time on team, my hypothesis was that the Reflective Manipulation will show larger increases in the dependent measures compared to the Procedural Manipulation.

5.1.3.1 Data Collection Timeline

Figure 5.5 below shows the timeline of this study. At the beginning of the semester (weeks 2-3), ENG1102 instructors created project teams for the semester and these teams began working on their design project for the semester. At the end of week 4, these teams completed at least one major deliverable for their project. At this time, the consent forms were distributed.

In the first Studio Session of week 5, students reflected on their initial team experiences by completing the trust and conflict pre-survey. In the second Studio Session during week five, students completed the early term peer evaluation using CATME, which included the team satisfaction and team interdependence subscales. During week 6, the anonymous results of their peer feedback were released to the students and they discussed the results with their teams in class. During Week 7, the two different feedback instruction manipulations were provided to the students in their studio session. The CATME mid-term peer evaluations with team satisfaction and interdependence subscales immediately followed this instruction in Week 8, with the final reflection survey on trust and conflict occurring week 9.

Figure 5.5. Spring 2020 Study Timeline

5.2 Results

I begin this section with a discussion of the differences on the CATME quantitative measures and the conflict and trust measures between the instructional manipulations. Additionally, I provided an overview of the qualitative results in section 5.2.3.

5.2.1 CATME Quantitative Measures

My hypothesis for this study was that the Reflective Manipulation would show larger increases in the dependent measures compared to the Procedural Manipulation. To begin, I used a one-way ANOVA to verify there were no differences between the initial (week

5) values of the dependent variables between the ENG1102 sections due to time of day (a possible confounding variable). As expected, no statistically significant differences were found as shown in Table 5.3 below.

Table 5.3. Pre-Intervention Results for Dependent Variables by Time of Day (Week 5)

Dependent Variable	F	df	p
CATME Score	0.252	(5, 464)	0.939
Team Interdependence	1.291	(5, 296)	0.268
Team Satisfaction	0.354	(5, 296)	0.879
Trust	1.818	(5, 456)	0.108
Conflict	1.499	(5, 456)	0.189

*p<0.05, **p<0.001

I then used an independent sample t-test to compare the post intervention (week 8) results on the dependent variables between the Reflective and Procedural Manipulations. As shown in Table 5.4 below, no statistically significant differences were detected between the two interventions. It is important to note that maximum possible value on the CATME scale is 1.05, the team satisfaction scale is 15, the team interdependence scale is 25, while both the trust and conflict scores have maximum values of 40. For many of these scales, the average score is close to the maximum, which could indicate a ceiling effect on these measures.

Table 5.4. Post Intervention results between Reflective and Procedural Manipulations (Week 8)

Dependent Variable	<u>Reflective</u>	<u>Procedural</u>	t (df)	p
	M (SD)	M (SD)		
CATME Score	1.00 (0.07)	1.00 (0.06)	-0.370 (468)	0.712
Team Interdependence	16.66 (2.93)	16.67 (3.01)	-0.060 (418)	0.952
Team Satisfaction	12.98 (2.22)	12.94 (2.07)	0.197 (418)	0.844
Trust	33.65 (4.03)	34.01 (4.35)	-0.862 (416)	0.389
Conflict	32.68 (4.37)	33.33 (4.39)	-1.517 (416)	0.130

*p<0.05, **p<0.001

I also used an independent sample t-test to compare the changes in dependent variables (post – pre) based upon intervention type. Again, no statistically significant differences were detected as shown in Table 5.5. For many of the scales, the differences were very small. For example, the team satisfaction, interdependence, trust, and conflict scales differed by less than one point. This would indicate that for any question on the scales, this represented less of a change than one question moving one rating (e.g., strongly agree to agree or vis versa).

Table 5.5. Changes in pre-post results between Reflective and Procedural manipulations (Week 8 – Week 5)

Dependent Variable	Reflective	Procedural	t (df)	p
	M (SD)	M (SD)		
Δ CATME Score	0.0025 (0.052)	-0.0002 (0.047)	0.578 (450)	0.564
Δ Team Interdependence	0.39 (2.47)	0.02 (2.21)	1.32 (282)	0.189
Δ Team Satisfaction	-0.10 (1.51)	-0.26 (1.66)	0.859 (282)	0.391
Δ Trust	0.68 (3.88)	0.73 (3.93)	-0.146 (411)	0.884
Δ Conflict	0.75 (4.29)	0.58 (4.25)	0.395 (411)	0.693

*p<0.05, **p<0.001

While there were no statistically significant differences between the manipulations on the dependent variables in this study, I investigated how students were using these quantitative measures to rate their teaming experience. I conducted a Pearson r correlation analysis to quantify the relationship, if any, that existed between the quantitative teaming measures as assessed prior to the manipulations (Table 5.6). I used the week 5 (pre-test) values for this test. As expected, the teaming dysfunction scales (trust and conflict) showed strong significant positive correlations with each other. Additionally, the CATME team satisfaction and team interdependence scales showed weak significant correlations with each other.

It is interesting to note that both the trust and conflict measures show strong significant positive correlations with the CATME team satisfaction measure. This

suggests that teams who have established trust and have productive conflict (higher scores on these scales) will demonstrate higher team satisfaction, which supports the use of the dysfunction scales as a measure of team success. These scales also show weak significant positive correlations with the team interdependence scale.

Table 5.6. Correlations and p-values between Teaming Measures (week 5)⁺

Measures	1	2	3	4	5
1. Trust	--	.781** (.000)	-.070 (.140)	.604** (.000)	.178** (.002)
2. Conflict	.781** (.000)	--	-.035 (.464)	.556** (.000)	.160** (.006)
3. CATME Score	-.070 (.140)	-.035 (.464)	--	-.066 (.253)	-.003 (.963)
4. Team Satisfaction	.604** (.000)	.556** (.000)	-.066 (.253)	--	.217** (.000)
5. Team Interdependence	.178** (.002)	.160** (.006)	-.003 (.963)	.217** (.000)	--

⁺ table notation Pearson r (p value)

* p < .05 ** p<0.005

What is disconcerting is the fact that the CATME peer rating does not significantly correlate with any of the other measures of teaming. This is concerning as this measure compares an individual's rating across the dimensions with the average rating of their teammates to determine how a student was rated in comparison with their teammates. According to Braender and Napels (2013), when all team members pull together or rank each other with the same contributions, this score is 1.00 (Braender and Napels, 2013). The lack of significance could indicate, as shown in Study 2, that students

were once again not differentiating between team members on the CATME behavioral codes, and that this measure of teaming is not a valid representation of team functionality.

The data for the Spring 2020 CATME ratings can be found in Figure 5.6 below. Approximately 12.4% ($n = 56$) of students were ranked lower than their teammates and approximately 48.5% ($n = 219$) of students were ranked higher than their teammates. This means that the remaining 39.1% ($n = 177$) of the students were ranked the same as their teammates. This also implies that over 39% of students are not using the CATME peer rating to provide differentiation in their ratings of each other, which again may mean that the CATME peer evaluation rating does not measure what it is intended to measure or that students don't understand or can't differentiate between the ratings.

Figure 5.6. CATME Peer Evaluation Score Results Spring 2020 ($n = 452$, $m=0.9996$, $SD = 0.06427$)

To further analyze these results, I explored the initial team interdependence and team satisfaction scores and the changes in these scores over the study time period. To begin, the initial team interdependence scores (shown in Figure 5.7) demonstrate a normal distribution, which is what I expected to see with this scale. The majority of teams would be at the center of the scale (approximately 15 out of 25), with few teams who might feel strongly toward the extremes (strongly agree or disagree).

Figure 5.7. CATME Team Interdependence Score Week 5 Results Spring 2020 ($n = 302$, $M = 16.38$, $SD = 2.96$)

However, when looking at the change of student responses over the testing period (Figure 5.8), this data also is represented by a normal distribution about 0.222. This indicates that few students are changing their responses to these questions in this time frame. Twenty-one percent ($n = 62$) of students did not change their answers at all. A total of 88 students (30.9%) changed their responses by an equivalent of answering one of the five questions with one degree more or less agreement (i.e., moving from neutral to

agree or disagree). Overall, only 40 students (14.1%) changed more than four points on the scale in either direction (the equivalent of switching from strongly agree to strongly disagree or vis versa on one question).

Figure 5.8. Change in Team Interdependence Value in Spring 2020 ($n = 284$, $m=0.222$, $SD = 2.36$)

The team satisfaction score results showed different characteristics than the team interdependence score. Figure 5.9 shows the initial (week 5) results, while Figure 5.10 shows the change in these values after the interventions. Looking at Figure 5.9, 82.7% ($n=250$) of the students either agreed or strongly agreed with all three team satisfaction questions at the onset of this study. This introduces a ceiling effect into the interpretation of these results, which could help explain the lack of difference between intervention groups over the course of the study. Looking at Figure 5.10, 51% ($n = 145$) of students

did not change their answers from their initial results with only 5 students changing 4 points or more in either direction.

Figure 5.9. CATME Team Satisfaction Score Week 5 Results Spring 2020 ($n = 302$, $M = 13.0$, $SD = 2.09$)

Figure 5.10. Change in Team Satisfaction Values in Spring 2020 ($n = 284$, $M = -.173$, $SD = 1.58$)

5.2.2 Dysfunction Measures

Like the CATME data analysis, running an independent sample t-test on the trust and conflict post intervention measures yielded no statistically significant results. Looking at the data for the trust measure (Figures 5.11 and 5.12), the same ceiling effect was observed with the team satisfaction measure. Sixty-six percent of the students ($n = 306$) rated agree or strongly agree on all eight of the trust survey questions with 52 (11.3%) of students strongly agreeing with all questions. Seventeen percent of the students ($n = 72$) did not change their answers on the post-intervention survey. There were 21.5% of students ($n = 89$) who changed their responses by more than 4 points or more in either direction, which includes 19 students (4.6%) who changed their responses by more than 8 points in either direction.

Figure 5.11. Trust Scale Score Week 5 Results Spring 2020 ($n = 462$, $M = 33.02$, $SD = 4.65$)

Figure 5.12. Change in Trust Value in Spring 2020 ($n = 413$, $M = 0.70$, $SD = 3.90$)

The conflict scale shows similar results (see Figures 5.13 and 5.14), although fewer students strongly agreed with all conflict scale scores (8.9% versus 11.3% on the trust survey). Additionally, only 59% of students ($n = 273$) rated agree or strongly agree on all eight of the conflict survey questions versus 66% on the trust scale. The change in conflict scores (although not significant from the pre-conflict score) demonstrated more variability than changes in the trust scores in the test period. Only 15% ($n = 62$) did not change their responses (versus 15% ($n = 72$) for trust). Additionally, 25.7% of students ($n = 106$) changed their responses by more than 4 points or more in either direction (17 more students than on the trust scale), including a total of 35 students (8.5%) who changed their responses by more than 8 points in either direction.

Figure 5.13. Conflict Scale Score Week 5 Results Spring 2020 ($n = 462$, $M = 32.27$, $SD = 4.74$)

Figure 5.14. Change in Conflict Values in Spring 2020 ($n = 413$, $M = 0.67$, $SD = 4.27$)

5.2.3 Peer Review Comments

Based on the minimal changes in the quantitative results and the low feedback responses from the fall 2019 studies, I focused the analysis of the peer review comments to initial development of the codebook (discussed in section 5.1.2.4) and a quantification of these codes with the Spring 2020 peer evaluation data. Further development of the codebook and refining of the codes occurred during the final study and will be discussed in Chapter 6.

In general, looking at one instructor's sections a greater percentage of students did provide comments on the end of term evaluations (77.6% for the Reflective Manipulation and 67.3% for the Procedural Manipulation). However, these comments lacked specific feedback related to the behaviors of the team members. For example, with the exception of the Contributing to Team's Work dimension in which students generated eight unique behaviors, students did not generate many unique behaviors for each of the dimensions of teaming (see Table 5.2). Additionally, the students did not provide comments across the dimensions (e.g., comment on the Contributing to Team's Work and other dimensions).

Figure 5.15 shows the cumulative percentage of students against increasing length of comment (measured by number of characters). In a random sample of 113 student codes from ENG1102, 50% of the students completed their peer feedback comments with 80 characters or less (approximately 15 words) and averaged 98 characters per response (approximately 18 words). Only 9% of students ($n = 10$) created feedback with more than 200 characters (approximately 51 words or more).

Figure 5.15. Length of Feedback Comments in ENG1102 Spring 2020 Sample

Overall, 16% of students ($n = 18$) did not comment on any dimensions of teaming, 40% ($n = 45$) mentioned one dimension, 36% ($n = 41$) commented on two dimensions, and 8% ($n = 9$) mentioned behaviors in three dimensions of teaming. No one commented on more than three dimensions. Figure 5.16 shows the distribution of these comments. In general, very few students commented on Keeping the Team on Track, K (9%, $n = 10$) or Expecting Quality, EQ (2%, $n = 2$). Additionally, less than half of the students commented on Interacting with Teammates (IT) or Having Related Knowledge and Skills (KSA).

Figure 5.16. CATME Behavioral Categories in ENG1102 Spring 2020 Sample

In general, there is a strong positive correlation between the number of dimensions of teaming the students identified in their comments and length of feedback comments measured by number of characters ($r = 0.518$, $p < 0.001$) and measured by number of words ($r = 0.493$, $p < 0.001$). This indicates that if there is a way to increase the amount of feedback the students provide; this should correspond to an increase in behaviors and categories the students identify.

5.3 Discussion

My goal for this project was to test two different instructional manipulations (Reflective and Procedural) to see if they had an impact on feedback towards other team members, team satisfaction, team interdependence, trust, and productive conflict in the first-year engineering classroom. For this study I hypothesized that the Reflective Manipulation

would show larger increases in the dependent measures compared to the Procedural Manipulation. Looking at the dependent measures, there were no significant changes in any of the dependent measures over the course of the semester for either instructional manipulation.

The team conflict, trust, satisfaction, and CATME ratings all exhibited a ceiling effect with a large number of students providing the highest rating at the week 5 measurement on the dependent measures scales and did not significantly change their ratings at the week 8 measurement. The team interdependence scale did not exhibit this ceiling effect, but there were still no significant differences between the week 5 and week 8 measure for either instructional manipulation.

The CATME adjusted factor did not significantly correlate with any of these teaming measures. Approximately 40% of the students were ranked the same as their teammates, which implies that the students are not using the CATME peer ratings to provide differentiation on how each team member acts with respect to the teaming behaviors. It also may indicate that students don't understand or can't differentiate between the ratings.

From Study 2, it was found that only 55% of students provided feedback (of any kind) in the comment section of the CATME survey and only 6% of students provided negative or constructive feedback to their teammates. In this study, while it was determined that a greater percentage of students provided comments in the peer evaluations, students did not comment on many specific behaviors related to the CATME

dimensions of teaming nor did they provide much detail in their comments at all (measured by length of response). There was a significant positive correlation between student comment length and the number of teaming dimensions addressed.

At this point, there was no evidence to support my hypothesis and I took a step back to rethink my approach. After critically analyzing these first three studies, I identified some limitations in the study design, measures, and the instructional manipulations to address. First, the study design did not have a clear control group, which would have controlled for the extraneous variable of time spent in group. In this study, I used two different instructional manipulations and compared these results to each other. However, this design did not tell if there was an effect by either manipulation, only if there were different effects between the manipulations.

Secondly, it appears that the dependent variables I chose were not effective measures to assess changes in rankings over time or differentiation between student rankings of each other. There was a strong ceiling effect exhibited with almost all of the dependent measures (team satisfaction, trust, conflict). Additionally, 39.1% ($n = 177$) of the ENG1102 students were ranked the same as their teammates on the CATME measure, which implies that students may not be using the quantitative CATME peer rating to differentiate between the contributions of different team members.

Third, the length and detail of comments in the peer feedback section may have been inhibited by the design of the CATME survey itself. Examining different survey instruments, increasing the visual size of the response box has been shown in several

studies to increase the length of response (Christian and Dillman, 2004; Smyth, Dillman, Christian, & McBride, 2009), while increased clarity of instructions has been shown to improve the thoroughness of responses (Smyth et al., 2009; Chaudhary & Israel, 2016). Based on the correlations between student feedback length and behavioral categories identified, increasing the amount of feedback may correspond to an increase in behaviors and categories the students identified.

Lastly, the use of the CTA Knowledge Audit prompts would be more effective if the students focused on a specific teaming event. In ENG1102, the CTA prompt instruction was provided with case studies that did not represent their semester work. While the students completed the peer evaluations after different aspects of their semester team project, they were not asked to reflect on the project when completing the evaluations. Having the students focus on a specific instance, these questions may encourage more specific examples about the teaming experience (Crandall, Klein, & Hoffman, 2006), and reduce memory errors linked with incorrect or biased responses (Cohen, Manion & Morrison, 2013). This change could result in added detail in the peer evaluation comments, which would indicate progress toward overcoming the hesitation to provide constructive feedback.

6 Study 4

My goals for this final study, which I completed in Fall 2020, were to address the main limitations found in the Spring 2020 study. First, I focused on having a true control group which eliminated the confound for time spent in project group. In addition, methodological changes were made to eliminate the potential confounds of time in group and delivery mode. The assignment of participants to experimental groups were balanced across instructor and class timing. In addition, the sections chosen for this study were all taught using the remote only instruction method to eliminate the potential confound of delivery mode that may have been present in the fall due to COVID-19 instructional restrictions.

Next, I modified the team project to enhance the teaming interactions and responsibilities of individual members for the overall team product. I also modified the reflective CTA activity to focus the students on evaluating their teaming experience rather than evaluating case studies, as the Reflective Manipulation showed the largest percentage of students commenting in the course peer evaluations. Finally, I modified the prompts and available space for feedback in the CATME survey to address the issues inherent in the survey instrument itself. I will summarize the specific methods, research design, hypotheses, and results of this study in the following sections.

6.1 Methods

6.1.1 Participants

This study was conducted within two classes of ENG1101: Engineering Analysis and Problem Solving during the Fall 2020 semester. These sections contained a total of 165 students in 45 teams (35% Female, 66% Male) and were taught by two different instructors teaching remotely due to the COVID-19 pandemic. All sections for these instructors were offered synchronously using Zoom (a video and audio-conferencing system).

During class time, students would sign into the class Zoom meeting, the instructor made any class announcements, verified understanding of the ESP materials (preparatory materials for class), and introduced the class activities for the day. Collaborative work time on these activities was facilitated through the use of Zoom breakout rooms (Hamlin et al., 2021). LEAP Leaders and the instructor visited the rooms to verify understanding and progress on the class materials. At the end of class, the instructor brought the class together to review and share the results of the collaborative activities as well as discuss upcoming assignments and due dates (Hamlin et al., 2021).

Seventy-nine percent of the students ($n = 131$) consented to participate in the survey. Teams were randomly assigned within instructor to one of the three groups (control, enhanced prompt group (EP), and experimental (CTA)) as shown in Table 6.1 below. Randomization occurred separately for each instructor's courses such that each instructor had approximately an equal number of teams in each category.

Table 6.1. Sample Distribution

	<u>Control</u>	<u>Enhanced Prompt (EP)</u>	<u>Experimental (CTA)</u>
Number of Students	45	39	47
Number of Teams	15	15	15
Gender Distribution (M / F)	(45% / 55%)	(24% / 76%)	(21% / 79%)

6.1.2 Materials

In this section I detail the manipulations to the survey instrument and instructional methods used in this study. Additionally, I discuss the instructional materials created for use in this study.

6.1.2.1 Survey Instrument Manipulation

For this study, CATME was used to collect demographic data and form teams using the TeamMaker functionality at the beginning of the semester. Later, I used the CATME peer evaluation survey to collect quantitative ratings for individual contributions on the five dimensions of teaming and anonymous qualitative feedback for each member of the student teams. This survey was recreated and administered through Survey Monkey so fields could be adapted and tested using the modification variable discussed above (e.g., more space to provide written responses and changing the prompts). Details of these adaptations are discussed in more detail below.

In addition, I added a scale question on team functionality to the peer evaluation survey for all groups: “At this current time, rate how satisfied you are with the functionality of your team experience”. This question was rated on a scale of 0 (*not*

satisfied) to 100 (*completely satisfied*). By providing a larger scale for functionality, I hoped that students would submit a broader range of responses and avoid the ceiling effect observed with the CATME Team Satisfaction scale. This survey was administered after the first teaming project (week 8 of the semester) and again at the end of the semester (week 14).

Within the CATME survey, students are provided with a list of observable behaviors that they rate themselves and their teammates (see Appendix B) for each of the five dimensions of teaming (**C**ontributing to Team's Work, **I**nteracting with Teammates, **K**eeping the Team on Track, **E**xpecting Quality, and **H**aving Related Knowledge and Skills). Following these rankings, students are then directed to a screen where they can provide comments on their own performance and comments to their teammates followed by a screen where they can provide confidential comments to the instructor, rate their team's functionality, and consent to the survey. This survey was recreated in Survey Monkey (see Appendix E) for this study, with the control group receiving the CATME survey without the modifications discussed below.

After recreating the CATME survey in Survey Monkey, it was manipulated in two main ways in an effort to increase the amount, and hopefully the utility, of the peer feedback. *Utility* was operationally defined as detailed, including references to specific team work, and providing actionable and/or impactful feedback. The first part of the manipulation was to increase the size of the comment boxes for the peer feedback. Although the CATME comment boxes can hold up to 3000 characters (www.catme.org), the visual size of these boxes is much smaller. A study by Smyth, Dillman, Christian and

McBride (2009) indicated that in a web-based survey (such as CATME or Survey Monkey) a larger response box can result in longer responses for less motivated individuals. For this manipulation of the CATME survey, prompt boxes were provided that were three times the height of the boxes in the original CATME survey.

The second part of the CATME tool manipulation was to increase the clarity of the instructions of the peer feedback prompts. Studies have indicated that increasing the question length (Kalton and Schuman, 1982) and providing clarifying instructions can improve the diligence in completing open-ended questions (Smyth et al., 2009; Chaudhary & Israel, 2016). To increase the clarity of the instructions, additional details regarding what they should provide in their comments about themselves and their teammates was provided. The original CATME prompt, shown in Figure 6.1, was used for the Control group, while the enhanced prompt, shown in Figure 6.2, was used for both the EP and the CTA groups. After each prompt, the students were provided with the response box to comment in for each of their teammates as described above.

Please provide constructive comments about your fellow teammates as well as yourself. The purpose of these comments is to give you the opportunity to explain how you rated your peers and if there was behavior or experiences in particular that influenced you when doing your peer and self-evaluations.

Figure 6.1. Original CATME Peer Comment Prompt

The purpose of these comments is to give you the opportunity to explain how you rated your peers and if there was behavior or experiences in particular that influenced you when doing your peer and self-evaluations. Please provide detailed, actionable, and constructive comments about your fellow teammates as well as yourself.

Detailed and actionable feedback include:

- focusing on specific behaviors, not personality (can be behaviors you would like to see more of or one's you would like to see stopped)
- providing examples of these specific behaviors,
- specifying how these behaviors affect your team or your team's work, and
- suggesting actions or behaviors that should be added, stopped, altered, or maintained as is.

Be sure to include comments on the five CATME behavioral categories: Contributing to Team's Work, Interacting with Teammates, Expecting Quality, Keeping the Team on Track, and Having Related Knowledge, Skills, and Abilities.

Figure 6.2. Enhanced CATME Peer Comment Prompt

6.1.2.2 *Peer Evaluation Introduction Video*

All groups received an instructional video on how to complete the peer evaluation survey.

While this video was created by me, it was based upon two help videos created by CATME (listed below). These CATME videos explain to the students what the CATME student interface looks like, what steps they need to take to complete each portion of the peer evaluation survey (ranking behaviors and giving feedback comments), and how they can submit and view the results.

1. “How to Take a Peer Evaluation Survey for Students” (3:15)

(https://www.youtube.com/watch?v=_vapni7K-mc&t=87s)

2. “Peer to Peer Comments Student Tutorial” (2:05)

<https://www.youtube.com/watch?v=TBWQried5o>

I downloaded and edited the transcripts of these videos to include interface information relevant to completion of the instrument in Survey Monkey. This video explains to the student what the Survey Monkey peer evaluation survey looks like, what steps they need to take to complete each portion of the peer evaluation survey (ranking behaviors and giving feedback comments), how to consent to the research, and how to submit their results. The instructional video transcript is shown in Appendix F, while the final video can be viewed at

<https://huskycast.hosted.panopto.com/Panopto/Pages/Viewer.aspx?id=69c3969a-63df-4fe7-86f0-ac4c00209f52>.

6.1.2.3 Teaming Project

I modified the team project for ENG1101 to enhance the teaming interactions and responsibilities of individual members for the overall team product. All ENG1101 students completed an integrated team project on the heat treatment of steel that required the team to work together in an interdependent manner. For this project, there were activities that required each student to complete background research or assignments to familiarize themselves with the material. Additionally, there were activities that the team completed together that utilized this background information. A complete list of activities for this project is listed below.

1. Day 1 Team Assignment: Steel Properties & Treatment Video

For this assignment, the teams created a 2-3 minute video where the teams defined a specific vocabulary term related to steel processing (tempering, quenching, hardness, impact strength, or edge holding) and why that term was important to knife making. The teams also explained what happens to steel in the tempering or quenching process or when the property they were given increases or decreases and why this information is important to engineers. As part of the assignment requirements, each person on the team needed to answer at least one of the questions.

2. Day 2 Individual Assignment: Steel Material Properties Spreadsheet

For this assignment, each student on a team researched a specific type of steel (440C, CPM154, CPM S35VN, CPM S30V) and created a spreadsheet that contained the material properties of the steel and the heat treatment response data for the steel.

3. Day 2 Individual Assignment: Review Vocabulary Video

Each student reviewed and took notes on all the vocabulary videos created by the teams in their section.

4. Day 3 Individual Assignment: Steel Properties and Processing Quiz

Each student completed a quiz that checked their understanding of the vocabulary information for this project.

5. Day 3: Individual Assignment: Plotting Steel Composition and Treatment

For this assignment, each student on a team plotted the material properties and heat treatment response data for the specific type of steel they researched.

6. Day 4 – 5 Individual Assignments: Curve Fitting (concept learning)

Each student learned the concepts necessary for the curve fitting calculations and modeling necessary for the final heat treatment analysis in this project.

7. Day 5 - 6 Team Assignment: Analyze Heat Treating Data from a Furnace

As a team, the students analyzed data from a heat treatment oven using MATLAB. Each team used the information they learned about curve fitting to generate the heating and cooling curves and the equations for these curves for the example heat treatment process. An example of the final plot (with the equations hidden) can be seen in Figure 6.3 below. After generating this plot, the team wrote a report that documented the team's analysis and provided recommendations for each steel that the team members researched in this project. For example, the team member who researched 154 CM steel would explain what properties that steel had if it had gone through the heat treatment process shown in the figure (e.g., the hardness increased to between 59-62 and a wear resistance value should be 49 mg). Additionally, the students would indicate what applications they would recommend for 154CM steel with those properties (e.g., cutlery, bearings, valve ports, or bushing).

Figure 6.3. Example Heat Treatment Analysis Plot

6.1.2.4 Task Manipulation

The last manipulation for this study was the manipulation of task. I modified the reflective CTA activity that the CTA group completed to focus the students on evaluating their teaming experience rather than evaluating generic case studies. To ensure that the timing (length of time between instruction video and when the peer evaluation was taken) was similar between the control, EP, and CTA groups, each group had an assignment to complete between the initial instructional video and completion of the peer evaluation form. These assignments were individually graded (not group graded) assignments. The

control and EP groups completed a study skills reflection assignment, while the CTA group completed the CTA teaming reflection assignment as discussed in more detail below.

6.1.2.4.1 Study Skill Reflection

The study skills reflection assignment, provided to the control and EP groups, took place in two parts. The first part was a review of study skill strategies using a premade Kahoot (an online game-based learning quiz). This Kahoot was set up as a challenge so students in ENG1101 could complete the review independently of the instructor. When prompted, students completed the study skills Kahoot without having the instructor facilitate the experience. Next, the students completed a study skills reflection. In this assignment, shown in Appendix G, students were asked to reflect on the types of study skills that could help them be successful in their upcoming midterm exams.

Students first reflected on past exams by listing which exams they had for which courses and how they studied for them. Then the students wrote a one to two sentence reflection to discuss the changes (if any) they should make to their study strategies for these courses in the future. Next the students looked at their future midterm exams, what study strategies they planned to use, and wrote a short reflection on their planned strategies. One LEAP Leader completed these activities to test them and reported that the Kahoot activity took 5 minutes, while the study skills reflection took approximately 20 minutes for a total of 25 minutes for these activities.

6.1.2.4.2 CTA Teaming Reflection

For the Fall 2020 CTA instructional manipulation (shown in Appendix G), delivered to the CTA group, I provided the students with a list of the main tasks for the heat treatment project. Students in the CTA group were instructed to reflect on their experience with their team in the completion of this project. They completed a table by defining what steps were needed to complete each task (e.g., team meeting, individual work) and who participated in each task. They were then directed to use the information from the table to answer the CTA reflection questions.

1. Can you describe an instance when you knew something was amiss with your team? Be specific and paint a picture with words (e.g., what happened? who was involved?)
2. Can you think of an example when working on this project when you noticed an opportunity for you or your teammates to do something better? Be specific and paint a picture with words (e.g., what happened? who was involved?)
3. Can you think of a time when you realized that you would need to change the way you or another team member were performing to get the job done? Be specific and paint a picture with words (e.g., what happened? who was involved?)

Students were then directed to use this information when they completed the Peer Evaluation form in Survey Monkey. The same LEAP Leader who tested the study skills

reflection activities tested this activity as well. She reported taking 22 minutes to complete this teaming reflection activity.

This teaming reflection was modified from the activity chosen for the Spring 2020 Reflective Manipulation (Study 3) in a couple of ways. First, the students were directed to reflect on their team's interactions (both individually and as a team) on tasks specific to the heat treatment project. In Study 3, students were asked to think back on their teaming interactions over the semester, without a specific instance to focus on. Secondly, the students in the current study were asked to use their experience on the heat treatment project to identify instances to address the CTA questions listed above. This is opposed to using the CTA reflection questions shown above in a more general context with case studies unrelated to their team.

These additional details provided by the CTA activity are important when providing feedback to an individual. If they cannot picture the behavior and the situation clearly, they cannot maintain or change their behavior (Harley, 2013). By having the students focus on a specific incident, these knowledge audit questions can help elicit brief stories and examples about the teaming experience (Crandall, Klein, & Hoffman, 2006), which could translate to added detail in the peer evaluation comments. Additionally, focusing on a specific event may help reduce memory errors linked with incorrect or bias responses (Cohen, Manion & Morrison, 2013).

6.1.2.5 Qualitative Coding

As mentioned previously, one of the main goals of this project was to increase the utility of the peer feedback generated by students with utility operationally defined as detailed, references specific team work, and providing actionable and/or impactful feedback. With the changes to the CATME prompts discussed in section 6.1.2.1 for the EP and CTA groups, the coding system needed to be refined to include codes not only for behaviors, but also for this added utility.

In Chapter 5, I discussed the coding methods used to develop the initial codebook. This book was further refined in Fall 2020 with assistance from the same Engineering Fundamentals Departmental Faculty member who helped code the Spring 2020 samples. Results from the peer feedback comments were coded using primary coding for teamwork behaviors and a secondary subcode for the increased utility of the feedback. This subcoding, or second-order tag, is often assigned to a primary code to add detail (Saldaña, 2016, pp. 91). I'll discuss the primary and secondary coding in more depth in the following sections.

6.1.2.5.1 Primary Coding

Initially, the results from the peer feedback were coded using primary coding for the five dimensions of teaming (Contributing to Team's Work, Interacting with Teammates, Keeping the Team on Track, Expecting Quality, and Having Related Knowledge and Skills) and non-behavioral codes shown in Table 6.2 below. For each CATME teaming behavior, we coded both positive behaviors (comments about things done well) and negative behaviors (comments about things that needed to change) that were included in

peer feedback comments. Additional primary codes were included for comments provided by students that did not specifically refer to behaviors. Several students gave generic positive comments for themselves or their teammates (coded as DA), or negative comments for themselves or their teammates (coded as DN), or gave comments that did not refer to the team or teamwork in any way (coded as NR).

The third column of the table is where we listed a description of these behaviors. These behaviors go beyond those originally defined by CATME. For example, for Contributing to Team's Work, the CATME measure discusses quality of work, contributions to the work, helping teammates, completing team's work, timeliness of work, and attendance at team meetings. Other behaviors that students indicated that I added to this category include preparation for class and team activities, and investment in the team's work. These additional items are highlighted in blue in Table 6.2.

For the most part, the students commented on behaviors that were originally identified in the CATME system. However, in addition to the behaviors in the Contributing to Team's Work category, students brought forth further behaviors for the Keeping the Team on Track and Having Related Knowledge and Skills categories as well. Planning and organizing team members or team activities, submitting assignments, and communicating with the instructor were all positive behaviors mentioned by peers that we placed into the Keeping the Team on Track category. Several students also

Table 6.2. Codebook for Peer Evaluation Analysis

Code	Definition	Description
C(-)	Contributing to Team's Work	Behaviors that demonstrate how the student is contributing to the team's work. These include the amount, timeliness, completion, and quality of work, helping others or asking for help, being prepared for class (e.g., doing the homework and ESP assignments, knowing the requirements for the assignments) and team activities, attendance in class or team activities, investment in team's work or activities.
IT(-)	Interacting with Teammates	Behaviors related to student personality and collaboration strategies. This includes communication between group members, asking for feedback, working with team members, communicating with each other , participation in team activities. Negative interactions - doing teamwork alone as opposed to working with others.
K(-)	Keeping the Team on Track	Behaviors related to leadership. Planning and organizing team members or team activities , keeping the team on track, taking a leadership role, submitting assignments or communicating with the instructor , giving feedback to teammates. Negatives - taking over an assignment and not letting others contribute or participate.
EQ(-)	Expecting Quality	Behaviors related to expecting quality of themselves or others. This includes motivating others to do quality work, encouraging themselves or others to go beyond the rubric requirements, encouraging others to improve.
KSA(-)	Having Related Knowledge, Skills and Abilities	Technical knowledge of course software or materials (e.g., MATLAB, NX, Excel). Personal skills such as organization, communication, creativity (e.g., bringing novel ideas) . Desire to learn. Knowing what to do.
DA	Affirmative tone, not specific to a behavior or actionable	Terse comments: "Good work", "good job", "I'm doing good" "good teammate" (if only comment) - gives no description of what they are doing well or should keep doing, no action items)
DN	Negative tone, not specific to a behavior or actionable	Terse comments. "Bad job" "You suck" (gives no description of what they are doing poorly or discuss specific behavior, impact, or action)
NR	Not relevant	Does not provide comment relative to team or behavior to team (e.g., nice face)

identified a negative behavior of an individual taking over an assignment and not letting others contribute or participate which was added to this category as well. Creativity and personal organization were added as skills to the Having Related Knowledge and Skills category.

I have also highlighted some phrases in the table in pink. These behaviors were very similar to one another and all focused on communicating in one aspect or another. However, the way that students described these behaviors led us to classify them in different categories. First, asking for help is clearly defined by CATME as Contributing to the Team's work, so that was placed there. Communicating with each other was placed into Interacting with Teammates. However, when students commented specifically on their own or their teammate's communication as a skill, it was coded in the Having Related Knowledge and Skills category.

6.1.2.5.2 Secondary Coding

The new secondary codes for utilized for rating the utility of comments are included in Table 6.3 below. A secondary code was created for each aspect of utility (references specific team work (FE), provides impactful feedback (FI), and provides actionable feedback (FA). These second-order tags were added to the primary codes discussed in the previous section. The following examples demonstrate how these codes were implemented.

- Example 1: "You are on topic during team meetings." Code: IT/FE This example would be given a primary code of IT and a secondary code of FE. Staying on

topic is an example of communication behavior (IT) and the team meetings provides a contextual teaming example (FE) where the behavior was taking place.

- Example 2: “It can be irritating when you don't show up to class as well as not paying attention to the things we learn while you are on your phone. It would be appreciated to pay attention the entire period so we can all learn the skills needed to do our projects.” Code: C-/FE/FI/FA. This would be coded with a C- as a behavior and negative as this behavior is negatively contributing to the team. It also provides a specific teaming example (not coming to class, not paying attention) as well as discusses the impact (all learn the skills) and a proposed action (pay attention).

Table 6.3. Secondary Code Definitions

Code	Definition	Description
FE	Example Feedback	Provides specific example of behaviors related to team member, team work, team product, or team process (e.g., helps others is too vague, helps others with understanding MATLAB is specific)
FI	Impact Feedback	Provides specific impacts or effects on team, quality of work or work product, or individual team members. Specifically, this should identify how the behavior helps or hurts any of the above. (e.g., helps others with understanding MATLAB does not define the impact, but helps others with understanding MATLAB to improve their skills does). Impacts can be to the individual such as "this helped me" or "I enjoy working with this team" or they can be to the team "this helped us when we were stuck". NOTE: "kept the team on track" is both an example and impact.
FA	Action Feedback	Provides a comment that explicitly states that something that should change (stop or start doing this) or stay the same.

While Table 6.3 provides a definition on how to implement these codes, it is important to note when implementing the FE code with the Fall 2020 peer evaluation data, multiple behaviors in a specific behavioral category were coded with this secondary code as well by agreement between myself and the other rater. The following example demonstrate how these codes were implemented.

- Example 3: “He always did more than his fair share of work and was very willing to help those who were stuck.” Code: C/FE. This would be coded with a C as a behavior and the secondary code FE as two behaviors (“doing more than fair share of work” and “very willing to help those who were stuck”) were provided.

6.1.2.5.3 Reliability

The midterm peer evaluations were coded by myself and the same EF departmental faculty who helped code the Spring 2020 sample and develop the new coding scheme. We used convergent coding in the development of this codebook. Each rater independently used the primary and secondary codes to rate all 577 midterm peer evaluation comments. After coding the entire sample, we met via Zoom and talked through approximately 22% of the data ($n=131$) to verify we were using the codes consistently and clarify the definitions when we were not. We would then update the codebook with this new information. Once we reached agreement, we would recode the remaining data with this updated codebook. This cycle was repeated with the next set of data (22%, $n = 131$), then the next 22% ($n = 131$), and finally the remaining 32% ($n = 184$).

To measure the inter-rater reliability between the coders, I calculated Cohen's kappa for each of the primary and secondary codes as shown in Table 6.4. Using the guidelines of interpreting Cohen's kappa recommended by Laerd Statistics (2015), shown in Table 6.5 below, all of these values demonstrate very good strength of agreement between the two raters. This is to be expected as we focused on developing this agreement through the repeating cycles of coding and evaluation.

Table 6.4. Inter-rater Reliability Measures on Midterm Peer Evaluation Coding

Codes	Cohen's kappa
Contributing to the Team's Work (C)	0.919
Interacting with Teammates (IT)	0.952
Keeping the Team on Track (K)	0.977
Expecting Quality (EQ)	0.943
Having Related Knowledge, Skills, and Abilities (KSA)	0.974
Non-descriptive, Affirmative (DA)	0.937
Non-descriptive, Negative (DN)*	1.000
Not Relevant (NR)*	1.000
Feedback with Team Example (FE)	0.877
Feedback with Impact (FI)	0.880
Feedback with Action (FA)	0.969

* Only one case was identified for DN and NR codes

Table 6.5. Classification of Cohen's kappa (Laerd, 2015)

Cohen's kappa	Strength of Agreement
<0.20	poor
0.21 – 0.40	fair
0.41 – 0.60	moderate
0.61 – 0.80	good
0.81 – 1.00	very good

To get a more independent measure of inter-rater reliability between our two coders, we divided the end of term peer evaluations between us and both rated 107 comments independently to use for comparison. These results are shown in Table 6.6 below. Although not all the ratings are as strong, all the CATME behavioral categories have a Cohen's kappa with a good strength of agreement (> 0.61) with the exception of Keeping the Team on Track, which has very good strength of agreement (>0.81). The three non-descriptive categories (DA, DN, and NR) all had very good strength of agreement as well, although there were not very many comments with these codes.

In general, the secondary codes for utility (references specific team work (FE), provides impactful feedback (FI), and provides actionable feedback (FA)) had moderate agreement. The lower agreement on these secondary codes may indicate a need for clearer definitions on what is a teaming example, impactful feedback or actionable feedback. It is also possible that there may not be enough examples ($n = 79$), impacts ($n = 44$), or actionable feedback ($n = 7$) in these 107 comments to determine inter-rater reliability.

Table 6.6. Inter-rater Reliability Measures – Independent Coding of Final Peer Evaluations (n = 17 students, 107 comments)

Codes	Cohen's kappa
Contributing to the Team's Work (C)	0.706
Interacting with Teammates (IT)	0.628
Keeping the Team on Track (K)	0.874
Expecting Quality (EQ)	0.795
Having Related Knowledge, Skills, and Abilities (KSA)	0.768
Negative Behaviors (-)	0.778
Non-descriptive, Affirmative (DA)	0.864
Non-descriptive, Negative (DN)*	1.000
Not Relevant (NR)	1.000
Feedback with Team Example (FE)	0.564
Feedback with Impact (FI)	0.557
Feedback with Action (FA)	0.541

* Only one case was identified for DN code

6.1.3 Design and Procedure

The design of this study included experimental manipulations of the type of feedback instruction, CATME survey modifications, and midterm/end-of-term timing of surveys. I used a 3 (instructional/survey manipulation) x 2 (survey timing) mixed design with instruction and survey type (control vs EP vs CTA) a between-subjects variable, and survey timing (midterm and end-of term) a within-subjects variable. My dependent variables included self-reported ratings of the CATME behavioral categories, team

functionality, and feedback comments. The CATME behavioral categories were measured using the CATME evaluation tool as adapted for this study. The feedback comments were measured using the content analysis coding scheme discussed in the previous section, which focuses on behavioral categories addressed in the comments. Finally, team functionality was measured as a self-reported scale from 0-100.

Figure 6.4 shows an overview of the project groups. Students were divided into Control, Enhanced Prompt (EP), and CTA groups and all teams completed a class project. After the project, all groups received baseline information on the CATME survey through a short instructional video I created. The control and EP groups completed a reflection on study skills while the CTA group completed the CTA reflection on their teaming experience (instructional manipulation). Finally, each group completed a peer evaluation survey. The Control group completed the original CATME peer evaluation (administered through Survey Monkey), while the EP and CTA groups completed the modified CATME peer evaluation (administered through Survey Monkey: survey manipulation). Essentially, changes in the dependent variables from the control to EP group should be the result of the survey manipulation. Additional changes between the EP and CTA group should be the result of the instructional manipulation.

1	CONTROL	2	ENHANCED PROMPT	3	CTA
1.	Students watch peer evaluation instructional videos	1.	Students watch peer evaluation instructional videos	1.	Students watch peer evaluation instructional videos
2.	Study Skills Reflection	2.	Study Skills Reflection	2.	CTA reflection on teaming experience
3.	Take CATME Survey	3.	Take CATME Survey with enhanced feedback prompts	3.	Take CATME Survey with enhanced feedback prompts

Figure 6.4. Fall 2020 Study Overview

6.1.3.1 Hypotheses

For this study, my hypotheses were that the modifications to the CATME survey should increase feedback utility. Utility will be operationally defined as detailed, references specific team work, and provides actionable and/or impactful feedback. Additionally, the use of the CTA teaming reflection will show increased feedback utility over the modifications to the CATME survey alone. Specifically:

H1: Students in the EP and CTA groups will demonstrate increased comments on CATME teaming behaviors in their feedback over the Control group. No differences are expected between the EP and CTA groups.

H2: Students in the CTA group will demonstrate increased comments referencing specific team tasks compared to both the control and EP groups. The EP group will demonstrate increased comments referencing specific team tasks as compared to the Control group.

H3: Students in the EP and CTA groups will demonstrate increased comments with actionable feedback over the Control group. No differences are expected between the EP and CTA groups.

H4: Students in the EP and CTA groups will demonstrate increased comments with feedback that explains the impact on the team over the Control group. No differences are expected between the EP and CTA groups.

H5: Students in the CTA group will demonstrate more complete feedback compared to both the Control group and the EP group. The EP group will demonstrate more complete feedback as compared to the Control group. More complete feedback will be operationally defined as feedback which provides more information to the student receiving the feedback and consists of not only the CATME behaviors, but also includes two or more of the following: specific teaming examples, the impact of the behavior, and actionable feedback.

6.1.3.2 Procedure

Figure 6.5 shows the timeline of the Fall 2020 study. At the beginning of the semester (weeks 2-3), ENG1101 instructors created project teams for the semester. During week 4, the teams began working on the Heat Treatment Project, which they completed week 7.

Figure 6.5. Fall 2020 Study Timeline

During week 8, teams were randomly assigned to the control, EP, or CTA groups within each instructor's sections and completed the Midterm Peer Evaluation. To make sure the students completed the correct assignments, the students were assigned to sections in Canvas (control, EP, and CTA) and assigned to a Zoom breakout room where they completed the activities assigned to their study group in class as shown in Figure 6.6 below. The students were directed to the module with their breakout room on it (1 = control, 2 = EP or 3 = CTA).

Figure 6.6. Zoom Breakout Room Activities

After completing the midterm peer evaluations, the comments were organized and returned anonymously to the individual students during week 9. For all groups, the final peer evaluation was administered during the final week of the Fall 2020 semester (week 14). As with the midterm peer evaluation, all groups were asked to rate their teammates on the five CATME dimensions of teaming (Contributing to Team's Work, Interacting with Teammates, Keeping the Team on Track, Expecting Quality, and Having Related Knowledge and Skills). After this quantitative rating, the Control group were given the original CATME prompt within the Survey Monkey Survey. The enhanced prompt was

used for both the EP and the CTA groups. After each prompt, the students were provided with the response box to comment in for each of their teammates.

6.2 Results

In this section, I focus on discussing the results from the peer evaluation data from the midterm and end of term evaluations. These results are organized by first focusing on the differences between the CATME behavioral codes between the experimental groups in Section 6.2.1. I start by evaluating the results from the primary codes and then move on to the secondary codes. In Section 6.2.2, I discuss the correlations calculated between the midterm self-evaluation and peer evaluation comments and team satisfaction.

6.2.1 CATME Behavioral Code Results

As stated in Table 6.1, there were a total of 45 teams evenly distributed between our experimental groups in this study. A total of 131 students consented to the survey (34% control, 30% EP, and 36% CTA) and generated a total of 406 comments for the midterm peer evaluation. The Control group generated 146 of these comments, 107 were generated by the EP group and 153 comments were generated by the CTA group. A total of 111 of the 131 students who consented to the survey also completed the final peer evaluation study (33% Control, 33% EP, and 34% CTA) and generated a total of 395 comments for the end-of-term peer evaluations. These comments are analyzed in the following sections.

I first reviewed the normality and variance of each of the dependent variables (primary and secondary code results and the new team satisfaction measure) using the

Explore function of SPSS. By inspection of the histograms, the majority of the data does not appear to follow a normal distribution. Most of the data are positively skewed (shifted toward lower numbers of behaviors coded with primary or secondary codes). Although the data do not follow a normal distribution, the tests performed for this study are considered “robust” to violations of normality. Additionally, since the distributions are all skewed in the same direction, I considered the assumptions to be met. Additionally, since I have only two categories in my within-subjects variable (timing), I interpreted my results as though the assumption of sphericity was met (www.laerd.com).

To begin, I ran a two-way repeated measure MANOVA to determine the effect of the group assignment (e.g., survey and instructional interventions; between subjects variable) over time (within subjects variable) on the feedback and team functionality. Although it is typical to use the Wilks’ λ test to determine significant effects, according to Field (2018), Roy’s largest root ($\phi_{\max}$) may be the most appropriate in some cases as it “represents the maximum possible between-group difference given the data collected” (Field, 2018), so this test was used in this instance.

A statistically significant effect of time on the combined dependent variables was found, $F(9,100) = 7.252, p < 0.001, \phi_{\max} = 0.653, \text{partial } \eta^2 = 0.395$. There was also a statistically significant effect of group on the combined dependent variables, $F(9,101) = 2.112, p = 0.035, \phi_{\max} = 0.188, \text{partial } \eta^2 = 0.158$. There was not a statistically significant interaction between time and group found on the dependent variables, $F(9,101) = 0.660, p = 0.743, \phi_{\max} = 0.059, \text{partial } \eta^2 = 0.056$. In the following sections, I will discuss both

the results of the effects of group and the effects of time on the dependent variables. I start by discussing the results for Hypothesis 1, which focused on the CATME behavioral (or primary) coding. I then discuss the results for the secondary coding, which address hypotheses 2-5.

6.2.1.1 Primary Code Results

Hypothesis 1 stated that: “students in the EP and CTA groups will demonstrate increased comments on CATME teaming behaviors in their feedback over the Control group. No differences are expected between the EP and CTA groups.” Increased comments was assessed based upon the number of CATME dimensions of teaming (Contributing to Team’s Work, Interacting with Teammates, Keeping the Team on Track, Expecting Quality, and Having Related Knowledge and Skills) discussed in each comment. For example, the following comment would have one dimension, Contributing to Team’s Work, C, addressed “I felt I do my part. I do get confused sometimes and need help but I help others when they need it and I understand that section.” On the other hand, this comment; “Helps others when they are struggling with something. Helps to keep the team on track and coordinates when you can meet up”, was coded as having two dimensions (C and K).

To begin, I looked at the differences between the study groups in quantity of behaviors addressed in the comments, which is shown in Figure 6.7 below. In general, it appears that the number of behavioral categories mentioned changed across the study groups, with the EP group showing an increase over the Control group (with the

exception of the Contributing to Team's Work category) and the CTA group showing an increase in number of categories over the EP group.

Figure 6.7. Distribution of Behavioral Category Comments by Study Group (Midterm Evaluations)

To test hypothesis one, I used a one-way ANOVA to compare the quantities of behavioral comments between the three study groups. Only two behavioral categories differed significantly between groups: Contributing to Team's Work (C) and Expecting Quality (EQ) behavioral categories. Significant differences in C, $F(2,403) = 3.190$, $p = .042$, were evaluated using the Tukey HSD post-hoc comparison. Significant differences ($p = 0.034$) existed between the control ($M = 0.75$, $SD = 0.52$) and the CTA group ($M = 0.90$, $SD = 0.50$) for this behavior. No differences were found between the EP group ($M = 0.84$, $SD = 0.52$) and the other two study groups. Figure 6.8 shows the results of this

analysis. To designate the groups on the graphs, the Control group is group 1, the EP group is group 2, and the CTA group is group 3.

Figure 6.8. Contributing to Team's work by Study Group (Midterm Evaluations)
 *Group 1 = Control, Group 2 = EP, Group 3 = CTA

The second behavioral category that showed significant results using a one-way ANOVA was the Expecting Quality (EQ) category, shown in Figure 6.9, $F(2,403) = 7.411, p = .001$. Post-hoc comparisons using the Tukey HSD test indicated that the CTA group ($M = 0.14, SD = 0.35$) provided a significantly higher number of comments in this category ($p < 0.001$) than the Control group ($M = 0.02, SD = 0.14$). Once again, there were no statistically significant differences between the EP group ($M = 0.10, SD = 0.31$) and the other study groups.

Figure 6.10 shows that students in the CTA group ($M = 2.27, SD = 1.41$) provided more comments in more categories on average than the EP group ($M = 2.07, SD = 1.39$) and the Control group ($M = 1.77, SD = 1.00$). While this shows a larger number of

categories addressed by the students in the EP and CTA groups over the control group, there were only statistically significant differences between the control and CTA groups ($p = 0.003$).

Figure 6.9. Expecting Quality by Study Group (Midterm Evaluations)
*Group 1 = Control, Group 2 = EP, Group 3 = CTA

Figure 6.10. Average Behaviors by Study Group (Midterm Evaluations)
*Group 1 = Control, Group 2 = EP, Group 3 = CTA

With this study design it is difficult to determine the specific effects of modifying the survey or the instructional manipulation using an ANOVA as the EP group has both manipulations. Therefore, I calculated the Cohen's d values to determine the effects for the following comparisons. First, I compared the Control group and the EP group to determine the effects of the survey manipulation. Then, I compared the results from the EP and CTA group to determine the effects of the instructional manipulation. Finally, I compared the results of the control and CTA group to determine if there are any combined effects of the survey and instructional manipulation.

Table 6.7 shows the effect sizes for the different manipulations. In general, the behavioral comments for Contributing to the Team's Work and Interacting with Teammates show a small effect due to the survey manipulation, while the Expecting Quality and Average Behaviors categories show a medium effect due to the survey manipulation. The instruction manipulation shows a small effect on all the behavioral categories except Having Related Knowledge, Skills, and Abilities, which does not show any effect due to the instructional manipulation. In addition, there appears to be a small effect on the Average Behaviors in the peer evaluation due to the instructional manipulation. With respect to the significant differences between the control and CTA groups in the Expecting Quality and Average Behavior categories listed, there appears to be a combined effect of both the survey and instructional manipulations.

Table 6.7. Effects of Study Manipulations on Behavioral Categories (Midterm Evaluations)

Codes	Cohen's d values		
	Survey	Instruction	Both
Contributing to the Team's Work (C)	0.18	0.11	0.29
Interacting with Teammates (IT)	0.18	0.08	0.25
Keeping the Team on Track (K)	0.05	0.10	0.15
Expecting Quality (EQ)	0.35	0.12	0.46
Having Related Knowledge, Skills, and Abilities (KSA)	0.02	0.00	0.02
Average Behaviors	0.25	0.13	0.39

In addition to a statistically significant effect of group on the combined dependent variables, there was also a statistically significant effect of time on the combined dependent variables (shown in Table 6.8). Looking at the behavioral codes, Interacting with Teammates (IT) and Having Related Knowledge, Skills, and Abilities (KSA) showed significant differences between the mid and end-of-term peer evaluations. These differences are shown in Figure 6.11 below.

In general, there were less comments provided on these categories at the end-of-term semester than the midterm peer evaluation. The general pattern of frequency of comments by category appears to be the same as on the midterm evaluations with the most comments on the Contributing to Team's Work (C), then IT, KSA, K, and EQ with the least comments. Looking at the two behavioral categories that showed significant differences in the repeated measures MANOVA, there was an overall 12% decrease in

the number comments on the Interacting with Teammates (IT) category. Additionally, 13% less comments were provided on the Having Related Knowledge, Skills, and Abilities (KSA) category.

Table 6.8. Changes in Primary Codes over Time

Dependent Variable	(df, Error)	F	p	partial η^2
Contributing to Team's Work, C	(1,108)	1.515	0.221	0.014
Interacting with Teammates, IT	(1,108)	5.568	0.020*	0.049
Keeping the Team on Track, K	(1,108)	3.894	0.051	0.035
Expecting Quality, EQ	(1,108)	2.509	0.116	0.023
Having Related Knowledge, Skills, and Abilities, KSA	(1,108)	1458.197	0.000**	0.931

* p < 0.05 ** p < 0.01

Figure 6.11. Distribution of Behavioral Category Comments over Time

6.2.1.2 Secondary Code Results

For this section, I discuss the results from Hypotheses 2-5. All of these hypotheses focus on the quantification of the secondary codes (FE, FI, FA) as a way to measure changes in utility of the peer evaluation measure and the effects of the experimental manipulation on the peer evaluation results. For the second hypothesis, it was expected that students in the CTA group will demonstrate increased comments referencing specific team tasks compared to both the control and EP groups. Additionally, it is expected that the EP group will demonstrate increased comments referencing specific team tasks as compared to the Control group. This was measured by the number of comments coded with the secondary code, FE. As shown in Figure 6.12 below, all groups averaged at least one comment that was coded as a specific teaming example within the midterm peer evaluation comments. However, a one-way ANOVA did not indicate any statistically significant differences between the study groups, $F(2,401) = 1.827, p = .162$.

Figure 6.12. Feedback Examples by Study Group (Midterm Evaluations)

*Group 1 = Control, Group 2 = EP, Group 3 = CTA

For the third hypothesis, it was expected that students in the EP and CTA groups would demonstrate increased comments with actionable feedback over the Control group. Additionally, no differences were expected between the EP and CTA groups. This was measured by the number of comments coded with the secondary code, FA. A one-way ANOVA indicated a significant effect on the comments coded with FA between groups, $F(2,401) = 5.798, p = .003$. Figure 6.13 shows the results of this analysis. Post-hoc comparisons using the Tukey HSD test indicated that students in the EP group ($M = 0.38, SD = 0.62$), showed a significantly higher number of comments with this code ($p = 0.008$) than the Control group ($M = 0.19, SD = 0.41$). Additionally, the CTA group ($M = 0.37, SD = 0.53$) shown a significantly higher number of comments with this code ($p = 0.012$) than the Control group as well. There were no statistically significant differences between the EP group and the CTA group ($p = 0.916$).

Figure 6.13. Actionable Feedback by Study Group (Midterm Evaluations)

*Group 1 = Control, Group 2 = EP, Group 3 = CTA

For hypothesis four, it was expected that students in the EP and CTA groups would demonstrate increased comments with feedback that explains the impact on the team over the Control group. Additionally, no differences were expected between the EP and CTA groups. This was measured by the number of comments coded with the secondary code, FI. Figure 6.14 shows the results of this analysis. Using a one-way ANOVA, it was determined that there was a significant effect on the comments on impact feedback, $F(2,401) = 3.814, p = .023$. Post-hoc comparisons using the Tukey HSD test indicated that students in the CTA group ($M = 0.59, SD = 0.70$) showed a significantly higher number of comments with this code ($p = 0.033$) than the Control group ($M = 0.40, SD = 0.58$). However, there were no statistically significant differences between the EP group ($M = 0.57, SD = 0.68$) and the CTA group ($p = 0.994$). The difference between the Control group and the EP group approaches significance ($p = 0.076$) with students in the EP group more likely to include comments on behavioral impacts in their peer evaluations.

Figure 6.14. Impact Feedback by Study Group (Midterm Evaluations)
 *Group 1 = Control, Group 2 = EP, Group 3 = CTA

Hypothesis 5 states that students in the CTA group would demonstrate more complete feedback compared to both the Control group and the EP group. Additionally, it was hypothesized that the EP group would demonstrate more complete feedback as compared to the Control group. In this case, “more complete feedback” was operationally defined as feedback which provided more information to the student receiving the feedback and consisted of not only the CATME behaviors, but included at least two of the secondary codes (FE, FI, and FA) as well. Figure 6.15 identifies the average number of comments which contained at least two of the three secondary codes in each study group. Figure 6.16 displays the average number of comments which contained all three secondary codes in each study group.

Using a one-way ANOVA, there was a significant effect on the comments providing at least two secondary codes, $F(2,401) = 6.356, p = .002$. Post-hoc comparisons

using the Tukey HSD test indicated that students in the EP group ($M = 0.74$, $SD = 0.98$) were significantly more likely ($p = 0.004$) to provide at least two secondary codes in their peer evaluation comments over the Control group ($M = 0.42$, $SD = 0.60$). Additionally, students in the CTA group ($M = 0.70$, $SD = 0.87$) were significantly more likely to provide at least two secondary codes in their peer evaluation comments over the Control group ($p = 0.011$). No significant differences were found between the EP and CTA groups when looking at this measure ($p = 0.843$).

Figure 6.15. Comments containing at least two modifiers by Study Group (Midterm Evaluations)

*Group 1 = Control, Group 2 = EP, Group 3 = CTA

Using a one-way ANOVA, there was a significant effect between groups on the comments providing all three secondary codes, $F(2,401) = 4.320$, $p = .014$. Post-hoc comparisons using the Tukey HSD test indicated no statistically significant differences ($p = 0.941$) between the EP ($M = 0.10$, $SD = 0.32$) and CTA group ($M = 0.10$, $SD = 0.31$)

when looking at all the average number of comments that that included all three secondary codes (Figure 6.16). There are statistically significant differences between the control ($M = 0.02$, $SD = 0.14$) and the CTA group ($p = 0.018$) with students in the CTA group providing more comments with all three secondary codes. There was a non-significant trend toward the EP group providing more comments with all the secondary codes than the Control group ($p = 0.079$).

Figure 6.16. Comments containing all three modifiers by Study Group (Midterm Evaluations)

*Group 1 = Control, Group 2 = EP, Group 3 = CTA

Once again, it is difficult to determine the individual and combined effects of the survey manipulation and the instructional manipulation using an ANOVA. To address this, I calculated the effect sizes (Cohen's d) for the different manipulations, which are shown in Table 6.9 below. In general, the survey manipulation shows a medium effect for all secondary codes individually (FE, FI, FA) and the use of two or all three secondary

codes in the peer evaluation comments. With the small, or negative, values for the instructional manipulation, it does not appear that there is an effect due to the instructional manipulation on the secondary codes. With the significant differences between the control and CTA group on the use of all three modifiers, there appears to be a combined effect of the survey and instructional manipulation in this case.

Table 6.9. Effects of Study Manipulations on Secondary Codes (FE, FI, FA) (Midterm Evaluations)

Codes	Cohen's d values		
	Survey	Instruction	Both
Feedback with Team Example (FE)	0.21	-0.02	0.18
Feedback with Impact (FI)	0.29	0.01	0.30
Feedback with Action (FA)	0.37	-0.04	0.36
Two Secondary Codes	0.41	-0.06	0.37
Three Secondary Codes	0.29	0.04	0.35

Table 6.10 shows the overall within-subjects effects of time on the dependent variables. Of the secondary codes, only Feedback with Action (FA) showed significant differences between the peer evaluation times. There were also significant differences on the team satisfaction measure between the midterm and end-of term peer evaluation ratings. These differences will be expanded upon below in Figure 6.17. As with the primary behavioral codes, there were less comments provided on the secondary codes at the end-of-term semester than the midterm peer evaluation. Additionally, there was significant decrease in the amount of actionable feedback (FA) by 35%. Team

Satisfaction, however showed a significant increase from the midterm ($M = 90.7$, $SD = 13.2$) to the end-of-term evaluation ($M = 95.0$, $SD = 6.5$).

Table 6.10. Changes in Secondary Codes over Time

Dependent Variable	(df, Error)	F	p	partial η^2
Feedback with Team Example, FE	(1,108)	1.784	0.184	0.016
Feedback with Impact, FI	(1,108)	0.645	0.424	0.006
Feedback with Action, FA	(1,108)	29.380	0.000**	0.214
Team Satisfaction	(1,108)	12.060	0.001**	0.100

* $p < 0.05$ ** $p < 0.01$

Figure 6.17. Distribution of Secondary Behavioral Comments over Time

6.2.2 Relationships

These previous analyses focused on exploring the differences between the experimental groups on the use of the behavioral comments, specific examples, impact statements, and

actionable feedback. This analysis focused on exploring the relationship between the use of the codes between groups. Additionally, since the Spring 2020 study (Study 3) did not indicate a relationship between the numerical CATME scores for the behavioral categories, I investigated the relationship between the frequency of comments provided by students in the different behavioral categories, comments that were coded with the secondary codes (FE, FI, FA) and their self-reported team satisfaction.

6.2.2.1 Self-Evaluations

To begin, I investigated the comments that were directed toward the person completing the evaluation (self-evaluation) and their self-reported team satisfaction. Tables 6.11–6.14 show the correlations between the frequency of these comments and team satisfaction by each group: control, EP, CTA, and overall. Some interesting things to note for the Control group. First, comments on Contributing to the Team’s Work are negatively correlated with Interacting with Teammates. This would indicate that students who were more likely to comment on Contributing to Team’s Work were less likely to comment on Interacting with Teammates or vis versa. Students who commented on Interacting with Teammates were more likely to also include a teaming example.

Additionally, students who commented on Keeping the Team on Track, were also significantly more likely to provide a team related example and an impact. In general, students who provided a team example, also provided an impact, but were less likely to include actionable feedback or those who provided actionable feedback were less likely to provide a team example or the impact of the behavior. It is interesting to note that comments on Contributing to the Team’s Work were positively correlated with Team

Satisfaction, which may indicate that when students observed others actively contributing to the work, they were more satisfied with their team. It was also interesting that actionable feedback was negatively correlated with Team Satisfaction for the Control group, which may indicate that there were many behaviors that students wished would change in their teammates and were currently unsatisfied with their teaming experience.

Looking at the EP group (Table 6.12), students were more likely to comment on more than one behavioral category. There are positive correlations between Contributing to the Team's Work, Expecting Quality, and Having Related Knowledge, Skills, and Abilities. Additionally, there are positive correlations between comments on Interacting with Teammates and Expecting Quality. With the exception of the Keeping the Team on Track Category, all the behavioral categories were positively correlated with providing teaming examples. Only the Keeping the Team on Track and Having Related Knowledge, Skills and Abilities categories have positive correlations with feedback with impact. There are also positive correlations between team feedback examples and feedback with impact. There were no significant correlations between any measure and actionable feedback or team satisfaction.

For the self-evaluations of the CTA group (Table 6.13), there were positive correlations between comments provided in the Expecting Quality, Interacting with Teammates, Keeping the Team on Track, and Having Related Knowledge, Skills, and Abilities categories. All of these behavioral categories were positively correlated with teaming feedback examples, however, only the Having Related Knowledge, Skills, and Abilities category had positive correlations with feedback with impact. As with the EP

group, there were no significant correlations between any measure and actionable feedback or team satisfaction.

Table 6.14 shows the correlations between all groups together. In general, there are many positive comments between the different teaming behavior categories. What is interesting to note is that if students do comment on the Expecting Quality category, they are also more likely to comment on all the other teaming behaviors. All behavior categories have positive correlations with teaming examples, but only Interacting with Teammates has actionable feedback. Additionally, only Contributing to Team's Work has significant correlations with Team Satisfaction.

In general, for the self-evaluations, the students who were most likely to have a higher satisfaction rating were also more likely to comment Contributing to Team's Work and less likely to include actionable feedback. This may indicate that students were more satisfied when they saw work being done for the team that did not require teammates to change the way they contributed. For the EP, there did not appear to be any significant contributions to team satisfaction from comments in any behavioral or secondary category. However, it appears that students were providing examples for most of the behaviors that occurred. Finally, the CTA group demonstrated that they were more likely to provide comments in all of the behavioral categories and include specific teaming examples. However, these results did not correlate with team satisfaction.

Table 6.11. Correlations between Qualitative Self-Evaluation Comments on Midterm Peer Evaluations (Control: n=45)⁺

Measures	1	2	3	4	5	6	7	8	9
1. Contributing to Team's Work (C)	--	-.391** (.008)	.239 (.114)	.081 (.597)	-.139 (.361)	.257 (.089)	.197 (.195)	-.121 (.428)	.330* (.027)
2. Interacting with Teammates (IT)	-.391** (.008)	--	.404** (.006)	-.147 (.337)	-.082 (.594)	.302* (.044)	.086 (.574)	.219 (.148)	-.207 (.173)
3. Keeping the Team on Track (K)	.239 (.114)	.404** (.006)	--	-.088 (.566)	-.049 (.750)	.651** (.000)	.549** (.000)	-.263 (.081)	.143 (.348)
4. Expecting Quality (EQ)	.081 (.597)	-.147 (.337)	-.088 (.566)	--	.080 (.601)	-.015 (.923)	-.148 (.331)	-.139 (.362)	.057 (.710)
5. Having Related Knowledge, Skills, and Abilities (KSA)	-.139 (.361)	-.082 (.594)	-.049 (.750)	.080 (.601)	--	.272 (.071)	.149 (.330)	-.141 (.356)	-.029 (.852)
6. Feedback with Team Example (FE)	.257 (.089)	.302* (.044)	.651** (.000)	-.015 (.923)	.272 (.071)	--	.509** (.000)	-.309* (.039)	.069 (.650)
7. Feedback with Impact (FI)	.197 (.195)	.086 (.574)	.549** (.000)	-.148 (.331)	.149 (.330)	.509** (.000)	--	-.304* (.043)	.111 (.466)
8. Feedback with Action (FA)	-.121 (.428)	.219 (.148)	-.263 (.081)	-.139 (.362)	-.141 (.356)	-.309* (.039)	-.304* (.043)	--	-.318* (.033)
9. Satisfaction with Team	.330* (.027)	-.207 (.173)	.143 (.348)	.057 (.710)	-.029 (.852)	.069 (.650)	.111 (.466)	-.318* (.033)	--

⁺ table notation Pearson r (p value)

* p < .05 ** p<0.01

Table 6.12. Correlations between Qualitative Self-Evaluation Comments on Midterm Peer Evaluations (EP: n = 39) ⁺

Measures	1	2	3	4	5	6	7	8	9
1. Contributing to Team's Work (C)	--	.145 (.379)	-.023 (.888)	.454** (.004)	.378* (.018)	.572** (.000)	.299 (.064)	.128 (.439)	.080 (.628)
2. Interacting with Teammates (IT)	.145 (.379)	--	-.071 (.666)	.386* (.015)	.300 (.064)	.448** (.004)	.043 (.796)	.173 (.293)	-.073 (.657)
3. Keeping the Team on Track (K)	-.023 (.888)	-.071 (.666)	--	.116 (.483)	-.065 (.696)	.159 (.333)	.473** (.002)	.153 (.352)	-.156 (.344)
4. Expecting Quality (EQ)	.454** (.004)	.386* (.015)	.116 (.483)	--	.530** (.001)	.768** (.000)	.231 (.158)	.150 (.361)	.127 (.442)
5. Having Related Knowledge, Skills, and Abilities (KSA)	.378* (.018)	.300 (.064)	-.065 (.696)	.530** (.001)	--	.579** (.000)	.326* (.043)	-.106 (.521)	.137 (.406)
6. Feedback with Team Example (FE)	.572** (.000)	.448** (.004)	.159 (.333)	.768** (.000)	.579** (.000)	--	.415** (.009)	.275 (.090)	.010 (.952)
7. Feedback with Impact (FI)	.299 (.064)	.043 (.796)	.473** (.002)	.231 (.158)	.326* (.043)	.415** (.009)	--	.254 (.119)	-.148 (.368)
8. Feedback with Action (FA)	.128 (.439)	.173 (.293)	.153 (.352)	.150 (.361)	-.106 (.521)	.275 (.090)	.254 (.119)	--	.003 (.987)
9. Satisfaction with Team	.080 (.628)	-.073 (.657)	-.156 (.344)	.127 (.442)	.137 (.406)	.010 (.952)	-.148 (.368)	.003 (.987)	--

⁺ table notation Pearson r (p value)

* p < .05 ** p < .01

Table 6.13. Correlations between Qualitative Self-Evaluation Comments on Midterm Peer Evaluations (CTA: n=47)⁺

Measures	1	2	3	4	5	6	7	8	9
1. Contributing to Team's Work (C)	--	.111 (.456)	.240 (.104)	.142 (.341)	.116 (.436)	.396** (.006)	.123 (.412)	-.182 (.221)	.256 (.082)
2. Interacting with Teammates (IT)	.111 (.456)	--	.352* (.015)	.323* (.027)	.130 (.384)	.305* (.037)	.118 (.429)	.161 (.279)	.273 (.063)
3. Keeping the Team on Track (K)	.240 (.104)	.352* (.015)	--	.520** (.000)	.261 (.077)	.485** (.001)	.326* (.025)	.046 (.758)	.146 (.326)
4. Expecting Quality (EQ)	.142 (.341)	.323* (.027)	.520** (.000)	--	.511* (.000)	.572** (.000)	.182 (.221)	.089 (.551)	.178 (.232)
5. Having Related Knowledge, Skills, and Abilities (KSA)	.116 (.436)	.130 (.384)	.261 (.077)	.511* (.000)	--	.687** (.000)	.329** (.009)	-.175 (.238)	.079 (.597)
6. Feedback with Team Example (FE)	.396** (.006)	.305* (.037)	.485** (.001)	.572** (.000)	.687** (.000)	--	.392** (.006)	-.038 (.801)	.225 (.129)
7. Feedback with Impact (FI)	.123 (.412)	.118 (.429)	.326* (.025)	.182 (.221)	.329** (.009)	.392** (.006)	--	-.059 (.693)	.009 (.953)
8. Feedback with Action (FA)	-.182 (.221)	.161 (.279)	.046 (.758)	.089 (.551)	-.175 (.238)	-.038 (.801)	-.059 (.693)	--	.091 (.545)
9. Satisfaction with Team	.256 (.082)	.273 (.063)	.146 (.326)	.178 (.232)	.079 (.597)	.225 (.129)	.009 (.953)	.091 (.545)	--

⁺ table notation Pearson r (p value)

* p < .05 ** p<0.01

Table 6.14. Correlations between Qualitative Self-Evaluation Comments on Midterm Peer Evaluations (All groups: n=131)⁺

Measures	1	2	3	4	5	6	7	8	9
1. Contributing to Team's Work (C)	--	-.054 (.537)	.176* (.044)	.212* (.015)	.119 (.177)	.402** (.000)	.209* (.017)	-.046 (.604)	.234** (.007)
2. Interacting with Teammates (IT)	-.054 (.537)	--	.277** (.001)	.234** (.007)	.125 (.155)	.347** (.000)	.096 (.276)	.188* (.032)	.023 (.796)
3. Keeping the Team on Track (K)	.176* (.044)	.277** (.001)	--	.302** (.000)	.101 (.252)	.458** (.000)	.441** (.000)	.004 (.959)	.085 (.350)
4. Expecting Quality (EQ)	.212* (.015)	.234** (.007)	.302** (.000)	--	.433** (.000)	.513** (.000)	.136 (.121)	.088 (.318)	.123 (.160)
5. Having Related Knowledge, Skills, and Abilities (KSA)	.119 (.177)	.125 (.155)	.101 (.252)	.433** (.000)	--	.551** (.000)	.303** (.000)	-.112 (.204)	.049 (.582)
6. Feedback with Team Example (FE)	.402** (.000)	.347** (.000)	.458** (.000)	.513** (.000)	.551** (.000)	--	.432** (.000)	.013 (.886)	.131 (.137)
7. Feedback with Impact (FI)	.209* (.017)	.096 (.276)	.441** (.000)	.136 (.121)	.303** (.000)	.432** (.000)	--	.005 (.955)	.001 (.991)
8. Feedback with Action (FA)	-.046 (.604)	.188* (.032)	.004 (.959)	.088 (.318)	-.112 (.204)	.013 (.886)	.005 (.955)	--	-.060 (.498)
9. Satisfaction with Team	.234** (.007)	.023 (.796)	.085 (.350)	.123 (.160)	.049 (.582)	.131 (.137)	.001 (.991)	-.060 (.498)	--

⁺ table notation Pearson r (p value)

* p < .05 ** p<0.01

6.2.2.2 *Peer-Evaluations*

Tables 6.15 - 6.18 show the correlations between the frequency of the peer evaluation comments and team satisfaction by each group: control, EP, CTA, and overall. Note, the items in these tables are comments directed to the teammates of the person completing the evaluation (peer-evaluation) and their self-reported team satisfaction. These are normalized by group size so these comments represent the average number of comments provided to their group members.

Looking at the comments of the Control group students towards their teammates (Table 6.15), there are no significant correlations between any of the behavioral categories. However, students were more likely to provide teaming examples for the Contributing to Team's Work and Interacting with Teammates categories and provide impact feedback for the Keeping the Team on Track category. Significant actionable feedback was not provided for any behavioral category. It is interesting to note that there is a negative correlation between Having Related Knowledge, Skills, and Abilities and Team Satisfaction. This would seem to indicate that that more comments on this category results in a lower Team Satisfaction rating or vis versa.

Looking at the EP group (Table 6.16), this group showed positive correlations with Expecting Quality and Interacting with Teammates, Keeping the Team on Track, and Having Related Knowledge, Skills, and Abilities. It is interesting to note that Interacting with Teammates is not significantly correlated with any other behavioral category besides Expecting Quality. This would appear to indicate that comments for this

category are significantly more likely to be provided by students who also comment on the Expecting Quality category. This may indicate that the interactions they are observing with their teammates are related to the quality of the work. Comments on Having Related Knowledge, Skills, and Abilities are also positively correlated with comments on Keeping the Team on Track. For the EP group, all behavioral categories are positively correlated with teaming examples, but none with impactful or actionable feedback or with Team Satisfaction.

For the CTA group (Table 6.17), there are few significant correlations between the behavioral categories. There are only positive correlations between the Keeping the Team on Track and Having Related Knowledge, Skills, and Abilities categories. However, all behavioral categories with the exception of Expecting Quality have significant positive correlations with teaming examples. Keeping the Team on Track also has impact feedback. Once again, none of the behavioral categories have significant correlations between actionable feedback or Team Satisfaction.

Looking at the overall correlations on peer-evaluation comments (comments on teammates that do not include the person completing the evaluation) in Table 6.18, shows again positive correlations between many of the behavioral categories with the exception of the Contributing to the Team's Work category. There were, however, positive correlations between all categories and team example feedback and (with the exception of the Having Related Knowledge, Skills, and Abilities category) with the impact feedback. However, none of these categories showed significant correlations with actionable feedback or Team Satisfaction.

In general, looking at the peer comments, the Control group students who provided teaming examples provided them in the Contributing to Team's Work and Interacting with Teammates categories. These are areas where it is most likely they would see observable behaviors within the team. It is also more likely that students who provided an impact of a behavior did so in the Keeping the Team on Track category. This is also an observable behavior. With the EP and CTA groups, these students were more likely to comment on more categories than the control group and provided examples with most of the behavioral categories. The CTA group also were more likely to provide impact feedback when commenting on Keeping the Team on Track. However, none of the groups showed significant correlations between their peer comments and actionable feedback or team satisfaction.

Table 6.15. Correlations between Qualitative Peer-Evaluation Comments on Midterm Peer Evaluations (Control: n=44)⁺

Measures	1	2	3	4	5	6	7	8	9
1. Contributing to Team's Work (C)	--	.161 (.298)	-.148 (.338)	-.100 (.518)	-.233 (.128)	.344* (.022)	.042 (.789)	-.187 (.223)	.077 (.618)
2. Interacting with Teammates (IT)	.161 (.298)	--	-.252 (.099)	.019 (.900)	.261 (.087)	.442** (.003)	.019 (.904)	.004 (.982)	-.173 (.261)
3. Keeping the Team on Track (K)	-.148 (.338)	-.252 (.099)	--	.148 (.337)	-.269 (.078)	.021 (.893)	.472** (.001)	-.030 (.846)	-.237 (.122)
4. Expecting Quality (EQ)	-.100 (.518)	.019 (.900)	.148 (.337)	--	.039 (.804)	.211 (.168)	-.177 (.251)	-.079 (.610)	.098 (.526)
5. Having Related Knowledge, Skills, and Abilities (KSA)	-.233 (.128)	.261 (.087)	-.269 (.078)	.039 (.804)	--	.275 (.070)	-.149 (.335)	.079 (.611)	-.321* (.034)
6. Feedback with Team Example (FE)	.344* (.022)	.442** (.003)	.021 (.893)	.211 (.168)	.275 (.070)	--	-.167 (.278)	.040 (.795)	-.046 (.768)
7. Feedback with Impact (FI)	.042 (.789)	.019 (.904)	.472** (.001)	-.177 (.251)	-.149 (.335)	-.167 (.278)	--	.016 (.916)	-.081 (.600)
8. Feedback with Action (FA)	-.187 (.223)	.004 (.982)	-.030 (.846)	-.079 (.610)	.079 (.611)	.040 (.795)	.016 (.916)	--	.002 (.992)
9. Satisfaction with Team	.077 (.618)	-.173 (.261)	-.237 (.122)	.098 (.526)	-.321* (.034)	-.046 (.768)	-.081 (.600)	.002 (.992)	--

⁺ table notation Pearson r (p value)

* p < .05 ** p<0.01

Table 6.16. Correlations between Qualitative Peer-Evaluation Comments on Midterm Peer Evaluations (EP: n=38)⁺

Measures	1	2	3	4	5	6	7	8	9
1. Contributing to Team's Work (C)	--	.159 (.341)	-.108 (.520)	.059 (.725)	.114 (.494)	.440** (.006)	.287 (.080)	-.180 (.279)	.097 (.562)
2. Interacting with Teammates (IT)	.159 (.341)	--	.243 (.141)	.449** (.005)	.300 (.067)	.659** (.000)	.257 (.120)	-.087 (.602)	.070 (.675)
3. Keeping the Team on Track (K)	-.108 (.520)	.243 (.141)	--	.493** (.002)	.445** (.005)	.435** (.006)	.281 (.088)	-.070 (.678)	-.031 (.853)
4. Expecting Quality (EQ)	.059 (.725)	.449** (.005)	.493** (.002)	--	.499** (.001)	.638** (.000)	.216 (.192)	.116 (.488)	.171 (.304)
5. Having Related Knowledge, Skills, and Abilities (KSA)	.114 (.494)	.300 (.067)	.445** (.005)	.499** (.001)	--	.556** (.000)	.246 (.137)	-.028 (.866)	.224 (.177)
6. Feedback with Team Example (FE)	.440** (.006)	.659** (.000)	.435** (.006)	.638** (.000)	.556** (.000)	--	.323* (.048)	.063 (.708)	.086 (.609)
7. Feedback with Impact (FI)	.287 (.080)	.257 (.120)	.281 (.088)	.216 (.192)	.246 (.137)	.323* (.048)	--	-.177 (.286)	-.014 (.933)
8. Feedback with Action (FA)	-.180 (.279)	-.087 (.602)	-.070 (.678)	.116 (.488)	-.028 (.866)	.063 (.708)	-.177 (.286)	--	-.167 (.316)
9. Satisfaction with Team	.097 (.562)	.070 (.675)	-.031 (.853)	.171 (.304)	.224 (.177)	.086 (.609)	-.014 (.933)	-.167 (.316)	--

⁺ table notation Pearson r (p value)

* p < .05 ** p<0.01

Table 6.17. Correlations between Qualitative Peer-Evaluation Comments on Midterm Peer Evaluations (CTA: n=47)⁺

Measures	1	2	3	4	5	6	7	8	9
1. Contributing to Team's Work (C)	--	.100 (.503)	-.119 (.427)	-.057 (.704)	-.066 (.662)	.333* (.022)	.099 (.509)	-.026 (.861)	.012 (.936)
2. Interacting with Teammates (IT)	.100 (.503)	--	-.044 (.771)	.236 (.110)	.150 (.313)	.376** (.009)	.088 (.555)	.080 (.595)	-.116 (.436)
3. Keeping the Team on Track (K)	-.119 (.427)	-.044 (.771)	--	.246 (.095)	.334* (.022)	.499** (.000)	.532** (.000)	-.139 (.351)	-.030 (.842)
4. Expecting Quality (EQ)	-.057 (.704)	.236 (.110)	.246 (.095)	--	.198 (.181)	.239 (.106)	.196 (.188)	-.082 (.582)	-.117 (.432)
5. Having Related Knowledge, Skills, and Abilities (KSA)	-.066 (.662)	.150 (.313)	.334* (.022)	.198 (.181)	--	.452** (.001)	.283 (.054)	.077 (.605)	.070 (.642)
6. Feedback with Team Example (FE)	.333* (.022)	.376** (.009)	.499** (.000)	.239 (.106)	.452** (.001)	--	.506** (.000)	-.203 (.171)	-.035 (.813)
7. Feedback with Impact (FI)	.099 (.509)	.088 (.555)	.532** (.000)	.196 (.188)	.283 (.054)	.506** (.000)	--	-.236 (.110)	-.234 (.114)
8. Feedback with Action (FA)	-.026 (.861)	.080 (.595)	-.139 (.351)	-.082 (.582)	.077 (.605)	-.203 (.171)	-.236 (.110)	--	.176 (.237)
9. Satisfaction with Team	.012 (.936)	-.116 (.436)	-.030 (.842)	-.117 (.432)	.070 (.642)	-.035 (.813)	-.234 (.114)	.176 (.237)	--

⁺ table notation Pearson r (p value)

* p < .05 ** p<0.01

Table 6.18. Correlations between Qualitative Peer-Evaluation Comments on Midterm Peer Evaluations (All groups: n=131)⁺

Measures	1	2	3	4	5	6	7	8	9
1. Contributing to Team's Work (C)	--	.163 (.065)	-.102 (.248)	.048 (.585)	-.074 (.402)	.379** (.000)	.186* (.035)	-.077 (.384)	.040 (.653)
2. Interacting with Teammates (IT)	.163 (.065)	--	.006 (.942)	.298** (.001)	.230** (.009)	.496** (.000)	.150 (.091)	.010 (.910)	-.093 (.296)
3. Keeping the Team on Track (K)	-.102 (.248)	.006 (.942)	--	.329** (.000)	.185* (.035)	.371** (.000)	.429** (.000)	-.067 (.449)	-.095 (.283)
4. Expecting Quality (EQ)	.048 (.585)	.298** (.001)	.329** (.000)	--	.259** (.003)	.412** (.000)	.212* (.016)	.069 (.436)	-.013 (.883)
5. Having Related Knowledge, Skills, and Abilities (KSA)	-.074 (.402)	.230** (.009)	.185* (.035)	.259** (.003)	--	.430** (.000)	.144 (.105)	.017 (.852)	-.021 (.812)
6. Feedback with Team Example (FE)	.379** (.000)	.496** (.000)	.371** (.000)	.412** (.000)	.430** (.000)	--	.349** (.000)	-.003 (.969)	-.012 (.889)
7. Feedback with Impact (FI)	.186* (.035)	.150 (.091)	.429** (.000)	.212* (.016)	.144 (.105)	.349** (.000)	--	-.111 (.212)	-.149 (.092)
8. Feedback with Action (FA)	-.077 (.384)	.010 (.910)	-.067 (.449)	.069 (.436)	.017 (.852)	-.003 (.969)	-.111 (.212)	--	.002 (.981)
9. Satisfaction with Team	.040 (.653)	-.093 (.296)	-.095 (.283)	-.013 (.883)	-.021 (.812)	-.012 (.889)	-.149 (.092)	.002 (.981)	--

⁺ table notation Pearson r (p value)

* p < .05 ** p<0.01

Using the information provided in Tables 6.11 – 6.18, I completed a regression model to determine which dependent variables are significant contributors to the new measure of team satisfaction. Of the self and peer codes, only self-comments on how an individual Contributes to the Team's Work contributed significantly to the prediction of Team Satisfaction with an R^2 value of 0.056, $F(1,127) = 7.597$, $p=0.007$. This would indicate that the number of comments related to how an individual contributes to their team contributes 5.6% to the prediction of their team satisfaction. This may indicate that their team satisfaction is not based on the number of comments, but the nature of the comments which should be investigated more thoroughly.

6.3 Discussion

There was a statistically significant effect of time and of group on the combined dependent variables, however there was not found to be a significant interaction between the two. In this section, I will therefore discuss the results for each of the five hypotheses on the peer evaluation comments by group (between-subjects) and over time (within-subjects).

First, it was hypothesized that changing the CATME survey by enhancing the prompt would result in significant differences on the number of CATME behavioral categories discussed in the peer comments. These differences should have been observed between the control and EP groups, while no significant differences should exist between the EP and CTA groups (H1). On the midterm evaluations, there were statistically significant results between two of the behavioral categories (Contributing to Team's

Work and Expecting Quality) and the average number of behavioral categories commented on between the control and the EP group, with the EP group providing more comments than the control group.

Looking at the effect size of the study manipulations, there is a small effect for Contributing to Team's Work (C) and a medium effect for Expecting Quality (EQ). Additionally, there is a medium effect of the average number of behavioral categories commented on due to the survey manipulation. As expected, on the midterm peer evaluations, there were no statistically significant differences between the EP and CTA groups. However, there are small effects due to the instructional manipulation on the average number of behavioral categories, which may be combined with the medium effect of the survey manipulation to produce the significant increase in the peer evaluation comments in the CTA group over the control group.

Looking at the ratings over time, in general, the amount of comments on all the behavioral categories decreased. In particular, comments on the Interacting with Teammates (IT) and Having Related Knowledge Skills and Abilities (KSA) categories showed significant decreases in the number of comments provided by students.

To summarize the results of H1, it was hypothesized that changing the CATME survey by enhancing the prompt would result in significant differences on the number of CATME behavioral categories discussed in the peer comments, which would result in observable differences between the control and EP groups. There were significant differences found in specific behavioral categories as well as the average number of

behavioral categories between the control and EP group, which matches my hypothesis. However, there were also significant differences in the average number of behavioral categories discussed in the midterm peer evaluations between the control and CTA group which may be the result of a combination of a small effect due to the instructional manipulation and the medium effect of the survey manipulation. These interactions should be investigated further.

Additionally, it appears that students are decreasing the amount of feedback provided on the end-of term evaluations. It may be due to the perception of the students on how these comments will be used. The midterm evaluation was clearly formative, with feedback returned to the students and the ability to affect their current team's performance may have increased their desired to provide comments. The end-of-term evaluations are more summative. It is unlikely that the current team will benefit from any comments on behaviors to teammates and may have been less likely to provide them. The reasons for this decline in comments should be investigated further.

For my second hypothesis, H2, it was hypothesized the CTA group would provide increased comments that referenced specific team tasks (FE) over both the EP and Control group due to the instructional manipulation, with the EP group providing more comments referencing team tasks over the Control group due to the survey manipulation. As with the behavioral codes, there was a decline in the number of FE comments over time, but not significantly so.

Looking at the between-subject results, I expected that the survey manipulation would increase the number of comments with team examples as the enhanced prompt specifically told the students to “provid[e] examples of these specific behaviors.” Although there was a medium effect of the survey manipulation on FE for the midterm peer evaluation, it did not translate into significant differences between any of the groups. I also expected increases between the EP and CTA group due to the instructional manipulation where students completed an activity which specifically detailed what each team member did for their mini-project, but there were no significant effects or differences found between any of the groups on the midterm peer evaluation. These results may be due to the choice to code more than one behavior in a category as a specific example (FE) during the coding process. This secondary code may need further revision to increase validity.

For the third hypothesis (H3), it was expected that students in the EP and CTA groups would demonstrate increased comments with actionable feedback (FA) over the Control group due to the survey manipulation. I expected that the survey manipulation would increase the number of comments with actionable feedback as the enhanced prompt specifically told the students that they should be “suggesting actions or behaviors that should be added, stopped, altered, or maintained as is.” As expected, there is a medium effect of the survey manipulation on actionable feedback (FA) and students in the EP and CTA groups were statistically more likely to provide more actionable feedback (FA) than the Control group on the midterm evaluations. Because no changes were made in the instructional manipulation to address actionable feedback, no

differences were expected between the EP and CTA groups on actionable feedback (FA). As expected, there were no statistically significant differences between the responses by the students in the EP and CTA groups on the midterm peer evaluation.

There was, however, a significant decrease in the number of actionable feedback provided on the end-of-term evaluation as compared to the midterm evaluation. Again, the reason for this decline should be investigated further, but it may be due to the nature of the assessment. Specific, actionable feedback may be interpreted by the students as things that would help the current team. As the evaluation happened at the end of the semester, there team function had ended.

For the fourth hypothesis (H4), it was expected that both the EP and CTA groups provided increased comments explaining the impact of behaviors on the team (FI) over the Control group due to the survey manipulation. I expected that the survey manipulation would increase the number of comments with impacts of behaviors (FI) as the enhanced prompt told the students they should be “specifying how these behaviors affect your team or your team’s work.” However, there are no significant differences between the control and EP group (although it is close, $p = 0.076$), despite having a medium effect due to the survey manipulation. There are significant differences between the control and CTA groups, which may indicate that there may be a significant effect due to the instructional manipulation or the combination of the survey and instructional manipulation which should be investigated further. There was a decline in the number of FI comments over time, but not significantly so.

For H5, it was expected that students in the CTA group would demonstrate more complete feedback compared to both the Control group and the EP group. Additionally, the EP group will demonstrate more complete feedback as compared to the Control group. As stated previously, “more complete feedback” was operationally defined as feedback which provided more information to the student receiving the feedback and consisted of not only the CATME behaviors, but included at least two of the secondary codes (FE, FI, and FA) as well. It was expected that the survey and instructional manipulations would both have significant effects if this hypothesis is supported. As expected, both the EP and CTA groups demonstrated significantly more comments that used at least two secondary codes in their peer evaluations over the Control groups. However, there were no significant differences between the EP and CTA groups which indicates that there only appears to be a medium effect of the survey manipulation on the use of two secondary codes.

This changes when looking at the use of all three secondary codes in the comments. The differences between the control and EP group using all three secondary codes approached significance ($p = 0.079$) with students in the EP group more likely to use all three secondary codes. Additionally, there are statistically significant differences between the control and CTA groups with the CTA group more likely to use all three secondary codes in the midterm peer evaluation comments. This may indicate a small significant effect of the instructional manipulation in addition to the medium effect due to the survey manipulation in the increase in the use of this code in the CTA group.

In addition to analyzing the effects of the survey and instructional manipulations on the peer evaluation comments, I investigated the use of these comments in predicting team satisfaction using a self-reported question for all groups: “At this current time, rate how satisfied you are with the functionality of your team experience”. This question was rated on a scale of 0 (*not satisfied*) to 100 (*completely satisfied*). In general, the students rate their satisfaction very high, there are no significant differences between groups on this measure. There were significant differences over time in their team satisfaction, with students in the final peer evaluations reporting statistically significant increases in team satisfaction.

Looking at the correlations and regression analysis, I expected more comments regarding teaming behaviors to have a positive relationship with this satisfaction measure. However, there were no consistent measures between the different groups. For example, there is a negative correlation between Having Related Knowledge, Skills, and Abilities and Team Satisfaction for the peer-evaluations in the Control group only. This would seem to indicate that that more comments on this category results in lower Team Satisfaction or vis versa. The only other measures that showed significant relationships with Team Satisfaction was a positive correlation with Contributing to the Team’s Work and a negative correlation with actionable feedback for the self-evaluations of the Control group. Looking at the hierarchical linear regression yielded only the Contributing to the Team’s Work contributed significantly to the prediction of Team Satisfaction, but only described 5.6% of the variance in team satisfaction. It would appear that the number of comments on a given behavioral category do not lead a good prediction of team

satisfaction. However, there are a lot of variables left unexplored that could be evaluated in a future study. For example, the effects of the number of positive or negative comments on any given category on the overall team satisfaction.

7 Project Summary

7.1 Overall Results

There exists a need for revision to the current peer evaluation methods as many students are unable or unwilling to provide meaningful feedback or assessment of their own or teammates contributions on a peer evaluation (Brawner, Murch, Ferguson, and Ohland, 2018; Cheville and Duvall, 2008; Dudevior, Laffely, and Mundy, 2010; Harley, 2013; Kaufman, Felder, and Fuller, 2000; Richmond, Satyamurthy, and DeFranco, 2016; Topping, 2003). Because of these challenges, the overall goal for my dissertation was to improve the utility of student peer evaluation feedback. In a series of four studies, following the Design-Based Research (DBR) approach for educational research, I evaluated the effects of a) alterations in the peer evaluation instrument (CATME), b) changes to instructions for students regarding peer evaluation, and c) new instructional strategies on the utility of the peer feedback produced on midterm and end-of-term evaluations by first-year engineering students. My overall conclusions are summarized below.

The current model for peer evaluation in Michigan Tech's first-year engineering program uses the CATME Peer Evaluation tool and was assessed during Study 2. With the training provided by CATME, students do not appear to differentiate their teammate's contributions within a teaming dimension (e.g., they rated their teammates the same on a behavioral category such as Expecting Quality). Additionally, approximately 39% of students do not differentiate across teaming dimensions (e.g., they rate their teammates the same across all teaming dimensions). Finally, only 55% of students provided

feedback (of any kind) in the comment section of the CATME survey and only 6% of students provided negative or constructive feedback to their teammates.

These results are consistent with literature. Cheville and Duvall (2008) noticed similar behaviors in students enrolled in their capstone design courses. Students who gave high scores on one area of the peer evaluation were more likely to also give high scores on the remaining areas of the peer evaluation. This would indicate this is not only a problem for first-year engineering students, it also affects upper-division students. With respect to qualitative comments, a 2018 study by Brawner, Murch, Ferguson, and Ohland, only about a third of the students provided comments to themselves or their teammates and these comments lacked details that explained their quantitative rating of their peer (Brawner, Murch, Ferguson, and Ohland, 2018).

With the CATME peer evaluation in use by 22,000 instructors at 2,400 institutions in 88 countries (info.catme.org), increasing the utility of this peer evaluation tool can have a significant impact on how teaming is assessed globally. Additionally, teamwork is a cornerstone of many courses and essential in the working environment for most engineering students. This work significantly impacts the validity of using peer evaluation or team assignments as part of course grading schemas or the assessment of individual effort on team projects, which is a major component of collaborative learning (Boud, Cohen, and Sampson, 1999; Loughry, Ohland, and Woehr, 2014; Svinicki and McKeachie, 2014).

The third and fourth study explored both a survey and instructional manipulation to increase the utility and meet the need outlined by literature. The survey manipulation increased the visual size of the response boxes for peer comments and increased clarity and detail of the instructional prompt in the CATME survey (as administered through Survey Monkey). The instructional manipulation used a Cognitive Task Analysis (CTA) Knowledge Audit procedure with specific probes designed to focus on a specific area of a task (Crandall, Klein & Hoffman, 2006; Militello and Hutton, 1998) In this case, the students were asked to focus on a specific team project. The final element created for these studies was the development of a qualitative codebook to identify and quantify the recurring ideas and themes brought forward by the students in their self and peer evaluation comments. This codebook was used to quantify the results of the manipulations.

The survey manipulation showed a significant positive effect on several of the individual behavioral categories, the average number of behavioral categories, and the amount of actionable feedback discussed in the midterm peer evaluation comments. These results are consistent with the studies that indicate that increasing the visual size of the response box increases the length of response (Christian and Dillman, 2004; Smyth, Dillman, Christian, & McBride, 2009) and that increasing the clarity of instructions improves the thoroughness of responses (Smyth et al., 2009; Chaudhary & Israel, 2016).

While there were no individual effects of the instructional manipulation measured due to the study design, there were combined effects of the instructional and survey manipulations that resulted in significant differences between the control and CTA group

on the average number of behavioral categories discussed in the midterm peer evaluation comments. Additionally, the CTA group demonstrated greater utility of feedback (operationally defined as detailed, references specific team work, and providing actionable and/or impactful feedback) over the Control group. This may demonstrate that having the students focus on a specific instance of teamwork before completing their peer evaluations may help more specific examples about the teaming experience (Crandall, Klein, & Hoffman, 2006), and reduce memory errors linked with incorrect or biased responses (Cohen, Manion & Morrison, 2013). However, the study design would need to be modified to unpack the individual effects of the survey and instructional interventions.

7.2 Future Work

These studies focused on the use of the CATME peer evaluation tool by Michigan Technological University first-year engineering students as a way to assess their teaming experiences. It is essential that these student teams learn effective strategies for peer feedback or they will not function optimally. These modifications to the CATME survey and the CTA instructional manipulation have the potential to increase the utility of this popular peer evaluation tool. While the results of these studies with respect to the qualitative peer evaluation comments are promising, there is still work to be done.

To begin, the codebook was developed using an inductive, open coding approach to develop the primary and secondary codes used in this study (Hsieh & Shannon, 2005; Saldaña, 2016). When using repeating cycles of coding and evaluation on the same data set, the inter-rater reliability between myself and another faculty member of Engineering

Fundamentals resulted in very good strength of agreement on all measures. However, when used independently, the secondary codes for utility (FE, FI, FA) showed only moderate strength of agreement, which demonstrates the need for more refinement in the codebook to ensure reliability before a broad utilization of this tool is encouraged.

Next, the modifications to the CATME tool should be tested in situ. For this final study, I recreated the survey instrument in Survey Monkey to manipulate the prompts and size of comment areas. A real test of the effectiveness of these manipulations would be to perform a test using the CATME tool to verify the increase of utility shown in this study. Also, this final study design did not allow me to separate the individual effects of the instructional and survey manipulations. An additional study with a Control group for the instructional manipulation would help address this limitation. Once this is complete, the effects of the CTA reflective instructional method can be unpacked from the survey manipulation. Additionally, the changes in student behaviors providing comments over time should be unpacked. Is the decrease in behavioral comments or comments on actionable feedback due to the perceived lack of usefulness of these comments to their current team?

Final, this series of studies indicated that there is a lot of work that needs to be done to evaluate the quantitative measures of teaming. Many of the instruments I used to evaluate teaming effectiveness or satisfaction such as the trust and conflict scale (Lencioni, 2007) and the team satisfaction scales (Van der Vegt, Emans, and Van De Vliert, 2001) have a ceiling effect associated with the results. This could be due to a reluctance by the students to give their teammates lower ratings in formal evaluations

(Brutus, Donia, and Ronen, 2013; Van Tyne, 2017) or it could indicate students are unable to discriminate between teammates due to a lack of observation of that dimension given the teaming tasks they had engaged in. It may also be due to the information a student is using to rate each other (e.g., personal attributes, quality of work) (Cheville and Duvall, 2008). Whatever the issue, evaluating team effectiveness remains a rich area of research.

8 Reference List

1. Adams, J. M., & DeGraaff, M. D., & Hohner, G. S. (2015, June), *Evaluating the Pre-Professional Engineer: Exploring the Peer Review Process* Paper presented at 2015 ASEE Annual Conference & Exposition, Seattle, Washington. 10.18260/p.24029
2. Anderson, T., & Shattuck, J. (2012). Design-based research: A decade of progress in education research?. *Educational researcher*, 41(1), 16-25.
3. Baker, D. F. (2008). Peer assessment in small groups: A comparison of methods. *Journal of Management Education*, 32(2), 183-209.
4. Bandura, A., & Walters, R. H. (1977). *Social learning theory* (Vol. 1). Englewood Cliffs, NJ: Prentice-hall.
5. Bernardin, H. J., & Villanova, P. (2005). Research streams in rater self-efficacy. *Group & Organization Management*, 30(1), 61-88.
6. Blackwell, L., Trzesniewski, K., & Dweck, C.S. (2007). Implicit Theories of Intelligence Predict Achievement Across an Adolescent Transition: A Longitudinal Study and an Intervention. *Child Development*, 78, 246-263. doi: 10.1111/j.1467-8624.2007.00995.x
7. Bonk, C., & Cunningham, D. (1998). Searching for constructivist learner-centered and sociocultural components for collaborative educational learning tools in C Bonk and K King (eds) *Electronic Collaborators: Learner-Centered Technologies for Literacy Apprenticeship and Discourse. Electronic collaborators: Learner-centered technologies for literacy, apprenticeship, and discourse. New York: Erlbaum.*
8. Boud, D., Cohen, R., & Sampson, J. (1999). Peer learning and assessment. *Assessment & evaluation in higher education*, 24(4), 413-426.
9. Braender, L. M., & Naples, M. I. (2013). Evaluating the impact and determinants of student team performance: Using LMS and CATME data. *Journal of Information Systems Education*, 24(4), 281-289.
10. Brawner, C. E., & Murch, O. W., & Ferguson, D. M., & Ohland, M. W. (2018, June), *Comparing Peer-to-Peer Written Comments and Teamwork Peer Evaluations*. Paper presented at 2018 ASEE Annual Conference & Exposition , Salt Lake City, Utah. 10.18260/1-2—30206
11. Brooks, C. M., & 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), 268-272.
12. Brutus, S., & Donia, M. B. (2010). Improving the effectiveness of students in groups with a centralized peer evaluation system. *Academy of Management Learning & Education*, 9(4), 652-662.
13. Brutus, S., Donia, M. B., & Ronen, S. (2013). Can business students learn to evaluate better? Evidence from repeated exposure to a peer-evaluation system. *Academy of Management Learning & Education*, 12(1), 18-31.

14. Burnette, J. L., O'Boyle, E. H., VanEpps, E. M., Pollack, J. M., & Finkel, E. J. (2013). Mind-sets matter: A meta-analytic review of implicit theories and self-regulation. *Psychological bulletin*, 139(3), 655.
15. Chaudhary, A. K., & Israel, G. D. (2016). Influence of importance statements and box size on response rate and response quality of open-ended questions in web/mail mixed-mode surveys. *Journal of Rural Social Sciences*, 31(3), 140.
16. Cheville, A., & Duvall, J. (2008, June), *Evaluating Different Aspects Of Peer Interaction Using An On Line Instrument* Paper presented at 2008 Annual Conference & Exposition, Pittsburgh, Pennsylvania. 10.18260/1-2—3558
17. Christian, L. M., & Dillman, D. A. (2004). The influence of graphical and symbolic language manipulations on responses to self-administered questions. *Public Opinion Quarterly*, 68(1), 57-80.
18. Clabaugh, G. K. (2010). The educational theory of Lev Vygotsky: A multi-dimensional analysis. *New Foundations*, 1-18.
19. Cohen, L., Manion, L., & Morrison, K. (2013). *Research methods in education*. routledge.
20. Crandall, B., Klein, G., Klein, G. A., & Hoffman, R. R. (2006). Working minds: A practitioner's guide to cognitive task analysis. Mit Press.
21. decisionwise. (2020) Using the SARA Model to Learn from 360-Degree Feedback <https://decision-wise.com/using-the-sara-model-to-learn-from-360-degree-feedback/>
22. Druskat, V. U., & Wolff, S. B. (1999). Effects and timing of developmental peer appraisals in self-managing work groups. *Journal of Applied Psychology*, 84(1), 58.
23. Dudevoir, G., & Laffely, A., & Mundy, A. J. (2010, June), *"Survivor" Meets Senior Project* Paper presented at 2010 Annual Conference & Exposition, Louisville, Kentucky. 10.18260/1-2—16267
24. Educational Broadcasting Corporation (2010) Workshop: Cooperative and Collaborative Learning Accessed 17 February 2020.
25. Felder, R. M., & Brent, R. (2001). Effective strategies for cooperative learning. *Journal of Cooperation & Collaboration in College Teaching*, 10(2), 69-75.
26. Felder, R. M., & Brent, R. (2016). *Teaching and learning STEM: A practical guide*. John Wiley & Sons.
27. Flanagan, Stephanie. (2017) *Penn State Performance Management: Giving and Receiving Feedback*. <https://hr.psu.edu/sites/hr/files/GivingAndReceivingFeedbackIndividualContributor.pdf>
28. Gladstein, D. L. (1984). Groups in context: A model of task group effectiveness. *Administrative science quarterly*, 499-517.
29. Hackman, J. R. (1990). Groups that work and those that don't (No. E10 H123). Jossey-Bass.
30. Hamlin, A. J., Kemppainen, A. J., Thiemann, K., Singer, A. M., Manser, N. D., Hamlin, B. H., and M. E. Jarvie-Eggart. (July, 2021). *Engaging students in synchronous remote or hybrid first-year engineering courses* Paper presented at 2021 Annual Conference & Exposition, Long Beach, CA. *Draft Submitted*

31. Hamlin, J. L. (2008). *Team effectiveness: A validation study of Lencioni's five functions of a team*. University of La Verne.
32. Harley, S. (2013). *How to Say Anything to Anyone*. Austin, TX: Greenleaf Book Group, 117.
33. Harley, S. (2018). *How People Respond to Negative Feedback - The Freak Out*. https://www.youtube.com/watch?v=TCWc_vbu-uA&t=16s
34. Hsieh, H. F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative health research*, 15(9), 1277-1288.
35. Jamnia, A. (2018). *Introduction to product design and development for engineers*. CRC Press.
36. Kalton, G., & Schuman, H. (1982). The effect of the question on survey responses: A review. *Journal of the Royal Statistical Society: Series A (General)*, 145(1), 42-57.
37. Karau, S. J., & Williams, K. D. (1993). Social loafing: A meta-analytic review and theoretical integration. *Journal of personality and social psychology*, 65(4), 681.
38. Kaufman, D. B., Felder, R. M., & Fuller, H. (2000). Accounting for individual effort in cooperative learning teams. *Journal of Engineering Education*, 89(2), 133-140.
39. Kemppainen, A. J. (2018). Exploring how Reflective Exercises affect First-Year Engineering Epistemological and Intelligence Beliefs.
40. Kilpatrick, D. J., Linville, M., & Stout, D. E. (2001). Procedural justice and the development and use of peer evaluations in business and accounting classes. *Journal of Accounting Education*, 19(4), 225-246.
41. Klein, G. (2000). Cognitive Task Analysis of Teams. In J. M. Schraagen, S. F. Chipman, & V. L. Shalin (Eds.). (pp. 417-429). *Cognitive task analysis*. Psychology Press.
42. Kuwabara, M., Einbinder, S. D., Sun, R., & Azizi, R. (2020). Collaborative Learning Techniques, Student Learning Outcomes, and Equal Workload within Groups in Different Teaching Modalities. *International Journal of Teaching and Learning in Higher Education*, 32(2), 293-304.
43. Laerd Statistics (2015). Cohen's kappa using SPSS Statistics. Statistical tutorials and software guides. Retrieved from <https://statistics.laerd.com/>
44. Lencioni, P. (2007). *The five dysfunctions of a team*. John Wiley & Sons.
45. Lencioni, P. M. (2012). *The five dysfunctions of a team: Team assessment*. John Wiley & Sons.
46. Levy, S. R., & Dweck, C. S. (1997). Implicit theory measures: Reliability and validity data for adults and children. *Unpublished manuscript, Columbia University, New York*.
47. Levy, S. R., Stroessner, S. J., & Dweck, C. S. (1998). Stereotype formation and endorsement: The role of implicit theories. *Journal of Personality and Social Psychology*, 74(6), 1421.
48. Loughry, M. L., Ohland, M. W., & Woehr, D. J. (2014). Assessing teamwork skills for assurance of learning using CATME team tools. *Journal of Marketing Education*, 36(1), 5-19.

49. Merriam, S. B., & Tisdell, E. J. (2015). *Qualitative research: A guide to design and implementation*. John Wiley & Sons.
50. Militello, L. G., Hutton, R. J., Pliske, R. M., Knight, B. J., & Klein, G. (1997). *Applied cognitive task analysis (ACTA) methodology*. KLEIN ASSOCIATES INC FAIRBORN OH.
51. Militello, L. G., & Hutton, R. J. (1998). Applied cognitive task analysis (ACTA): a practitioner's toolkit for understanding cognitive task demands. *Ergonomics*, 41(11), 1618-1641.
52. Miller, E., & Hirshfield, L., & Chachra, D. (2016, June), *Classifying Dissatisfaction: Student Perspectives on Teammate Performance* Paper presented at 2016 ASEE Annual Conference & Exposition, New Orleans, Louisiana. 10.18260/p.26499
53. Ohland, M. W., & Loughry, M. L. (2006, October). Designing a Peer Evaluation Instrument that is Simple, Reliable, and Valid: The Comprehensive Assessment of Team-Member Effectiveness. In *STEM Assessment Conference* (p. 202-208).
54. Ohland, M. W., Loughry, M. L., Woehr, D. J., Bullard, L. G., Felder, R. M., Finelli, C. J., & 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 & Education*, 11(4), 609-630.
55. Pritchard, A., & Woollard, J. (2013). *Psychology for the classroom: Constructivism and social learning*. Routledge.
56. Richmond, S. S., & Satyamurthy, K., & DeFranco, J. F. (2016, June), *Exploring the Value of Peer Assessment* Paper presented at 2016 ASEE Annual Conference & Exposition, New Orleans, Louisiana. 10.18260/p.26872
57. Romero, C., Master, A., Paunesku, D., Dweck, C. S., & Gross, J. J. (2014, February 10) Academic and Emotional Functioning in Middle School: The Role of Implicit Theories. *Emotion*. Advance online publication. <http://dx.doi.org/10.1037/a0035490>
58. Saldaña, J. (2015). *The coding manual for qualitative researchers*. Sage.
59. Smith, B. L., & MacGregor, J. T. (1992). What is collaborative learning. *Towards the Virtual University: International Online Learning Perspectives*, 217-232.
60. Smyth, J. D., Dillman, D. A., Christian, L. M., & McBride, M. (2009). Open-ended questions in web surveys: Can increasing the size of answer boxes and providing extra verbal instructions improve response quality?. *Public Opinion Quarterly*, 73(2), 325-337.
61. Sorby, S. A., Monte, A. E., & Hein, G. L. (2001). Implementing a common first year engineering program at Michigan Tech. *age*, 6, 1.
62. Svinicki, M. D., & WJ, M. (2014). *McKeachie's teaching tips 14 th Ed. Wadsworth: Cengage Learning*.
63. Tekleab, A. G., Quigley, N. R., & Tesluk, P. E. (2009). A longitudinal study of team conflict, conflict management, cohesion, and team effectiveness. *Group & Organization Management*, 34(2), 170-205.
64. Topping, K. (2003). Self and peer assessment in school and university: Reliability, validity and utility. In *Optimising new modes of assessment: In search of qualities and standards* (pp. 55-87). Springer, Dordrecht.

65. Tuckman, B. W. (1965). Developmental sequence in small groups. *Psychological bulletin*, 63(6), 384.
66. U.S. Department of Education. (2017). *Collaborative Problem Solving: Considerations for the National Assessment of Educational Progress*. Retrieved from https://nces.ed.gov/nationsreportcard/pdf/researchcenter/collaborative_problem_solving.pdf
67. Van den Akker, J., Gravemeijer, K., McKenney, S., & Nieveen, N. (Eds.). (2006). *Educational design research*. Routledge.
68. Van der Vegt, G. S., Emans, B. J., & Van De Vliert, E. (2001). Patterns of interdependence in work teams: A two-level investigation of the relations with job and team satisfaction. *Personnel Psychology*, 54(1), 51-69.
69. Van Tyne, N. C. (2017, June), *Board # 77 : A Sea of Variations: Lessons Learned from Student Feedback about the Role of Trust in First-year Design Teams* Paper presented at 2017 ASEE Annual Conference & Exposition, Columbus, Ohio. 10.18260/1-2—27922
70. Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard university press.
71. Wilson, C. (2010). Bruce Tuckman's Forming, Storming, Norming & Performing Team Development Model. Retrieved 24 February 2020.
72. Wink, J., & Putney, L. G. (2002). *A vision of Vygotsky*. Allyn & Bacon.

A Identifying Teaming Issues

Table A.1. Faculty / LEAP Leader Issue Identification Survey

Please take 5 minutes and help identify which of the following negative behaviors, actions, or consequences do you see most frequently in first-year engineering teams. Please narrow your list to the top five most frequently occurring items that you have personally observed in the teams you have instructed.

First-year engineering teams most commonly exhibit the following negative behaviors, actions, or consequences: (choose your top 5)

Issue Description	Issue Selection
1. conceal their weaknesses and mistakes from one another	☐
2. hold grudges	☐
3. create environments where back-channel politics and personal attacks thrive	☐
4. fail to tap into all the opinions and perspectives of team members	☐
5. breeds lack of confidence and fear of failure	☐
6. revisits discussion and decisions again and again	☐
7. places an undue burden on the team leader as the sole source of discipline	☐
8. encourages mediocrity	☐
9. fail to recognize and tap into one another's skills and experiences	☐
10. is easily distracted	☐
11. waste time and energy managing their behaviors for effect	☐
12. misses deadlines and key deliverables	☐
13. creates resentment among team members who have different standards of performance	☐
14. hesitate to offer help outside their own areas of responsibility	☐
15. jump to conclusions about the intentions and aptitudes of others without attempting to clarify them	☐
16. dread meetings and find reasons to avoid spending time together	☐
17. encourages team members to focus on their own individual goals	☐
18. hesitate to ask for help or provide constructive feedback	☐
19. ignore controversial topics that are critical to team success	☐

Adapted from Lencioni, P. (2007). The five dysfunctions of a team. John Wiley & Sons.

B CATME Peer Evaluation Instrument

			Description of Rating
☐	☐	☐	<ul style="list-style-type: none"> Does more or higher-quality work than expected. Makes important contributions that improve the team's work. Helps teammates who are having difficulty completing their work.
☐	☐	☐	Demonstrates behaviors described immediately above and below.
☐	☐	☐	<ul style="list-style-type: none"> Completes a fair share of the team's work with acceptable quality. Keeps commitments and completes assignments on time. Helps teammates who are having difficulty when it is easy or important.
☐	☐	☐	Demonstrates behaviors described immediately above and below.
☐	☐	☐	<ul style="list-style-type: none"> Does not do a fair share of the team's work. Delivers sloppy or incomplete work. Misses deadlines. Is late, unprepared, or absent for team meetings. Does not assist teammates. Quits if the work becomes difficult.

Figure B.1. CATME Rating: Contributing to Team's Work (Ohland et al., 2012)

			Description of Rating
☐	☐	☐	<ul style="list-style-type: none"> Asks for and shows an interest in teammates' ideas and contributions. Makes sure teammates stay informed and understand each other. Provides encouragement or enthusiasm to the team. Asks teammates for feedback and uses their suggestions to improve.
☐	☐	☐	Demonstrates behaviors described immediately above and below.
☐	☐	☐	<ul style="list-style-type: none"> Listens to teammates and respects their contributions. Communicates clearly. Shares information with teammates. Participates fully in team activities. Respects and responds to feedback from teammates.
☐	☐	☐	Demonstrates behaviors described immediately above and below.
☐	☐	☐	<ul style="list-style-type: none"> Interrupts, ignores, bosses, or makes fun of teammates. Takes actions that affect teammates without their input. Does not share information. Complains, makes excuses, or does not interact with teammates. Is defensive. Will not accept help or advice from teammates.

Figure B.2. CATME Rating: Interacting with Teammates (Ohland et al., 2012)

			Description of Rating
☐	☐	☐	<ul style="list-style-type: none"> • Watches conditions affecting the team and monitors the team's progress. • Makes sure that teammates are making appropriate progress. • Gives teammates specific, timely, and constructive feedback.
☐	☐	☐	Demonstrates behaviors described immediately above and below.
☐	☐	☐	<ul style="list-style-type: none"> • Notices changes that influence the team's success. • Knows what everyone on the team should be doing and notices problems. • Alerts teammates or suggests solutions when the team's success is threatened.
☐	☐	☐	Demonstrates behaviors described immediately above and below.
☐	☐	☐	<ul style="list-style-type: none"> • Is unaware of whether the team is meeting its goals. • Does not pay attention to teammates' progress. • Avoids discussing team problems, even when they are obvious.

Figure B.3. CATME Rating: Keeping the Team on Track (Ohland et al., 2012)

			Description of Rating
☐	☐	☐	<ul style="list-style-type: none"> • Motivates the team to do excellent work. • Cares that the team does outstanding work, even if there is no additional reward. • Believes that the team can do excellent work.
☐	☐	☐	Demonstrates behaviors described immediately above and below.
☐	☐	☐	<ul style="list-style-type: none"> • Encourages the team to do good work that meets all requirements. • Wants the team to perform well enough to earn all available rewards. • Believes that the team can fully meet its responsibilities.
☐	☐	☐	Demonstrates behaviors described immediately above and below.
☐	☐	☐	<ul style="list-style-type: none"> • Satisfied even if the team does not meet assigned standards. • Wants the team to avoid work, even if it hurts the team. • Doubts that the team can meet its requirements.

Figure B.4. CATME Rating: Expecting Quality (Ohland et al., 2012)

			Description of Rating
☐	☐	☐	<ul style="list-style-type: none"> • Demonstrates the knowledge, skills, and abilities to do excellent work. • Acquires new knowledge or skills to improve the team's performance. • Able to perform the role of any team member if necessary.
☐	☐	☐	Demonstrates behaviors described immediately above and below.
☐	☐	☐	<ul style="list-style-type: none"> • Demonstrates sufficient knowledge, skills, and abilities to contribute to the team's work. • Acquires knowledge or skills as needed to meet requirements. • Able to perform some of the tasks normally done by other team members.
☐	☐	☐	Demonstrates behaviors described immediately above and below.
☐	☐	☐	<ul style="list-style-type: none"> • Missing basic qualifications needed to be a member of the team. • Unable or unwilling to develop knowledge or skills to contribute to the team. • Unable to perform any of the duties of other team members.

Figure B.5. CATME Rating: Having Related Knowledge, Skills, and Abilities (Ohland et al., 2012)

C Team Evaluation Follow-up Questions

Table C.1. Team Satisfaction follow-up Questions

Place a check in the column that identifies the extent to which you agree or disagree with the statement. 1 = Strongly Disagree, 5 = Strongly Agree

Survey Items	1	2	3	4	5
1. I am satisfied with my present teammates.					
2. I am pleased with the way my teammates and I work together.					
3. I am very satisfied with working in this team.					

Adapted from Van der Vegt, G. S., Emans, B. J., & Van De Vliert, E. (2001). Patterns of interdependence in work teams: A two-level investigation of the relations with job and team satisfaction. *Personnel Psychology*, 54(1), 51-69.

Table C.2. Team Interdependence follow-up Questions

Place a check in the column that identifies the extent to which you agree or disagree with the statement. 1 = Strongly Disagree, 5 = Strongly Agree

Survey Items	1	2	3	4	5
1. My teammates and I have to obtain information and advice from one another in order to complete our work					
2. I depend on my teammates for the completion of my work					
3. I have a one-person job; I rarely have to check or work with others (scale reversed)					
4. I have to work closely with my teammates to do my work properly					
5. In order to complete our work, my teammates and I have to collaborate extensively					

Adapted from Van der Vegt, G. S., Emans, B. J., & Van De Vliert, E. (2001). Patterns of interdependence in work teams: A two-level investigation of the relations with job and team satisfaction. *Personnel Psychology*, 54(1), 51-69.

D Assessing Team Trust and Conflict

Table D.1. Team Reflection Survey

Assign a rating to each statement. It is essential to the accurate scoring of this instrument. Evaluate the statements honestly and be as objective as possible. Be thoughtful about your responses, but don't agonize over each response. Your initial "gut feel" is usually best. Place a check in the column that most closely applies to your team. 1 = Never, 5 = Always.

Survey Items	1	2	3	4	5
Trust					
1. Team members admit their mistakes.					
2. Team members acknowledge their weaknesses to one another.					
3. Team members ask for help without hesitation.					
4. Team members ask one another for input regarding their areas of responsibility.					
5. Team members acknowledge and tap into one another's skills and expertise.					
6. Team members willingly apologize to one another.					
7. Team members are unguarded and genuine with one another.					
8. Team members can comfortably discuss their personal lives with one another.					
Conflict					
1. Team members are passionate and unguarded in their discussion of issues.					
2. Team meetings are interesting and compelling (not boring).					
3. During team meetings, the most important - and difficult - issues are described.					
4. Team members voice their opinions even at the risk of causing disagreement.					
5. During discussions, team members challenge one another about how they arrived at their conclusions and opinions.					
6. Team members solicit one another's opinions during meetings.					
7. Team members communicate unpopular opinions to the group.					
8. When conflict occurs, the team confronts and deals with the issue before moving to another subject.					

Adapted from Lencioni, P. (2007). The five dysfunctions of a team. John Wiley & Sons.

Table D.2. Reflection Scoring Interpretation

Scales	High	Medium	Low
	Average score of 3.75 and above	Average score of 3.25 -3.74	Average score of 3.24 and below
Trust	Your team has created an environment where vulnerability and openness are the norm.	Your team may need to get more comfortable being vulnerable and open with one another about individual strengths, weaknesses, mistakes, and needs for help.	Your team lacks necessary levels of openness and vulnerability about individual strengths, weaknesses, mistakes, and needs for help.
Conflict	Your team is comfortable engaging in unfiltered discussion around important topics.	Your team may need to learn to engage in more unfiltered discussion around important topics.	Your team is not comfortable engaging in unfiltered discussion around important topics.

Adapted from Lencioni, P. M. (2012). *The five dysfunctions of a team: Team assessment*. John Wiley & Sons.

E Assessing Teamwork with Adapted Surveys

Fall 2020 Midterm Peer Evaluation

Your goal is to answer each question as accurately as possible. Please complete the evaluation in private and do not discuss your answers with anyone else.

- * 1. Enter your team information and names of your teammates (First and Last). Note, if you do not have the number of teammates listed, just put NA in the space for their name and do not rate those team members in the tables below.

Section and Team Number	<input type="text"/>
Your name	<input type="text"/>
Teammate 1	<input type="text"/>
Teammate 2	<input type="text"/>
Teammate 3	<input type="text"/>
Teammate 4	<input type="text"/>

Each row in the following tables show a list of behaviors, while each column lists teammates whom you've identified above. For example, if teammate 1 is identified as Sally above, be sure to rate Sally everywhere it says teammate 1 below.

For each teammate column, select one row that best describes that individual's behaviors during the time period evaluated.

- * 2. Contributing to the Team

Rate each of your teammates (including yourself) by clicking the row that best describes how that team member behaved during the period being evaluated.

	Teammate Yourself	Teammate 1	Teammate 2	Teammate 3	Teammate 4
Does more or higher-quality work than expected.	☐	☐	☐	☐	☐
Makes important contributions that improve the team's work.	☐	☐	☐	☐	☐
Helps teammates who are having difficulty completing their work.	☐	☐	☐	☐	☐
Demonstrates behaviors described immediately above and below.	☐	☐	☐	☐	☐
Completes a fair share of the team's work with acceptable quality.	☐	☐	☐	☐	☐
Keeps commitments and completes assignments on time.	☐	☐	☐	☐	☐
Helps teammates who are having difficulty when it is easy or important.	☐	☐	☐	☐	☐
Demonstrates behaviors described immediately above and below.	☐	☐	☐	☐	☐
Does not do a fair share of the team's work. Delivers sloppy or incomplete work.	☐	☐	☐	☐	☐
Misses deadlines. Is late, unprepared, or absent for team meetings.	☐	☐	☐	☐	☐
Does not assist teammates. Quits if the work becomes difficult.	☐	☐	☐	☐	☐

* 3. Interacting with Teammates

Rate each of your teammates (including yourself) by clicking the row that best describes how that team member behaved during the period being evaluated.

	Teammate Yourself	Teammate 1	Teammate 2	Teammate 3	Teammate 4
Asks for and shows an interest in teammates' ideas and contributions.	☐	☐	☐	☐	☐
Makes sure teammates stay informed and understand each other.	☐	☐	☐	☐	☐
Provides encouragement or enthusiasm to the team.	☐	☐	☐	☐	☐
Asks teammates for feedback and uses their suggestions to improve	☐	☐	☐	☐	☐
Demonstrates behaviors described immediately above and below.	☐	☐	☐	☐	☐
Listens to teammates and respects their contributions.	☐	☐	☐	☐	☐
Communicates clearly. Shares information with teammates.	☐	☐	☐	☐	☐
Participates fully in team activities.	☐	☐	☐	☐	☐
Respects and responds to feedback from teammates.	☐	☐	☐	☐	☐
Demonstrates behaviors described immediately above and below.	☐	☐	☐	☐	☐
Interrupts, ignores, bosses, or makes fun of teammates.	☐	☐	☐	☐	☐
Takes actions that affect teammates without their input. Does not share information.	☐	☐	☐	☐	☐
Complains, makes excuses, or does not interact with teammates.	☐	☐	☐	☐	☐
Is defensive. Will not accept help or advice from teammates.	☐	☐	☐	☐	☐

* 4. Keeping the Team on Track

Rate each of your teammates (including yourself) by clicking the row that best describes how that team member behaved during the period being evaluated.

	Teammate Yourself	Teammate 1	Teammate 2	Teammate 3	Teammate 4
Watches conditions affecting the team and monitors the team's progress.	☐	☐	☐	☐	☐
Makes sure that teammates are making appropriate progress.	☐	☐	☐	☐	☐
Gives teammates specific, timely, and constructive feedback.	☐	☐	☐	☐	☐
Demonstrates behaviors described immediately above and below.	☐	☐	☐	☐	☐
Notifies changes that influence the team's success.	☐	☐	☐	☐	☐
Knows what everyone on the team should be doing and notices problems.	☐	☐	☐	☐	☐
Alerts teammates or suggests solutions when the team's success is threatened.	☐	☐	☐	☐	☐
Demonstrates behaviors described immediately above and below.	☐	☐	☐	☐	☐
Is unaware of whether the team is meeting its goals.	☐	☐	☐	☐	☐
Does not pay attention to teammates' progress.	☐	☐	☐	☐	☐
Avoids discussing team problems, even when they are obvious.	☐	☐	☐	☐	☐

*** 5. Expecting Quality**

Rate each of your teammates (including yourself) by clicking the row that best describes how that team member behaved during the period being evaluated.

	Yourself	Teammate 1	Teammate 2	Teammate 3	Teammate 4
Motivates the team to do excellent work.	☐	☐	☐	☐	☐
Cares that the team does outstanding work, even if there is no additional reward.	☐	☐	☐	☐	☐
Believes that the team can do excellent work.	☐	☐	☐	☐	☐
Demonstrates behaviors described immediately above and below.	☐	☐	☐	☐	☐
Encourages the team to do good work that meets all requirements.	☐	☐	☐	☐	☐
Wants the team to perform well enough to earn all available rewards.	☐	☐	☐	☐	☐
Believes that the team can fully meet its responsibilities.	☐	☐	☐	☐	☐
Demonstrates behaviors described immediately above and below.	☐	☐	☐	☐	☐
Satisfied even if the team does not meet assigned standards.	☐	☐	☐	☐	☐
Wants the team to avoid work, even if it hurts the team.	☐	☐	☐	☐	☐
Doubts that the team can meet its requirements.	☐	☐	☐	☐	☐

*** 6. Having Related Knowledge, Skills, and Abilities**

Rate each of your teammates (including yourself) by clicking the row that best describes how that team member behaved during the period being evaluated.

	Yourself	Teammate 1	Teammate 2	Teammate 3	Teammate 4
Demonstrates the knowledge, skills, and abilities to do excellent work.	☐	☐	☐	☐	☐
Acquires new knowledge or skills to improve the team's performance.	☐	☐	☐	☐	☐
Able to perform the role of any team member if necessary.	☐	☐	☐	☐	☐
Demonstrates behaviors described immediately above and below.	☐	☐	☐	☐	☐
Demonstrates sufficient knowledge, skills, and abilities to contribute to the team's work.	☐	☐	☐	☐	☐
Acquires knowledge or skills as needed to meet requirements.	☐	☐	☐	☐	☐
Able to perform some of the tasks normally done by other team members.	☐	☐	☐	☐	☐
Demonstrates behaviors described immediately above and below.	☐	☐	☐	☐	☐
Missing basic qualifications needed to be a member of the team.	☐	☐	☐	☐	☐
Unable or unwilling to develop knowledge or skills to contribute to the team.	☐	☐	☐	☐	☐
Unable to perform any of the duties of other team members.	☐	☐	☐	☐	☐

Figure E.1. Instructions and CATME Behavior Scale Questions for Fall 2020 Peer Evaluation Surveys

Fall 2020 Midterm Peer Evaluation 1

Peer to Peer Comments

Please provide constructive comments about yourself and your fellow teammates (as identified on the previous page; ex. Sally = Teammate 1). Note if you do not have 4 teammates, just put an NA in those comment boxes.

The purpose of these comments is to give you the opportunity to explain how you rated your peers and if there was behavior or experiences in particular that influenced you when doing your peer and self-evaluations.

Note, this survey is currently set up so that your teammates will read your comments but will not know that it was you that wrote them.

* 7. Comments for Yourself

* 8. Comments for Teammate 1 (Enter teammate 1 name and comments here)

* 9. Comments for Teammate 2 (Enter teammate 2 name and comments here)

* 10. Comments for Teammate 3 (Enter teammate 3 name and comments here)

* 11. Comments for Teammate 4 (Enter teammate 4 name and comments here)

Figure E.2. Instructions and Peer Evaluation Prompts for Control Group

Fall 2020 Midterm Peer Evaluation 2

Peer to Peer Comments

Please provide constructive comments about yourself and your fellow teammates (as identified on the previous page; ex. Sally = Teammate 1). Note if you do not have 4 teammates, just put an NA in those comment boxes.

The purpose of these comments is to give you the opportunity to explain how you rated your peers and if there was behavior or experiences in particular that influenced you when doing your peer and self-evaluations. Please provide detailed, actionable, and constructive comments about your fellow teammates as well as yourself. Detailed and actionable feedback include:

- focusing on specific behaviors, not personality (can be behaviors you would like to see more of or one's you would like to see stopped)
- providing examples of these specific behaviors,
- specifying how these behaviors affect your team or your team's work, and
- suggesting actions or behaviors that should be added, stopped, altered, or maintained as is.

Be sure to include comments on the five CATME behavioral categories: Contributing to Team's Work, Interacting with Teammates, Expecting Quality, Keeping the team on Track, and Having Related Knowledge, Skills, and Abilities.

Note, this survey is currently set up so that your teammates will read your comments but will not know that it was you that wrote them.

* 7. Comments for Yourself

* 8. Comments for Teammate 1 (Enter teammate 1 name and comments here)

* 9. Comments for Teammate 2 (Enter teammate 2 name and comments here)

* 10. Comments for Teammate 3 (Enter teammate 3 name and comments here)

* 11. Comments for Teammate 4 (Enter teammate 4 name and comments here)

Figure E.3. Instructions and Peer Evaluation Prompts for Enhanced and CTA Groups

* 12. Confidential comments to the instructor

* 13. At this current time, rate how satisfied with the functionality of your team experience

0 - Not Satisfied at All
100 - Completely Satisfied

☐

Study Overview

Your instructor is studying how to improve the student team experience. We have made some changes to how we assess teaming, but we need to find out if they are really working.

This peer evaluation, administered now and at the end of the term, can help us measure how student's evaluate their team experience. By consenting to allow us to store and use your anonymized team evaluation answers as part of our study, you can help us to improve how students work together in teams.

Please read through the consent form below and indicate whether or not you consent to allow us to use your answers in our study. If you would like a copy, print this page or ask your instructor to send it to you for your records. Again, all identifying information will be removed from our data sets. You will do the same amount of work whether or not you consent. By consenting, you are allowing us to use your anonymized answers in order to make this course better.

* 14. Select one of the boxes below to indicate whether you consent to the study or not.

☐ I understand the procedures described below. My questions have been answered to my satisfaction, and I confirm that I am at least age 18 years or older and I agree to participate in this study.

☐ I understand the procedures described below. My questions have been answered to my satisfaction, and I confirm that I am at least age 18 years or older and I do not agree to participate in this study.

* 15. Select one of the boxes below to indicate whether you allow the researcher access to your FERPA data.

☐ I give permission for the researchers to access FERPA regulated private information (ACT/SAT scores, high school GPA, etc).

☐ I do not give permission for the researchers to access FERPA regulated private information (ACT/SAT scores, high school GPA, etc).

Figure E.4. End of Peer Evaluation Survey and Consent Information

F Peer Evaluation Instructional Video Transcript

This video will demonstrate how to take your peer evaluation survey

Peer evaluation surveys help your instructor assess team member effectiveness.

They also provide you with feedback from your teammates on your teamwork behavior.

The peer evaluation survey asks you to rate your team members on five different behavioral criteria, with an additional screen for you to provide comments to your teammates and to your instructor. Note that if you leave the survey in the process of completing it, Survey Monkey will not save your responses, and you must restart.

Be sure to read all instructions carefully.

To start, enter your name and the names of your teammates into the survey.

Next, you will be asked to rate your teammates against different criteria for each of the 5 dimensions of teaming.

For each category, the behaviors are organized into five different columns. Please answer these questions to the best of your ability by selecting the column that best describes that individual's behavior during this time period.

For example, if teammate one

- Completes a fair share of the team's work with acceptable quality.
- Keeps commitments and completes assignments on time.
- Helps teammates who are having difficulty when it is easy or important.

You would rate them by checking this circle. If teammate 2 demonstrates those behaviors, but also some from the behavioral definition, you would rate them by clicking this column.

You should have one column selected for each teammate you have.

If you do not have a teammate, enter N/A.

Be sure to rate yourself and your teammates on all five of these teaming dimensions. When you have finished, click next on the survey.

This page provides spaces to make comments about yourself and your teammates. Be sure to read all instructions carefully.

With constructive feedback, each teammate can build on their strengths and improve in areas where they fall short, making the team more effective and enjoyable.

Remember as you are writing your comments to keep following guidelines in mind:

1. Be respectful – You shouldn't say anything in the written comments that you would not say to the person's face.
2. Be real – Your comments should be honest and constructive.
3. Be trustworthy – Do not violate the trust of your teammates and instructors by publicizing your comments about others or theirs about you.
4. Make sure your comments only relate to teammates' behavior and contribution to the team. You should not make comments related to gender, ethnicity, national origin, appearance, or anything else unrelated to the team's work.
5. Your comments should be yours alone.

Each comment box needs to be filled out before you are able to move on.

Your comments will be seen by your instructor and will be released to your teammates anonymously.

There is also a comment box to add confidential comments to your instructor. Only you and your instructor will have access to these comments.

The last question on this page is a rating on how satisfied you are with your team's function. Use the slider to rate your team on a scale from 0 to 100.

When you have completed this page click the next button.

The final page provides a summary of a research project being done to improve the student team experience. We have made some changes to how we assess teaming, but we need to find out if they are really working.

This peer evaluation, administered now and at the end of the term, can help us measure how students evaluate their team experience. By consenting to allow us to store and use your anonymized team evaluation answers as part of our study, you can help us to improve how students work together in teams.

Please read through the consent form below and indicate whether or not you consent to allow us to use your answers in our study. If you would like a copy, print this page or ask your instructor to send it to you for your records. Again, all identifying information will be removed from our data sets. You will do the same amount of work whether or not you consent. By consenting, you are allowing us to use your anonymized answers in order to make this course better.

After reading the consent form, answer questions 14 and 15 to indicate your willingness to participate in this research.

Click the "done," button to complete the survey.

G Instructional Activities – Fall 2020

Name: _____

Team: ____ Section: ____

We're right in the middle of midterm exams and it is good to reflect on the types of study skills that can help you be successful. For this reflection, I've provided you with space to list out any exams you've had in the past week and upcoming exams. First, I'd like you to reflect on any exams you've had in the past 7 days. Complete Table 1 below to add the course, exam date, topics and how you studied for the exam. Go ahead and complete the table; add or delete rows as necessary. After completing the table, answer the reflection question below.

Table 1. Past Exam Reflection (last seven days)

Course	Exam Date	Topics Covered	Study Strategies

Past Exam Reflection Question

Looking at your study habits and strategies and how the exam went, in a couple of sentences discuss what (if any) changes will you make to your study strategies for these courses in the future.

Next, I'd like you look forward to the exams you have in the next 7 days. Complete Table 2 and identify the classes, exam dates, topics they cover, and study strategies you feel will be most effective for this/these course(s). Go ahead and complete the table; add or delete rows as necessary. Once you have completed the table, go ahead and submit your assignment. After completing the table, answer the reflection question below.

Table 2. Future Exam Reflection(last seven days)

Course	Exam Date	Topics Covered	Study Strategies

Future Exam Reflection Question

Based on your experiences with your past exams and your planned study strategies, write a couple of sentences discussing your study plans for these future exams.

Figure G.1. Study Skills Reflection Assignment

Name: _____ Team: ____ Section: ____

For this reflection, I've provided you with a list of all the main tasks for this project. You are going to reflect on your experience with your team in the completion of this project. You will use this information to complete your peer evaluations. Go ahead and complete the table for this project; add or delete rows as necessary. Once you have completed the table, answer the reflection questions at the end and submit your assignment. You will use this information when you complete the Peer Evaluation form.

Table 1. Steel Project Task Reflection

Main Task and Designation	Steps to Complete Task (e.g., team meetings, individual work)	Who participated?
Steel Properties and Treatment Video (Team)		
Peer Review of Videos (Individual)		
Steel Material Properties Spreadsheet (Individual)		
Plotting Steel Composition and Treatment Response (Individual)		
Heat Treatment MATLAB Analysis (Team)		
Heat Treatment Report (Team)		

Looking at the information from the table above, answer the following questions.

1. **Can you describe an instance when you knew something was amiss with your team?** Be specific and paint a picture with words (e.g., what happened? who was involved?)
2. **Can you think of an example when working on this project when you noticed an opportunity for you or your teammates to do something better?** Be specific and paint a picture with words (e.g., what happened? who was involved?)
3. **Can you think of a time when you realized that you would need to change the way you or another team member were performing to get the job done?** Be specific and paint a picture with words (e.g., what happened? who was involved?)

Figure G.2. CTA Teaming Reflection

H Copyright Documentation

Michigan Tech

Amber Kemppainen <amber@mtu.edu>

Use of Teaming Graphic in Dissertation

2 messages

Amber Kemppainen <amber@mtu.edu>

Fri, Nov 13, 2020 at 9:12 PM

To: info@coachingcultureatwork.com

Cc: "Amber" <amber@mtu.edu>

Hello,

My name is Amber Kemppainen from Michigan Technological University. I am finishing my dissertation on peer evaluation methods in teaming and would like to use the teaming graphic (shown below) of Tuckman's stages of teaming in your article "Bruce Tuckman's Forming, Storming, Norming & Performing Team Development Model" in my Dissertation.

Can I have your permission to use this image in my work? Please let me know! Thank you.

Amber

--

Ms. Amber Kemppainen (she/her/hers)
Principal Lecturer, Engineering Fundamentals
co-Director LEAP Program and EF Learning Center
Michigan Technological University
112D Dillman Hall
1400 Townsend Dr.
Houghton, MI 49931
<https://www.facebook.com/MTUEngineeringFundamentals>

Carol Wilson <carolwilson@coachingcultureatwork.com>
To: Amber Kemppainen <amber@mtu.edu>, Info <info@coachingcultureatwork.com>

Sat, Nov 14, 2020 at 2:31 AM

Hi Amber,

You are welcome. Here is another version which you can use if you prefer, and a link to the latest version of the article:
<http://www.coachingcultureatwork.com/bruce-tuckmans-forming-storming-norming-performing-team-development-model/>

Good luck with your dissertation!

Kind regards

Carol

Carol Wilson

CEO, Fellow of ILM, Association for Coaching & Professional Speaking Association

Mobile: +44 7730 006244

Author "Performance Coaching: A Complete Guide to Best Practice Coaching and Training"

Figure H.1. Permission Documentation for Model of Teaming (Wilson, 2010)

Marketplace™

John Wiley & Sons - Books - License Terms and Conditions

This is a License Agreement between Amber Kemppainen ("You") and John Wiley & Sons - Books ("Publisher") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by John Wiley & Sons - Books, and the CCC terms and conditions.

All payments must be made in full to CCC.

Order Date	03-Jan-2021	Type of Use	Republish in continuing education materials
Order license ID	1087633-1	Publisher	JOHN WILEY & SONS, INCORPORATED
ISBN-13	9781555421878	Portion	Chart/graph/table/figure

LICENSED CONTENT

Publication Title	Groups that work (and those that don't) : creating conditions for effective teamwork	Country	United States of America
Author/Editor	Hackman, J. Richard.	Rightholder	John Wiley & Sons - Books
Date	01/01/1989	Publication Type	Book
Language	English		

REQUEST DETAILS

Portion Type	Chart/graph/table/figure	Distribution	Worldwide
Number of charts / graphs / tables / figures requested	1	Translation	Original language of publication
Format (select all that apply)	Print, Electronic	Copies for the disabled?	Yes
Who will republish the content?	Academic institution	Minor editing privileges?	Yes
Duration of Use	Life of current edition	Incidental promotional use?	No
Lifetime Unit Quantity	Up to 4,999	Currency	USD
Rights Requested	Main product		

NEW WORK DETAILS

Title	Enhancing Peer Feedback Strategies to Improve Team Effectiveness	Publisher	Michigan Technological University / ProQuest
Author	Amber Kemppainen	Expected presentation date	2021-01-15

ADDITIONAL DETAILS

Order reference number	N/A	The requesting person / organization to appear on the license	Amber Kemppainen
------------------------	-----	---	------------------

REUSE CONTENT DETAILS

Title, description or numeric reference of the portion(s)	Table 1. Points of Leverage for Creating Conditions That Enhance Group Task Performance	Title of the article/chapter the portion is from	Introduction
Editor of portion(s)	Hackman, J. Richard.	Author of portion(s)	Hackman, J. Richard.
Volume of serial or monograph	N/A	Issue, if republishing an article from a serial	N/A
Page or page range of portion	13	Publication date of portion	2020-11-13

Figure H.2. Permission Documentation for Points of Leverage for Creating Conditions that Enhance Group Task Performance (Tuckman, 1990)

Figure H.3. Permission Documentation for SARA model for Receiving Feedback (decisionwise, 2020)