2023 Annual Conference & Exposition
Baltimore Convention Center, MD | June 25 - 28, 2023
The Harbor of Engineering Education for 130 Years
Paper ID #38551

Board 310: Impact of Student/Team Characteristics on Design Project Outcomes in Senior Design Courses

Mr. Hrushikesh Godbole, Rochester Institute of Technology

Hrushi Godbole holds an undergraduate degree in Production Engineering and a masters degree in Industrial and Systems Engineering. He is currently a graduate student at Rochester Institute of Technology pursuing a PhD in Mechanical and Industrial Engineering. Prior to starting the PhD program, he has gained five years of industry experience developing new products in the smart lighting industry performing various roles including product management, engineering and operations. His research interests include systems engineering, product design process and knowledge management in development teams.

Dr. Elizabeth A. Debartolo, Rochester Institute of Technology

Elizabeth A. DeBartolo, PhD is the Director of the Multidisciplinary Senior Design Program at the Rochester Institute of Technology, where students from Biomedical, Computer, Electrical, Industrial, and Mechanical Engineering work together on multidisciplinary projects. She is active in the national Capstone Design Community, and received her BSE in Mechanical Engineering and Materials Science from Duke University and her MS and PhD from Purdue University.

Dr. Shun Takai, Northern Illinois University
Marcos Esterman Jr., Rochester Institute of Technology

Impact of Student/Team Characteristics on Design Team Outcomes

Abstract

The ability to effectively work in teams is one of the desired outcomes of engineering and engineering technology programs. Unfortunately, working in teams is still challenging for many students. Social loafing, a tendency to work less when part of a team than when working individually, tends to destroy both teamwork performance and individual learning, especially in solving ill-structured problems, such as design. Furthermore, a bad experience on a past team is a significant concern as it could generate negative feelings about future team projects. The formation of collaborative teams is a critical first step in team-project-based design courses as team composition directly affects not only teamwork processes and outcomes, but also teamwork skills and experience.

This NSF sponsored project aims to enhance students' teamwork experiences and teamwork learning through 1) understanding how to form better student design teams and 2) identifying exercises that will effectively improve team member collaboration. We do this by comparing student team characteristics and design task characteristics with the quality of the design team outcome and examining the resulting correlations. Student characteristics cover six categories: 1) background information, 2) work structure preferences, 3) personality, 4) ability, 5) motivation, and 6) attitude. Task characteristics and design team outcomes are characterized using the Creative Product Semantic Scale (CPSS) and the Design Quality Rubric (DQR).

In this article, we present correlations between a subset of student team characteristics and task characteristics with design team outcomes for 2020-2021 senior design teams at Northern Illinois University and the Rochester Institute of Technology.

Introduction and Motivation

Teams are ubiquitous in today's work environment. Acting through the interdependent actions of individuals, teams embody the phrase "the whole is greater than the sum of its parts." The development of most contemporary products and services requires collaboration between individuals from various disciplines. Teams enable leveraging economies of scale and specialization, which can improve efficiency and performance of work output. Academicians from various disciplines including engineering and social sciences have continually improved their understanding of teams [1]–[6]. The recent leaps made in agile development highlight that even industry has realized the importance of effective teams and is striving to improve team processes [7].

Recognizing the importance of being able to work effectively in teams, the Accreditation Board for Engineering and Technology (ABET) requires engineering programs to provide students with exposure to teamwork as part of the undergraduate engineering curriculum [8]. As a result, almost all engineering programs in the US implement team based design project capstone courses [9], which provide over 100,000 engineering students with the opportunity to practice teamwork skills.

However, student experiences of working in teams are a mixed bag. A survey of students from various disciplines, including engineering, has indicated that 27% of students were unsatisfied with their teams and the division of tasks among the team members, and 32% of students experienced poor or very poor group work [10]. Especially common in design problems that are inherently ill-structured, social loafing tends to negatively affect team performance as well as individual learning [11]. To make matters worse, a bad experience on a past team project increases chances of negative feelings toward future team projects [12], [13].

In an effort to improve understanding of student design teams, team effectiveness models and instruments to measure inputs, processes, and outputs have been comprehensively reviewed, leading to a generalized design team effectiveness model [14]. This research work operationalizes a subset of the proposed general model in the context of undergraduate engineering capstone design courses, in order to better understand the correlations between project inputs (team characteristics, task characteristics) and outputs (design team outcome). In particular, this paper investigates the following research questions:
RQ1 - Which project task characteristics correlate with the design team outcome characteristics?
RQ2 - Which student team characteristics correlate with design team outcome characteristics?

Subjects

Student and design team characteristic data were collected in Fall 2020, when Northern Illinois University (NIU) was operating remotely and the Rochester Institute of Technology (RIT) was offering partial in-person classes on campus. Table 1 summarizes the students' disciplines and Table 2 summarizes the types of design teams in the senior design courses at NIU and RIT.

Table 1: Student participant distribution by engineering discipline

	NIU	RIT
Overall	187	331
Mechanical	113	124
Electrical	65	71
Biomedical	9	44
Industrial	-	45
Computer	-	47

Table 2: Summary of Team Demographics by University

Team Type	NIU	RIT
Overall	61 teams	64 teams
Multi-disciplinary	30	63
Single-disciplinary	31	1

Instruments

Task Characteristics - The CPSS was used to evaluate the initial problem statement provided to each team at the start of the course. CPSS measures artifacts representing the product, such as ideas, proposals, processes, prototypes, or tangible product itself, along three dimensions: novelty, resolution, and elaboration & synthesis. For this study, we used only the novelty dimension (3 sub-scales, 15 items) and complexity sub-scale of the elaboration & synthesis dimension (5 items). These 15 Novelty items consist of five items in each of the Original, Surprising and Germinal sub-scales. Similarly, the complexity subscale consists of five items. In all cases, items are scored on a 7-point Likert scale.

Personality - The five-factor model (the ‘Big Five’) is one of the contemporarily dominant personality taxonomies which consists of five factors: extraversion, agreeableness, conscientiousness, neuroticism, and intellect [15]. The mini-International Personality Item Pool (mini-IPIP) is a survey instrument with four items to measure each of the five traits and has been shown to be reliable and robust with a relatively short questionnaire [16].

Motivation - Motivation (component of conscientiousness) of team members has been shown to be a determinant in team success [17], [18]. In the context of education, the academic motivation scale is an English adaptation of the original measure of motivation developed in the French language [19]. In this survey instrument, motivation is further classified into intrinsic motivation, extrinsic motivation and amotivation. Intrinsic motivation may be driven by a need to know, a need to accomplish, or a need for experience. Extrinsic motivation may be externally regulated (i.e., do something because you’re told), introjected (i.e., do something to avoid feeling guilt), or identified (i.e., internalizing the extrinsic forces). Each of these factors is considered separately in the scale.

Social Loafing Tendency - In a group setting, when pressures to work come from outside the group and individual work is hard to identify, then the division of external pressure amongst group members leads to each member working less than if working alone. This drop in individual performance is referred to as social loafing [20]. In the academic context, a four item instrument has been developed to measure social loafing tendency [21].

Sucker Effect – The sucker effect is the reduction of individual efforts while working in a team context owing to a perception that others are free-riding [22]. While social loafing is an outcome focused phenomenon, the sucker effect focuses on the above mechanism. Hence, it is measured using an instrumental factor, an ethical factor and an equity factor. These factors were constructed based on the Australian Work Ethic Scale [23] and the Protestant Work Ethic Scale [24].

Social Compensation - Social compensation can be described as the tendency of individuals, especially those with low interpersonal trust, to work harder in a team environment in order to compensate for the lower performing teammates [25]. In turn, interpersonal trust, the expectancy that the spoken statements of others can be relied upon, has been shown to be a useful parameter in the context of student-teams [26] and has been refined into a well-researched and validated survey instrument [27].

Tolerance to Uncertainty and Ambiguity - In technical problem solving research, uncertainty and ambiguity are often treated as the same [28]. However, uncertainty is characterized by lack of information about value of known relevant parameters while ambiguity is characterized by lack of clarity about the relevant parameters and their relationship [28]. With the above differentiation, it can be said that all development projects are an exercise in uncertainty reduction and that one may expect engineers to be at least somewhat tolerant to uncertain situations. To characterize the degree of tolerance, a 20 item survey instrument was developed [29]. Tolerance to ambiguity has a been a widely studied area over the last 50 years [30], [31]. Budner conceives tolerance to ambiguity as a personality trait [32] and defines it as the tendency to perceive situations which cannot be adequately structured or categorized by the individual, because of the lack of sufficient cues, as sources of threat as intolerance to ambiguity [32]. From this perspective, phenomenological denial (repression and denial), phenomenological submission (anxiety and discomfort), operative denial (destructive or reconstructive behavior), or operative submission (avoidance behavior) as a response to novelty, complexity or insolvability of a situation is interpreted as intolerance to ambiguity [32]. These latent phenomena have been captured in a 16-item scale to measure tolerance to ambiguity [32].

Design Team Outcome - Both universities in this study require student teams to create a poster describing the prototype (final project outcome). This poster (design artifact) was the subject of rater evaluation using the CPSS and DQR. The CPSS was used to assess the novelty and complexity of the design outcome [33]. The novelty dimension of the scale consists of 15 items that are rated on a 7-point Likert scale. These 15 items consist of five items each of Original, Surprising and Germinal sub-scales. Similarly, the complexity subscale consists of five items, and it measures the complexity of the developed solution on a 7-point Likert scale. The DQR was created by Sobek [34] for design assessment in engineering education by consolidating 23 evaluation rubrics collected from various universities and design competitions. The DQR measures five dimensions of the project outcome: meeting the technical criteria and the customer requirements; being feasible in its application and fabrication / assembly; incorporating original and novel ideas, non-intuitive approaches, or innovative solutions; being simple, avoiding any unnecessary sophistication and complexity; and the overall impression. In 2005, Meyer et al. [35] conducted a thorough review of instruments for quantitative evaluation of capstone design outcomes. Among the various instruments reviewed, the DQR was found to be suitable for evaluating student projects for its ability to evaluate a diverse range of design projects in a project-independent and process-independent manner. Other researchers in the field of engineering education have relied upon this instrument as well [36], [37].

Team Characteristics - Since the late 1990s, there has been research on operationalizing individual measures into team measures. The relation between individual team member characteristics - characterized by their mean, min, max and variance - and team effectiveness has been studied [3]. Depending upon the task type - additive, compensatory, conjunctive, or disjunctive [38], different member characteristic aggregation mechanisms may be suitable. Figure 1 summarizes the inputs and outputs that were measured for each student team. Table 3 summarized what a larger value of each measure indicates.

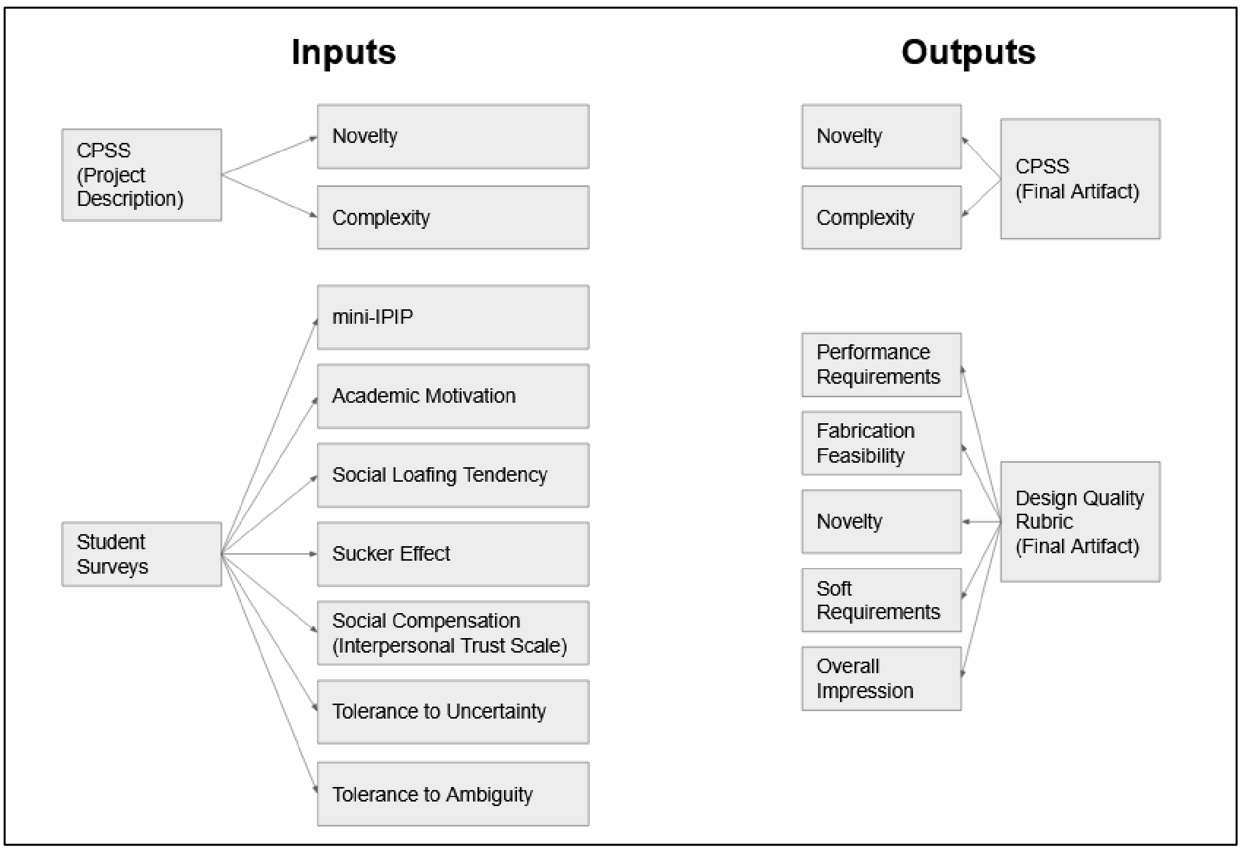

Figure 1: Summary of input and output measures

Table 3: Interpretation of rating scale of various measures

Category	Measure	Meaning of a larger value
Personality (Big 5)	Extraversion	Higher extraversion
	Agreeableness	Higher agreeableness
	Conscientiousness	Higher conscientiousness
	Neuroticism	Higher neuroticism
	Intellect	Higher intellect
Motivation	Intrinsic Motivation - To Know	Higher motivation from participation and doing the activity
	Intrinsic Motivation - Towards Accomplishment	Higher motivation from desire to be competent and create
	Intrinsic Motivation - To Experience Stimulation	Higher motivation from stimulation that the activity provides
	Extrinsic Motivation - Identified	Higher motivation from internalized (for individual good), external means
	Extrinsic Motivation - Introjected	Higher motivation from internalized (for general good), external means
	Extrinsic Motivation - External Regulation	Higher motivation from external means such as rewards and constraints
	Amotivation	More extreme lack of motivation
Social Loafing	Social Loafing Tendency	Higher tendency to work less when in a group than when alone
Sucker Effect	Ethical Factor	Higher belief in moral righteousness of hard work
	Instrumental Factor	Higher belief that work should be rewarded
	Equity Factor	Higher belief in effort-proportionate rewards
Social Compensation	Interpersonal Trust Scale	Higher belief in reliability of behavior, promises, or statements of other individuals
Tolerance to Uncertainty	Tolerance to Uncertainty	More comfort and better thriving in uncertainty
Tolerance to Ambiguity	Tolerance to Ambiguity	More comfort and better thriving in ambiguity

Data collection

The survey instruments discussed in the literature review section were compiled into a web-based survey platform. This included the mini-international personality item survey [16], academic motivation scale survey [19], social loafing survey [21], sucker effect survey [39], and interpersonal trust scale [27]; as well as CPSS for both task characteristics and design team outcome, and DQR for design team outcome. Task characteristics and design team outcomes were evaluated by three raters, two from RIT and one from NIU, in order to reduce the potential for individual bias. Some items were flipped in accordance with the original instruments to reduce likelihood of order bias and skewing. Institutional Review Board (IRB) approval was sought upon NSF grant approval and upon IRB approval the surveys were sent out to students in October 2020.

At NIU, student characteristics were collected in a single survey. At RIT, student characteristics were split across three surveys to keep the length of the reasonable and to prevent rater fatigue. Each of these three surveys was expected to take around 10 minutes to complete. Reminders, sent to students for the following three weeks, urged students to participate in the survey and notified them that the participation was voluntary. The surveys were closed after three weeks. Table 4 summarizes survey response rate. Out of the 61 teams at NIU and 64 teams at RIT, three teams and NIU and 20 teams at RIT had two or more student respondents per team for all survey instruments.

Table 4: Summary of data collection at the two universities

		NIU (Single online survey)	RIT (Split into 3 online surveys)
Response rate	Survey	36/187 (19%)	-
	Survey 1	-	116/331 (35%)
	Survey 2	-	94/331 (28%)
	Survey 3	-	78/331 (24%)
Response by team	One response	28	50
	More than one response	3	20

Analysis and Results

Teams with two or more student respondents were retained for further analysis to stay aligned with the purpose of studying ‘team characteristics’ as opposed to ‘individual student’ characteristics. Based on the small number of NIU teams with at least one response per team, we only present the analysis of RIT teams in this paper.

Python, distributed through Anaconda, was used to perform the data analysis. Raw data from the various survey instruments were collected in .csv file format. Several of the instruments included questions with reversed scales, and after correcting the flipped survey items using Microsoft Excel, the data was imported into Spyder IDE. The Pandas library was used for data preparation. The actual correlation coefficient and p-values were calculated using the ‘stats’ module of the SciPy library. In this calculation, the p-value is calculated with the null hypothesis that the distributions underlying the samples are uncorrelated and normally distributed, and the alternative hypothesis is two-sided [40].

Table 5 summarizes Cronbach’s alpha of the survey given at RIT which includes all student responses and those students on teams with >1 respondents. Almost all survey instruments had Cronbach’s alpha higher than the acceptable value of 0.7, indicating that the scales used are internally consistent.

Table 5: Cronbach alpha of all survey instruments measuring individual characteristic

Category	Measure	Cronbach's alpha (all)	Cronbach's alpha (n>1)
Personality (Big 5)	Extraversion	0.87	0.89
	Agreeableness	0.77	0.80
	Conscientiousness	0.62	0.59
	Neuroticism	0.73	0.61
	Intellect	0.68	0.79
Motivation	Intrinsic Motivation - To Know	0.84	0.81
	Intrinsic Motivation - Towards Accomplishment	0.87	0.87
	Intrinsic Motivation - To Experience Stimulation	0.74	0.73
	Extrinsic Motivation - Identified	0.82	0.87
	Extrinsic Motivation - Introjected	0.87	0.88
	Extrinsic Motivation - External Regulation	0.87	0.82
	Amotivation	0.84	0.86
Social Loafing	Social Loafing Tendency	0.71	0.70
Sucker Effect	Ethical Factor	0.65	0.75
	Instrumental Factor	0.78	0.80
	Equity Factor	0.78	0.80
Social Compensation	Interpersonal Trust Scale	0.73	0.74
Tolerance to Uncertainty	Tolerance to Uncertainty	0.90	0.89
Tolerance to Ambiguity	Tolerance to Ambiguity	0.41	0.46

Inter-rater reliability, calculated using the method described by James et al. [41], is summarized in Table 6. Inter-rater reliability for each outcome measure for each team was calculated based on ratings of each of the three raters. Then, using a threshold $r_{w,g}$ of 0.8 as a signifier of agreement, the percentage of teams for which raters agree with each other was reported. Similarly, the average $r_{w,g}$ for each outcome measure was reported.

Table 6: Inter rater reliability summarized by teams, for both task characteristics and design team outcome ($r_{wg} \geq 0.8$ was interpreted as 'agree')

	CPSS_ProjDesc_Novelty	CPSS_ProjDesc_Complexity	CPSS_FinalArtf_Novelty	CPSS_FinalArft_Complexity	DQR_CustReq	DQR_Feasibility	DQR_Innovative	DQR_Simplicity	DQR_Overall
Average r_{wg}	0.98	0.96	0.99	0.97	0.82	0.90	0.89	0.91	0.86
Percentage of r_{wg} > 0.8	100%	100%	100%	100%	65%	85%	90%	95%	80%

Tables 7-9 summarize the correlation analysis. Several team characteristics were found to have statistically significant correlations with design team outcome measures. Tables 7-9 summarize correlations between the team characteristics (aggregated from individual member characteristics using various statistical measures) and design team outcome measures. In all figures, a green cell represents a positive correlation with significance level p<0.1, and a red cell represents a negative correlation with p<0.1.With n=20 and p=0.1, the corresponding absolute value of r for statistically significant correlation is 0.378. The actual correlation values are summarized in Appendix A.

Table 7: The Big Five Personality Traits correlated with design team outcome measures

Measure	Statistic	Novelty_CPSS	Complexity_CPSS	DQR_TechCrit	DQR_Feasiblity	DQR_Innovative	DQR_Simplicity	DQR_Overall
Extraversion	Mean			+				
Extraversion	Max	+		+	+			
Extraversion	Min							
Extraversion	Range							
Extraversion	StDev							
Agreeableness	Mean			+	+	+	+	+
Agreeableness	Max					+	+	+
Agreeableness	Min					+		
Agreeableness	Range							
Agreeableness	StDev							
Conscientiousness	Mean	−						
Conscientiousness	Max							
Conscientiousness	Min	−	−					
Conscientiousness	Range	+	+		+			
Conscientiousness	StDev	+	+		+			
Neuroticism	Mean					+		
Neuroticism	Max			+	+	+	+	+
Neuroticism	Min							
Neuroticism	Range			+	+	+	+	+
Neuroticism	StDev			+	+	+	+	+
Intellect	Mean			+	+	+	+	+
Intellect	Max			+	+		+	+
Intellect	Min			+		+		
Intellect	Range							
Intellect	StDev							

Table 8: Motivation correlated with design team outcome measures

Measure	Statistic	CPSS_Novelty	CPSS_Complexity	DQR_TechCrit	DQR_Feasiblity	DQR_Innovative	DQR_Simplicity	DQR_Overall
Intrinsic Motivation - To Know	Mean							
Intrinsic Motivation - To Know	Max							
Intrinsic Motivation - To Know	Min							
Intrinsic Motivation - To Know	Range							
Intrinsic Motivation - To Know	StDev							
Intrinsic Motivation - Towards Accomplishment	Mean							
Intrinsic Motivation - Towards Accomplishment	Max							
Intrinsic Motivation - Towards Accomplishment	Min							
Intrinsic Motivation - Towards Accomplishment	Range							
Intrinsic Motivation - Towards Accomplishment	StDev							
Intrinsic Motivation - To Experience Stimulation	Mean							
Intrinsic Motivation - To Experience Stimulation	Max							
Intrinsic Motivation - To Experience Stimulation	Min			−				
Intrinsic Motivation - To Experience Stimulation	Range							
Intrinsic Motivation - To Experience Stimulation	StDev							
Extrinsic Motivation - Identified	Mean							
Extrinsic Motivation - Identified	Max							
Extrinsic Motivation - Identified	Min			−				
Extrinsic Motivation - Identified	Range			+				+
Extrinsic Motivation - Identified	StDev			+				+
Extrinsic Motivation - Introjected	Mean			−				
Extrinsic Motivation - Introjected	Max							
Extrinsic Motivation - Introjected	Min			−				−
Extrinsic Motivation - Introjected	Range			+				
Extrinsic Motivation - Introjected	StDev			+				+
Extrinsic Motivation - External Regulation	Mean							
Extrinsic Motivation - External Regulation	Max							
Extrinsic Motivation - External Regulation	Min							
Extrinsic Motivation - External Regulation	Range							
Extrinsic Motivation - External Regulation	StDev							
Amotivation	Mean							
Amotivation	Max							
Amotivation	Min				−			
Amotivation	Range							
Amotivation	StDev							
Average Academic Motivation	Mean							
Average Academic Motivation	Max							
Average Academic Motivation	Min							
Average Academic Motivation	Range							
Average Academic Motivation	StDev			+				

Table 9: Social loafing tendency, sucker effect, social compensation, tolerance to uncertainty, and tolerance to ambiguity correlated with design team outcome measures

Measure	Statistic	Novelty_CPSS	Complexity_CPSS	DQR_TechCrit	DQR_Feasiblity	DQR_Innovative	DQR_Simplicity	DQR_Overall
Social Loafing	Mean						−	−
Social Loafing	Max						−	
Social Loafing	Min						−	
Social Loafing	Range							
Social Loafing	StDev							
Ethical	Mean							
Ethical	Max							
Ethical	Min							
Ethical	Range							
Ethical	StDev							
Instrumental	Mean							
Instrumental	Max							
Instrumental	Min							
Instrumental	Range							
Instrumental	StDev							
Equity	Mean						−	
Equity	Max	−						
Equity	Min							
Equity	Range	−						
Equity	StDev							
Social Compensation (Interpersonal Trust)	Mean					+		
Social Compensation (Interpersonal Trust)	Max							
Social Compensation (Interpersonal Trust)	Min						+	
Social Compensation (Interpersonal Trust)	Range							
Social Compensation (Interpersonal Trust)	StDev							
Tolerance To Uncertainty	Mean							
Tolerance To Uncertainty	Max							
Tolerance To Uncertainty	Min							
Tolerance To Uncertainty	Range							
Tolerance To Uncertainty	StDev							
Tolerance To Ambiguity	Mean			+	+	+	+	+
Tolerance To Ambiguity	Max			+	+		+	+
Tolerance To Ambiguity	Min						+	+
Tolerance To Ambiguity	Range							
Tolerance To Ambiguity	StDev							

Discussion

The Big 5 measures were found to have the most statistically significant correlations with design team outcomes, as shown in Table 7. The distribution (both standard deviation and range) of conscientiousness had positive correlation while the min conscientiousness had negative correlation with both novelty and complexity measures of the CPSS evaluation. Agreeableness and neuroticism had a positive correlation with DQR measures. A plausible causal mechanism is that agreeableness would improve team collaboration and thereby design team outcome. Interestingly, neuroticism showed a positive correlation with the DQR framework in both the team maximum and the interval/standard deviation. This implies that having a team member with high neuroticism is a positive factor, but it is also beneficial to balance that with team member(s) with lower neuroticism scores. Intellect was found to have statistically significant positive correlation with all measures of the DQR framework - this finding is in-line with the commonly accepted notion that intelligent people lead to better outcomes.

Correlations of various motivation measures with design team outcome measures are presented in Table 8. Extrinsic motivation had a correlation with the "meeting technical requirements" measure of the DQR framework but not with other DQR measures like innovativeness, feasibility, and simplicity. This finding aligns with the commonly observed student behavior of checking all boxes towards an acceptable deliverable but not going beyond with grading based (extrinsic) reinforcement. Similarly, amotivation (lack of motivation) was found to have a negative correlation with the feasibility measure of the DQR framework.

Social loafing tendency, the sucker effect, social compensation, tolerance to uncertainty, and tolerance to ambiguity correlations with design team outcomes are presented in Table 9. As expected, lower social loafing, lower sucker effect and higher social compensation would lead to better project outcomes. The mean equity factor (sucker effect) for teams had a negative correlation with the simplicity measure of DQR. If the team, on average, says they are less likely to want to reward everyone equally regardless of effort, then the solution has less unwanted or unnecessary complexity. Further, the max equity score negatively correlates to high novelty. The lower the maximum score (the less likely you are to have a team member who believes reward should be proportional to effort) the more likely the design team outcome will have high novelty. A possible explanation is that the team members are more willing to entertain wild ideas without fear when they would be discounted as not being valuable contributions. The interval for equity also negatively correlates to high novelty. In other words, if team members have similar views toward how rewards will be distributed, the output novelty is higher.

Also shown in Table 9, tolerance to ambiguity has a positive correlation with all DQR measures of design team outcome. This suggests that teams comfortable with ambiguous situations tend to deliver better design project outcomes. This in turn justifies the conventional wisdom of helping students to deal with the ambiguity present in open-ended design projects. However, the reliability of the tolerance to ambiguity survey yielded a low Cronbach Alpha value of 0.41 which hints at the need to revisit the survey instrument.

Social loafing has a negative correlation with the simplicity measure of DQR. A plausible causal mechanism could be that, to meet minimally acceptable customer requirements, the students develop independent subsystem-level solutions to design challenges, but with reduced person-hours (due to social loafing) not enough effort can be put on integration of the subsystems. This can lead to a needlessly complex solution with poorly defined interfaces. Such solutions would then get rated poorly for DQR simplicity measure.

Outcome measures of the CPSS evaluation framework (novelty, complexity) did not correlate with the outcome measures of the DQR evaluation framework. CPSS is an extensively validated survey instrument while DQR has received relatively less academic attention. In fact, seemingly corresponding items in the two instruments yielded contrasting correlations. This could indicate that these evaluation frameworks are measuring different latent outcome attributes. Post-hoc discussion among the raters (authors of this paper) revealed that rating DQR items was easier (raters were more confident with their rating) compared to CPSS items.

Interestingly, except for the Big 5 and equity factor of sucker effect, no other team characteristic had a statistically significant correlation with CPSS evaluation. For an exhaustive list of statistically significant correlations, refer to Tables 7-9 in the analysis section.

As seen in Tables 7-9, range and standard deviation of inputs tend to show correlations (or not) together. In previous academic works, standard deviation has been used as the measure of variability among team member characteristics. However, in the case of small team sizes with 5 or less members (as is typical of most engineering capstone design teams [9]), it is useful to use range as the measure of variability in order to save a degree of freedom. An argument can be made that a range statistic is sensitive to outliers, but in the case of small student teams, a single student with an individual characteristic outside the usual value is often seen to heavily influence team dynamics. So, in fact, such outlier values of individual characteristic and their influence on the team level characteristic would be useful to retain.

There was some challenge among raters to confidently rate project descriptions (task characteristics) using the CPSS instrument. Consistently reliable and holistic evaluation of student project outcomes continues to be challenging, but the DQR holds promise as it is tailored to academic context.

From a holistic perspective, engineering education needs to improve the understanding of the ‘nature’ (individual characteristics) of a diverse set of students and provide appropriate ‘nurture’ (team processes) for each set of students. More diversity in the way students approach problems is encouraged but there may be limits. For example, a high range of equity factor among team members produced worse design team outcomes. A team process, by definition, applies to all team members and a high diversity of individual characteristics within a team may render an optimal team process to be sub-optimal for each individual member. Personalized team processes (different activities for each individual team member) may help foster synergy among diverse team members. Research into such aspects of effective teams would produce better design team outcomes as well as nurture each individual student to become a better future engineer.

Conclusion

Following the team effectiveness model discussed by Takai et. al [14], the correlations between task characteristics and student team characteristics with design team outcomes were explored. The effectiveness of various survey instruments in linking inputs to design team outcomes was found to be widely distributed. In line with previous research, the five-factor model (Big 5), operationalized using the mini-IPIP survey, yielded the most statistically significant correlations with design team outcomes. A full list of other statistically significant correlations has been established and many measures of team characteristics and project characteristics appear to have no statistically significant correlation with the quality of the design outcome. Lastly, further research is warranted to establish normative metrics to evaluate aspects of engineering capstone design projects.

Limitation and Future Work

In this study, the response rate was relatively low (about 20% at NIU and 24% at RIT). This also resulted in a small number of teams with more than one response (4 teams at NIU and 20 teams at RIT). Furthermore, 13 of these 20 teams at RIT had just two student respondents per team (modal team size = 4). While not included in this paper, the student surveys were administered again in Fall 2021 on paper at NIU, which increased response rates to 85%. At RIT, surveys were still administered electronically during Fall 2021, but the last author paid a brief visit to each team to explain the study and encourage participation, increasing the response rate to 38%.

With the larger number of students’ and teams’ data, we will conduct comprehensive analysis including studying correlation between cognitive ability, psychological safety, or decision making and design team performance. The comprehensive study of both individual students’ and teams’ data should enable us to find team formation methodologies that lead to better design team performance. This comprehensive study is left for future work.

Acknowledgments

This work was supported, in part, by the National Science Foundation grants NSF-DUE-2021497 awarded to NIU and NSF-DUE-2021434 awarded to RIT. The authors would also like to acknowledge Dr. Donald Peterson and Ms. Justyna Kielar at NIU and Mr. Chandler Libby at RIT for assisting with data collection. The authors would also like to thank all faculty and staff associated with the RIT capstone program.

References

[1] J. Marschak, “Elements for a theory of teams,” *Manage. Sci.*, vol. 1, no. 2, pp. 127–137, 1955.

[2] B. Oakley, R. M. Felder, R. Brent, and I. Elhajj, “Turning student groups into effective teams,” *J. student centered Learn.*, vol. 2, no. 1, pp. 9–34, 2004.

[3] M. R. Barrick, G. L. Stewart, M. J. Neubert, and M. K. Mount, “Relating member ability and personality to work-team processes and team effectiveness.,” *J. Appl. Psychol.*, vol. 83, no. 3, p. 377, 1998.

[4] S. I. Tannenbaum, R. L. Beard, and E. Salas, “Team building and its influence on team effectiveness: An examination of conceptual and empirical developments,” in *Advances in psychology*, vol. 82, Elsevier, 1992, pp. 117–153.

[5] J. Mathieu, M. T. Maynard, T. Rapp, and L. Gilson, “Team effectiveness 1997-2007: A review of recent advancements and a glimpse into the future,” *J. Manage.*, vol. 34, no. 3, pp. 410–476, 2008.

[6] E. Sundstrom, K. P. De Meuse, and D. Futrell, “Work teams: Applications and effectiveness.,” *Am. Psychol.*, vol. 45, no. 2, p. 120, 1990.

[7] K. Beck *et al.*, “Manifesto for agile software development,” 2001.

[8] ABET, “Criteria for Accrediting Engineering Programs, 2022 – 2023,” 2021. [Online]. Available: https://www.abet.org/accreditation/accreditation-criteria/criteria-for-accrediting-engineering-programs-2022-2023/.

[9] S. Howe, L. Rosenbauer, and S. Poulos, “The 2015 capstone design survey results: Current practices and changes over time,” *Int. J. Eng. Educ.*, vol. 33, no. 5, p. 1393, 2017.

[10] D. Hall and S. Buzwell, “The problem of free-riding in group projects: Looking beyond social loafing as reason for non-contribution,” *Act. Learn. High. Educ.*, vol. 14, no. 1, pp. 37–49, 2013.

[11] N. M. Webb, “Assessing students in small collaborative groups,” *Theory Pract.*, vol. 36, no. 4, pp. 205–213, 1997.

[12] P. Lewis, D. Aldridge, and P. M. Swamidass, “Assessing teaming skills acquisition on undergraduate project teams,” *J. Eng. Educ.*, vol. 87, no. 2, pp. 149–155, 1998.

[13] E. Pfaff and P. Huddleston, “Does it matter if I hate teamwork? What impacts student attitudes toward teamwork,” *J. Mark. Educ.*, vol. 25, no. 1, pp. 37–45, 2003.

[14] S. Takai and M. Esterman, “A review of team effectiveness models and possible instruments for measuring design-team inputs, processes, and outputs,” *Int. J. Eng. Educ.*, vol. 35, no. 6, pp. 1684–1697, 2019.

[15] L. R. Goldberg, “An alternative" description of personality": the big-five factor structure.,” *J. Pers. Soc. Psychol.*, vol. 59, no. 6, p. 1216, 1990.

[16] M. B. Donnellan, F. L. Oswald, B. M. Baird, and R. E. Lucas, “The mini-IPIP scales: tiny-yet-effective measures of the Big Five factors of personality.,” *Psychol. Assess.*, vol. 18, no. 2, p. 192, 2006.
[17] E. G. French, “Some characteristics of achievement motivation.,” *J. Exp. Psychol.*, vol. 50, no. 4, p. 232, 1955.
[18] F. W. Schneider and J. G. Delaney, “Effect of individual achievement motivation on group problem-solving efficiency,” *J. Soc. Psychol.*, vol. 86, no. 2, pp. 291–298, 1972.
[19] R. J. Vallerand, L. G. Pelletier, M. R. Blais, N. M. Briere, C. Senecal, and E. F. Vallieres, “The Academic Motivation Scale: A measure of intrinsic, extrinsic, and amotivation in education,” *Educ. Psychol. Meas.*, vol. 52, no. 4, pp. 1003–1017, 1992.
[20] B. Latané, K. Williams, and S. Harkins, “Many hands make light the work: The causes and consequences of social loafing.,” *J. Pers. Soc. Psychol.*, vol. 37, no. 6, p. 822, 1979.
[21] M. C. Schippers, “Social loafing tendencies and team performance: The compensating effect of agreeableness and conscientiousness,” *Acad. Manag. Learn. Educ.*, vol. 13, no. 1, pp. 62–81, 2014.
[22] J. M. Jackson and S. G. Harkins, “Equity in effort: An explanation of the social loafing effect.,” *J. Pers. Soc. Psychol.*, vol. 49, no. 5, p. 1199, 1985.
[23] R. Ho and J. I. Lloyd, “Development of an Australian work ethic scale,” *Aust. Psychol.*, vol. 19, no. 3, pp. 321–332, 1984.
[24] H. L. Mirels and J. B. Garrett, “The Protestant ethic as a personality variable.,” *J. Consult. Clin. Psychol.*, vol. 36, no. 1, p. 40, 1971.
[25] K. D. Williams and S. J. Karau, “Social loafing and social compensation: The effects of expectations of co-worker performance.,” *J. Pers. Soc. Psychol.*, vol. 61, no. 4, p. 570, 1991.
[26] J. B. Rotter, “A new scale for the measurement of interpersonal trust.,” *J. Pers.*, 1967.
[27] J. P. Robinson, P. R. Shaver, and L. S. Wrightsman, “Criteria for scale selection and evaluation,” *Meas. Personal. Soc. Psychol. attitudes*, vol. 1, pp. 1–16, 1991.
[28] S. Schrader, W. M. Riggs, and R. P. Smith, “Choice over uncertainty and ambiguity in technical problem solving,” *J. Eng. Technol. Manag.*, vol. 10, no. 1–2, pp. 73–99, Jun. 1993.
[29] M. H. Freeston, J. Rhéaume, H. Letarte, M. J. Dugas, and R. Ladouceur, “Why do people worry?,” *Pers. Individ. Dif.*, vol. 17, no. 6, pp. 791–802, 1994.
[30] A. Furnham and T. Ribchester, “Tolerance of ambiguity: A review of the concept, its measurement and applications,” *Curr. Psychol.*, vol. 14, pp. 179–199, 1995.
[31] A. Furnham and J. Marks, “Tolerance of ambiguity: A review of the recent literature,” *Psychology*, vol. 4, no. 09, pp. 717–728, 2013.
[32] N. Y. Stanley Budner, “Intolerance of ambiguity as a personality variable 1,” *J. Pers.*, vol. 30, no. 1, pp. 29–50, 1962.
[33] K. O’Quin and S. P. Besemer, “The development, reliability, and validity of the revised creative product semantic scale,” *Creat. Res. J.*, vol. 2, no. 4, pp. 267–278, 1989.
[34] D. Sobek and V. Jain, “Two instruments for assessing design outcomes of capstone projects,” in *2004 Annual Conference*, 2004, pp. 9–1327.
[35] D. G. Meyer, “Capstone design outcome assessment: Instruments for quantitative evaluation,” in *Proceedings Frontiers in Education 35th Annual Conference*, 2005, pp. F4D-F4D.
[36] R. Gerlick *et al.*, “Assessment structure and methodology for design processes and

products in engineering capstone courses," in *2008 Annual Conference & Exposition*, 2008, pp. 13–240.
[37] S. Santana, "Instrumentation for Evaluating Design-learning and Instruction Within Courses and Across Programs," in *2021 ASEE Virtual Annual Conference Content Access*, 2021.
[38] I. D. Steiner, *Group process and productivity*. Academic press New York, 1972.
[39] S. Abele and M. Diehl, "Finding teammates who are not prone to sucker and free-rider effects: The protestant work ethic as a moderator of motivation losses in group performance," *Gr. Process. Intergr. Relations*, vol. 11, no. 1, pp. 39–54, 2008.
[40] P. Virtanen *et al.*, "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python," *Nat. Methods*, vol. 17, pp. 261–272, 2020.
[41] L. R. James, R. G. Demaree, and G. Wolf, "Estimating within-group interrater reliability with and without response bias.," *J. Appl. Psychol.*, vol. 69, no. 1, p. 85, 1984.

Appendix A – Coefficients of statistically significant correlations.
With n=20 and p-values of p=0.1 and p=0.05, the corresponding absolute values of statistically significant correlation are 0.378 and 0.444. Regular fonts are used for p<0.1 and bold fonts for p<0.05. Positive correlations are shown in green fonts and negative correlations in red fonts.

Measure	Statistic	Novelty_CPSS	Complexity_CPSS	DQR_TechCrit	DQR_Feasibility	DQR_Innovative	DQR_Simplicity	DQR_Overall
Extraversion	Mean			0.40				
Extraversion	Max	0.44		**0.49**	**0.47**			
Extraversion	Min							
Extraversion	Range							
Extraversion	StDev							
Agreeableness	Mean			0.43	0.42	**0.46**	0.39	0.43
Agreeableness	Max					**0.45**	**0.53**	**0.53**
Agreeableness	Min					0.44		
Agreeableness	Range							
Agreeableness	StDev							
Conscientiousness	Mean	**-0.45**						
Conscientiousness	Max							
Conscientiousness	Min	**-0.54**	-0.39					
Conscientiousness	Range	**0.57**	**0.48**		**0.50**			
Conscientiousness	StDev	**0.59**	**0.48**		**0.46**			
Neuroticism	Mean					**0.54**		
Neuroticism	Max			**0.50**	0.44	**0.52**	**0.48**	**0.54**
Neuroticism	Min							
Neuroticism	Range			**0.50**	0.43	0.42	**0.48**	**0.52**
Neuroticism	StDev			**0.49**	0.44	0.42	**0.48**	**0.52**
Intellect	Mean			**0.63**	**0.50**	**0.46**	**0.51**	**0.51**
Intellect	Max			**0.61**	**0.56**		**0.45**	0.44
Intellect	Min			0.39		0.44		
Intellect	Range							
Intellect	StDev							
Social Loafing	Mean						**-0.61**	-0.42
Social Loafing	Max						**-0.46**	
Social Loafing	Min						**-0.50**	
Social Loafing	Range							
Social Loafing	StDev							
Ethical	Mean							
Ethical	Max							
Ethical	Min							
Ethical	Range							
Ethical	StDev							
Instrumental	Mean							
Instrumental	Max							
Instrumental	Min							
Instrumental	Range							
Instrumental	StDev							
Equity	Mean						-0.42	
Equity	Max	-0.44						
Equity	Min							
Equity	Range	-0.43						
Equity	StDev							
Social Compensation (Interpersonal Trust)	Mean					0.41		
Social Compensation (Interpersonal Trust)	Max							
Social Compensation (Interpersonal Trust)	Min						0.40	
Social Compensation (Interpersonal Trust)	Range							
Social Compensation (Interpersonal Trust)	StDev							

Appendix A (Continued)

Measure	Statistic	Novelty_CPSS	Complexity_CPSS	DQR_TechCrit	DQR_Feasiblity	DQR_Innovative	DQR_Simplicity	DQR_Overall
Intrinsic Motivation - To Know	Mean							
Intrinsic Motivation - To Know	Max							
Intrinsic Motivation - To Know	Min							
Intrinsic Motivation - To Know	Range							
Intrinsic Motivation - To Know	StDev							
Intrinsic Motivation - Towards Accomplishment	Mean							
Intrinsic Motivation - Towards Accomplishment	Max							
Intrinsic Motivation - Towards Accomplishment	Min							
Intrinsic Motivation - Towards Accomplishment	Range							
Intrinsic Motivation - Towards Accomplishment	StDev							
Intrinsic Motivation - To Experience Stimulation	Mean							
Intrinsic Motivation - To Experience Stimulation	Max							
Intrinsic Motivation - To Experience Stimulation	Min			-0.41				
Intrinsic Motivation - To Experience Stimulation	Range							
Intrinsic Motivation - To Experience Stimulation	StDev							
Extrinsic Motivation - Identified	Mean							
Extrinsic Motivation - Identified	Max							
Extrinsic Motivation - Identified	Min			-0.43				
Extrinsic Motivation - Identified	Range			**0.60**				**0.45**
Extrinsic Motivation - Identified	StDev			**0.62**				**0.45**
Extrinsic Motivation - Introjected	Mean			-0.38				
Extrinsic Motivation - Introjected	Max							
Extrinsic Motivation - Introjected	Min			**-0.56**				**-0.45**
Extrinsic Motivation - Introjected	Range			**0.46**				
Extrinsic Motivation - Introjected	StDev			**0.46**				0.38
Extrinsic Motivation - External Regulation	Mean							
Extrinsic Motivation - External Regulation	Max							
Extrinsic Motivation - External Regulation	Min							
Extrinsic Motivation - External Regulation	Range							
Extrinsic Motivation - External Regulation	StDev							
Amotivation	Mean							
Amotivation	Max							
Amotivation	Min				**-0.47**			
Amotivation	Range							
Amotivation	StDev							
Average Academic Motivation	Mean							
Average Academic Motivation	Max							
Average Academic Motivation	Min							
Average Academic Motivation	Range							
Average Academic Motivation	StDev			0.38				
Tolerance To Uncertainty	Mean							
Tolerance To Uncertainty	Max							
Tolerance To Uncertainty	Min							
Tolerance To Uncertainty	Range							
Tolerance To Uncertainty	StDev							
Tolerance To Ambiguity	Mean			**0.60**	**0.56**	0.44	**0.66**	**0.72**
Tolerance To Ambiguity	Max			**0.62**	**0.55**		**0.57**	**0.68**
Tolerance To Ambiguity	Min						**0.48**	**0.49**
Tolerance To Ambiguity	Range							
Tolerance To Ambiguity	StDev							