What Factors Impact Psychological Safety in Engineering Student Teams? A Mixed-Method Longitudinal Investigation

Courtney Cole
The Pennsylvania State University
343 Leonhard Building,
University Park, PA 16802-1401
Email: cmc6503@psu.edu

Abigail O'Connell
The Pennsylvania State University
343 Leonhard Building,
University Park, PA 16802-1401
Email: aeo5146@psu.edu

Zibing Gong
The Pennsylvania State University
343 Leonhard Building,
University Park, PA 16802-1401
zmg5089@psu.edu

Kathryn Jablokow
School of Graduate Professional Studies,
The Pennsylvania State University,
30 E. Swedesford Road,
Malvern, PA 19355,
Email: KWL3@psu.edu
ASME Fellow

Susan Mohammad
Department of Psychology,
The Pennsylvania State University,
617 Moore Building,
University Park, PA 16802
Email: sxm40@psu.edu

Sarah Ritter
School of Engineering Design, Technology and Professional Programs,
The Pennsylvania State University,
213 Hammond Building,
University Park, PA 16802,
Email: scr15@psu.edu

Katie Heininger
The Pennsylvania State University
343 Leonhard Building,
University Park, PA 16802-1401
Email: katieheininger@psu.edu

Jacqueline Marhefka
Department of Psychology,
The Pennsylvania State University,
140 Moore Building,
University Park, PA 16802,
Email: jtm40@psu.edu

Scarlett R. Miller
The Pennsylvania State University
213-P Hammond Building
University Park, PA 16802-1401
scarlettmiller@psu.edu
ASME Member

ABSTRACT

Although teamwork is being integrated throughout engineering education because of the perceived benefits of teams, the construct of psychological safety has been largely ignored in engineering research. This omission is unfortunate, because psychological safety reflects collective perceptions about how comfortable team members feel in sharing their perspectives and it has been found to positively impact team performance in samples outside of engineering. While prior research has indicated that psychological safety is positively related to team performance and outcomes, research related to psychological safety in engineering teams is less established. There is also a lack of comprehensive methodologies that capture the dynamic changes that occur throughout the design process and at each time point. In light of this, the goal of the current study was to understand how psychological safety might be measured practically and reliably in engineering student teams over time. In addition, we sought to identify factors that impact the building and waning of psychological safety in these teams over time This was accomplished through a study with 260 engineering students in 68 teams in a first-year engineering design class. The psychological safety of the teams was captured for each team over five time points over the course of a semester long design project. The results of this study provide some of the first evidence on the reliability of psychological safety in engineering teams and offer insights as to how to support and improve psychological safety.

1.0 INTRODUCTION

Engineering organizations around the globe are increasingly relying on team-based activities [1, 2] based on teams' abilities to solve complex problems more effectively than individuals alone [2, 3]. This increase in performance has been attributed to the wide range of knowledge, experience, and expertise present in a team [3]. In other words, the "wisdom of the collective" [4] (p.39) is often better than the output of the sum of the individual team members [5, 6]. In an engineering context, this improvement is essential because companies need to produce new solutions to problems to thrive in the fast-paced market dependent on innovation [7]. Because of this demand from engineering industry, engineering education has also recognized and incorporated team processes into the engineering classroom [8, 9] especially through cornerstone (first-year) and capstone (final year) undergraduate design courses [10].

Despite the rise in the integration of teamwork in engineering education, how to cultivate teamwork skills in the classroom remains unclear [11]. While meta-analysis substantiates that various factors that could affect psychological safety in different scenarios [12, 13], the relative importance to engineering design education remains unclear. To have a better understanding of team performance, there should be more information on the dynamic assembly of teams, which is also lacking in current literature [14]. Moreover, there is also little information on *what, how, and when* to successfully intervene in team activities for better team performance.

To answer these questions, the study of the longitudinal trajectory of psychological safety in an engineering context can be applied to the engineering classroom. Specifically, psychological safety is, "a shared belief that the team is safe for interpersonal risk taking" ([15] p. 354). When people feel psychologically safe, they are more likely to speak up about their own opinion without fear of failure, which would allow them to learn and contribute more efficiently [12]. Psychological safety as a predictor has shown its consistency and generalizability when measuring outcomes from individuals, teams, and groups [12]. Specifically, meta-analysis outside of this has substantiated the relationship between psychological safety and performance [15]. When tasks are complex, knowledge-intensive, and innovative, this relationship is strongest. These are the traits that characterize the tasks in engineering design. However, there is limited research on psychological safety in an engineering context.

Despite the lack of studies of psychological safety in engineering, research in innovation management points to why this relationship may be critical to explore. Particularly, prior work has shown that psychological safety shares a relationship with creativity through promoting individuals to feel safe to share novel ideas and give constructive feedback [16, 17]; skills that are critical to foster in engineering education [11]. These skills guide teams throughout the engineering design process, which relies on an iterative process of three phases: generation, evaluation, and communication [18]. Low performance throughout these phases can hurt overall performance of the final design [18], substantiating the importance of investigating how psychological safety develops over time.

In light of this prior work, the goal of the current paper was to explore the role of psychological safety in student interactions at each of the conceptual phases of the engineering design process. Specifically, qualitative responses were assessed to understand what positive and negative interactions contributed to teams' ratings of psychological safety. The results of this study provide empirical evidence for the factors that contribute to the building and/or waning of psychological safety across the trajectory of an engineering design project. Furthermore, these results can be used to develop intervention methods geared towards particular issues (e.g., communication issues) at specific points in the design process.

2.0 RELATED WORK

In this section, we review the literature on psychological safety, focusing on how it is validated as a construct, and expand upon how it may impact processes and interactions that occur throughout the engineering design process. We also review studies of psychological safety to support the need to extend empirical knowledge to the engineering design context.

2.1 Measuring, Validating, and Establishing Psychological Safety in Teams

Psychological safety is a team construct established through deep interactions and conversations that facilitate how team members perceive their treatment from team members [19]. In other words, teams with high psychological safety share a general belief that nobody will be punished or humiliated for sharing ideas, making mistakes, or asking questions, for example. The construct first came to focus in organizational research in the 1990s and it has since been identified as a pivotal factor of teamwork, team learning, voice, and organizational learning [12]. The concept can be analyzed at the individual, group, and organizational level [12], making it applicable in a variety of industries such as healthcare [20] and manufacturing [14]. Particularly, meta-analysis has identified various constructs that feed into and come from psychological safety at the individual and team levels [12]. Some of these inputs are problem-solving efficacy, task conflict, social interactions, team characteristics (e.g., personality), and team leadership [12]; all of which can contribute to how psychological builds, wanes, or establishes as a team construct. From there, psychological safety can impact information sharing within a team, how teams handle conflict, and how teams learn–ultimately impacting performance [12]. Because building psychologically safe teams can be dependent on these constructs, identifying when teams' understandings of their psychological safety become established, and how and why psychological safety changes within teams over time remains a fundamental goal for expanding its research in engineering design.

The first step in building a psychologically safe team is understanding when a shared understanding of psychological safety exists across teams. This is critical to establish, as psychological safety is a *team-level* construct [15], where team members possess a shared belief of feeling safe for interpersonal risk-taking, meaning that team members can make decisions that affect the team without fear of repercussions if plans fail. However, this shared understanding may take time to establish. Specifically, psychological safety is a *dynamic* construct, meaning that it takes time to manifest through team interactions [21], which helps teams develop a shared mental model within the team [22, 23]. Therefore, analysts must justify aggregating individual psychological safety ratings to the team level to validate the construct. This requires a two-step process, where the first step requires interrater reliability (IRR) to be calculated as Cronbach's alpha [24]. This identifies when individuals' *perceptions of the what the scale is measuring* matches the *intended construct to be measured* [25]. In some cases, lower Cronbach's alpha values below the $\alpha = 0.70$ threshold [26] can indicate poor inter-relatedness, or there are heterogeneous constructs present in the scale [24]. The second step requires interrater agreement indices to be calculated, such as r_{wg} and ICC values, which justify whether the scale can be aggregated to the team-level [25]. Specifically, when all the individual members' ratings are similar with other individuals on their team (here, the individual psychological safety ratings), then the team construct (team psychological safety) can be considered a shared team level construct [27].

Beyond aggregations to the team level, understanding *when* a shared perception of psychological safety becomes established within a team can pose a challenge. While meta-analysis showed that psychological safety is consistent, generalizable, and a multi-level predictor of various outcomes in various contexts [12], many of these studies, examined psychological safety from a "snap-shot" view; failing to measure psychological safety over time. This is problematic when considering recent work in engineering design [28], as design teams must work together through different tasks at each stage of the design process over a duration of time [18]. Although many studies show the importance of psychological safety in fostering team performance, *when* to address issues due to low psychological safety, as well as *types* of intervention methods to promote psychological safety, remains unclear. This calls to attention the importance of identifying whether psychological safety is stable across the different time points of a team's lifespan.

2.2 Potential Impact of Psychological Safety on Phases of the Engineering Design Process

While psychological safety has seen limited treatment in engineering design, there are multiple benefits that support its potential utility in engineering design education. One major theme of psychological safety is that it facilitates the contribution of ideas and thus stimulates the team performance [12]. Specifically, meta-analytic evidence demonstrated that the relationship between psychological safety and learning, as well as performance is the strongest, especially when having complex, knowledge-intensive tasks that involve creativity and sense-making [13]. Higher psychological safety has the potential to increase creativity by allowing people to express their opinions in a dignified and respectful manner [29]. With creativity influence on engineering design, it is worth noticing the relationship between creativity and psychological safety. Studies also suggest that the development of psychological safety could enable creativity by encouraging participants to speak up with their novel ideas (concept generation), provide feedback to other members (concept selection), and challenge solutions throughout the process (e.g. during prototyping) [16, 30]. Particularly, the generation [31] and selection [32] of ideas are important for success of the final design, as prior research suggests that creativity and the likelihood of people speaking up could play an important role during the design process. These could also have an impact on the elucidation of sound decisions during the development stages of the project which are also related to the success of the final design [33]. However, within these stages, task conflict could occur. Task conflict has been demonstrated to have a relationship with team performance [34], showing that team members that are less likely to agree on tasks easily tend to produce creative outputs, promoting team performance [15]. Particularly in engineering design, this conflict provides an opportunity for team members to challenge others' ideas to improve performance [35]. However, psychological safety can mediate whether this conflict either hurts or helps teams [34]. For these reasons, investigating how psychological safety wanes and/or builds throughout the engineering design stages is critical to understanding how to leverage team performance.

In addition to the dynamics of the various stages of the engineering design process, it is also important to consider team behaviors in relation to psychological safety. As prior work indicates, psychological safety that results from team behaviors can impact the level of safety within an environment where individuals participate [12]. This implies that psychological safety could be subject to several factors. For example, interpersonal relationships, group dynamics (also an output of psychological safety that impacts further workplace practices [13]), leadership, and organizational norms are just four of factors that have been identified as inputs of psychological safety [36]. Additionally, specific team behaviors or learning behaviors could positively relate to psychological safety. For example, when team members are being respected by other members, they are more likely to gain confidence and speak up, knowing others will not hold their mistakes against them [12]. In addition, listening to each other [37] and the frequency of communication [38] are both proposed to have an influence on psychological safety. At this individual-to-individual level of psychological safety, other factors could feed into why team members may choose to treat specific members a certain way. Particularly, surface-level diversity, such as gender and ethnic background, have been shown to have differential effects on team performance and relationship conflict in teams [39]. Beyond surface-level diversity, factors grouped under deep-level diversity can also contribute to such differential effects [39]. Specifically, deep-level diversity can come in various forms, such differences due to cognitive style and level in the form of cognitive gap [40, 41]; a contributor of conflict in teams [41]. These are important to keep in mind as confounding factors of psychological safety, as diversity can impact how team members treat others and feel treated.

In addition to team behaviors, there are several other factors that can impact psychological safety. In lieu of the limited research on industry-based engineering teams, we turn our focus to student teams, as these groups cannot be compared on the same grounds [42]. Importantly, student teams tend to shorter in duration, which could weaken the effectiveness of team development [43]. To investigate these teams, we can apply Salas et al.'s Nine Critical Considerations of Teamwork (9 C's) [44]. Importantly, not all factors apply to our student teams (see 4.4 Content Analysis), therefore the current discussion focuses on role of team composition, communication, coordination, cooperation, and conflict. For our study, team *composition* refers to the individual factors that are relevant to team performance such as knowledge, skills and attitudes, stemming from the team characteristics (personality) construct [12]. For example, the "Big Five" personality traits [45], such as openness and neuroticism, as well as emotional stability, have been posited to be correlated with psychological safety. For example, openness can potentially increase the possibility of an individual feeling safe to take risks [46]. In addition, team *communication* is used to reflect factors such as giving and receiving feedback [12], being respectful to others [12, 47], and listening to each other [37]. This was based on the social interaction and team characteristics constructs, which can impact information sharing [12]. Moreover, *conflict*, which refers to the incompatibility of team members' interests, beliefs, or views, may also impact team psychological safety as it has been shown to have a relationship with team performance [34]. This stemmed from the task conflict construct that is mediated by psychological safety and impacts how teams handle conflict [12]. Interestingly, higher psychological safety can help teams leverage the benefits of conflict to produce creative outputs [48]. Differing from conflict, *coordination* focuses on the participants' behavior patterns when performing tasks, promoting team effectiveness [44], which come from the problem-solving efficacy and team characteristics constructs [12]. Specifically, prior work has shown that coordination can impact the efficiency and success of designs in engineering [49, 50]. Finally, stemming from the team leadership construct, *cooperation* allows teams to develop a problem-solving orientation that promotes leaders and members to discuss and learn from mistakes [12, 51].

In addition to these Critical Considerations of Teamwork, the *cohesiveness* and *creativity* of the team may also affect their psychological safety. *Cohesiveness* refers to interpersonal attraction, commitment to task, and/or group pride [52]. While meta-analysis has pointed to the importance of these factors in relation to psychological safety [13], the results are mixed. For example, where one study indicated that team cohesiveness can reduce the likelihood of interpersonal risk-taking in psychological safety [53], another study linked cohesiveness in teams to feeling safe in sharing opinions [54]. These studies bring to question the impact of team cohesiveness in building psychological safety. Finally, *creativity*, which can emerge as a construct from psychological safety, may also impact team psychological safety [12]. Creativity is heavily emphasized in engineering design [55], however, its impacts on the team psychological safety remain unclear.

While empirical work shows support that the aforementioned factors may impact psychological safety, some important questions remain. Specifically, *how, what,* and *when* should we intervene in team activities to reach higher psychological safety and better team performance? To respond to the research needs, we measured the reliability and applicability of psychological safety in an engineering design project in this study. In addition, we explored the factors that could contribute to the construct of psychological safety, as it has been identified as a key research need [12, 13].

3.0 RESEARCH OBJECTIVES

The purpose of this study was three-fold. First, it sought to understand the longitudinal reliability of the psychological safety construct in engineering student teams. This is critical, as reliability certifies whether a scale is measuring what it claims to measure under specific conditions, such as whether psychological safety can be measured in engineering design. Second, it sought to explore the factors that impact the building or waning of psychological safety in engineering design student teams. Third, it sought to examine which factors affected psychological safety during each of the five Time Points. The results of this study provide a better understanding of *what kind*s of intervention might lead to better team performance, *how* to successfully intervene, and *when* is a better time to intervene. This work also provides some of the first evidence of the reliability of psychological safety in an engineering context.

4.0 METHODOLOGY

To answer the research questions stated as part of this study, a longitudinal study was conducted over the course of two years with 260 participants. See Table 1 for the detailed number of teams and participants in each semester. Data collection started during the Summer 2018 term and concluded in the Spring 2020 term. Over the condensed summer terms, the duration of the study was 4 weeks, whereas over the fall and spring terms, the study was 8 weeks instead, but the activities and timeline remained the same. Importantly, all teams spent the same amount of time together, where the summer sessions met twice as long each week, see Figure 1 for study timeline. The remainder of this section describes the data that was collected as part of the current investigation.

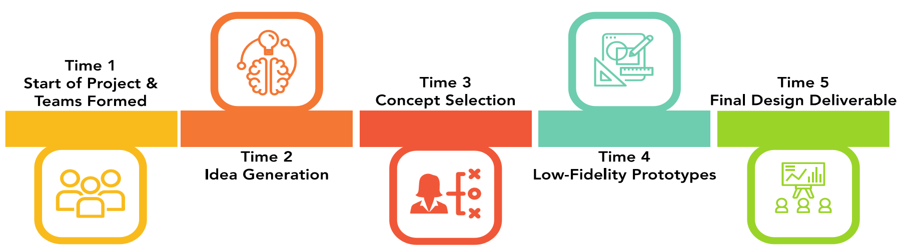

Figure 1: Timeline of the study – psychological safety was captured at the end of each time point.

Table 1: Descriptions of design challenges based on instructor and semester, as well as total number of participants and teams.

Semester	Instructor	Number of Teams	Number of Participants	Project Description
Summer 2018	A	12	48	Tackle food insecurity in developing countries as a result of climate, conflict, unstable markets, food waste, and lack of investment in agriculture
Spring 2019	A and B	13	49	Ensure healthy lives and promote the well-being at all ages through addressing diseases, pollution, and traffic injuries.
Summer 2019	A	12	47	Ensure healthy lives and promote the well-being at all ages through addressing diseases, pollution, and traffic injuries.

Fall 2019	A and C	16	61	Ensure healthy lives and promote the well-being all ages through addressing diseases, pollution, and traffic injuries.
Spring 2020	A and D	15	58	Develop a new water toy for children ages 3 to 5 to teach STEM in a fun, safe, novel way.

4.1 Participants

In total, 260 participants between the ages of 18 to 19 formed into 68 teams and participated in the study (189 males and 71 females), see Table 1 for summary. The participants were all enrolled as engineering majors at a large northeastern university.

4.2 Procedure

The study was completed within a first-year introduction to engineering design class (EDSGN 100). A total of 11 sections of students were involved in the current investigation. While 7 sections took place during a typical semester (8 weeks), 4 sections happened over condensed summer terms (4 weeks). The study data was collected over five Time Points, including (1) start of the project, (2) idea generation, (3) concept selection, (4) prototyping, and (5) final design deliverable (see Figure 2). Specifically, at the end of each time point, a survey including seven questions developed by Edmondson (p. 382 [15]) and two open-ended questions were completed by the participants. The first seven questions centered around the level of how comfortable team members felt when they made mistakes without criticism, provided suggestions with good intentions, as well as whether they felt accepted as a member of the group [15], see Supplementary Materials for survey items. Psychological safety was measured via calculating the average of the scale items from the Team Psychological Safety survey for each individual, and then team scores were aggregated to the team level by taking an average of the team members' individual psychological safety scores. These questions were followed by two open-ended questions, which explicitly asked students to state the positive and negative interactions that had an impact on their answers to the psychological safety survey items. All participants consented at the start of the study (Time Point 1) under the Institutional Review Board guidelines set forth at the university. Importantly, time in between Time Points could last one class session in between to four or five, depending instructors' needs to cover other assignments besides the main project. Each main design session occurred within one class period.

During ***Time Point 1***, participants were assigned to a group of 3 to 4 people based on Kirton's Adaption-Innovation inventory (KAI) [40]. Although not the focus of this paper, half of the teams were constructed to be homogeneous and half of the teams were heterogeneous, as based on the KAI scores. Later, teams engaged in a design challenge, which differed by the instructors of each section, see Table 1. Then, students performed research on the contextual area of the design problem to understand the problem for their design projects. At the end of the class, students completed the first psychological safety survey.

During ***Time Point 2***, students were given a lecture on how to investigate customer needs and how to develop problem statements. Next, instructors introduced a module to the students on the importance of creativity. Later, students were guided through methods of idea generation, and were asked to generate as many ideas as possible in 15 minutes. For the next 10 minutes, the teams were encouraged to brainstorm as a group by making modifications to the ideas generated by each team member. At the end of the session, students took the second psychological safety survey.

During ***Time Point 3***, teams were led through concept selection activities where the participants were asked to assess ideas generated by their respective team members in Time Point 2. Specifically, participants were asked to assess ideas individually in a random order by categorizing the ideas using the individual concept assessment sheets into "Consider" or "Do Not Consider," see Figure 3. Ideas in the "Consider" category were concepts that the participants felt would most likely satisfy the design goals, while the "Do Not Consider" category contains ideas that the participants felt had little to no likelihood of satisfying the design goals. Afterwards, the students were asked to do a final round of screening, which they decided on the final ideas as a group. At the end of the class, all of the participants completed the psychological safety survey.

Idea #	Brief Description of Idea	Is this idea worth considering for further design?	
		Consider	Do Not Consider
1	Plastic sheet with grid	✓	
2	Snap off UTI test strips		✓
1	Air Cannon	☐	☑
3	Delivery Tube	☑	☐
2	Unmanned Helicopter	☑	☐
4	Zipline Supply System	☐	☑

Figure 2: Example Individual Concept Selection Sheet

During ***Time Point 4*,** participants were asked to embody the ideas selected in Time Point 3 by building low-fidelity prototypes for 20 minutes while using commonly available materials (e.g., foam cores, cardstock, post-it notes, etc.). From there, participants were asked to give an "elevator pitch" to their classmates from other teams introducing their prototypes. After their classmates asked some critical questions, such as those related to the technical feasibility and economic impacts, participants gathered the feedback from their classmates and made additional changes to their design. At the end of the class, all students completed the psychological safety survey.

During ***Time Point 5***, teams presented their final design ideas with a PowerPoint presentation and a final rendered prototype using CAD (see Figure 3). At the end of class, students completed the psychological safety survey.

Figure 3: Design evolution of Team 5's ideas, where they designed a floating farm to combat food insecurity problems in Madagascar due to flooding and high wind from cyclones.

4.4 Content Analysis

To analyze the open-ended prompts in the survey, "Please describe any positive/negative team interactions or activities that impacted the rating," a qualitative analysis was conducted. Specifically, prior literature showed that factors such as group dynamics [36], interpersonal relationships [36], communication [38], and personal traits [56] are known to have influence on psychological safety. Consequently, a codebook was developed through an directed content analysis approach [57], building upon prior work, see Supplemental Materials. Specifically, the responses were coded using five of the Critical Considerations for Teamwork Framework (Composition, Communication, Coordination, Cooperation, Conflict) [44], and two additional Critical Considerations (Creativity and Cohesiveness [52]) that are relevant to this study and engineering design teams in general. We call this the "Seven Critical Considerations (7 C's) of Psychological Safety in Engineering Design Teamwork" model. Although psychological safety is derived from an aggregate score for all items [15], items that map to the constructs under investigation probe respondents for how they are able to bring up problems (Conflict and Communication), feel safe to take risks, and make mistakes (Conflict and Cooperation), and difficulty in asking others for help (Cooperation, Communication, and Coordination). Additionally, the scale gets at feeling rejected for being different (Cohesiveness and Composition), not feeling that efforts are undermined (Cohesiveness and Cooperation), and feeling that unique skills and talents are utilized (Creativity and Coordination). After coding the qualitative data into the themes above, we performed further axial-coding and re-grouped the answers into secondary themes.

5.0 RESULTS AND DISCUSSION

This section highlights the results from exploring longitudinal study on psychological safety. During the study, the average psychological safety scores for all 68 teams from Time Point 1 to Time Point 5 were 5.59, 6.02, 6.10, 6.12, 6.07, respectively. The interrater reliability of the qualitative data was assessed via two raters; one PhD Industrial Engineering student and one undergraduate Industrial Engineering student, coded using the same codebook at an overlap of 20% of the open-ended questions using NVivo Pro. An interrater reliability (Cohen's Kappa) of 0.76 was reached. The remainder of this section presents the results in reference to our research questions. The statistical data were analyzed via the SPSS v.26.

RQ1: How applicable and reliable is psychological safety in engineering student teams over time?

Before conducting further analyses with the psychological safety scores, the first step is to ensure that the scale applies to the engineering design team environment. Without this validation, further analyses with survey data and open-ended responses would be subject to usability. To justify the reliability of the psychological safety scale used for RQ1, Cronbach's alpha (α) was computed based on the data collected from five different time points throughout the course for all teams within each time point. Cronbach's alpha measures the internal consistency of a scale by identifying how similar two responses are to randomly selected sets of items within the scale [26]. Generally, a Cronbach's alpha value greater than 0.7 is acceptable, regardless of random error calculated using [24].

The results computed for each of the five time points demonstrate that the scale was reliable since Cronbach's alpha values range from 0.70 to 0.82 across Time Points 1 through 5, as shown in Figure 4. These values were equal to 0.70, 0.72, 0.75, 0.74, and 0.82 for each time point, respectively. Among these values, Time Point 1 exhibited the lowest value at $\alpha = 0.7$. This may be because psychological safety in a team takes time to develop over the course of a project, because it begins at the individual level and manifests at the team level over time [23, 27]. Because Time Point 1 represented the first meeting, this signaled that there may not have been enough time for the perception of psychological safety to manifest. After Time Point 1, the Cronbach's alpha increased over time while the survey length remained the same. Therefore, we can conclude that at Time Point 1 when the team construct might not have been stabilized, it could have caused difficulties in examining the psychological safety in this longitudinal study.

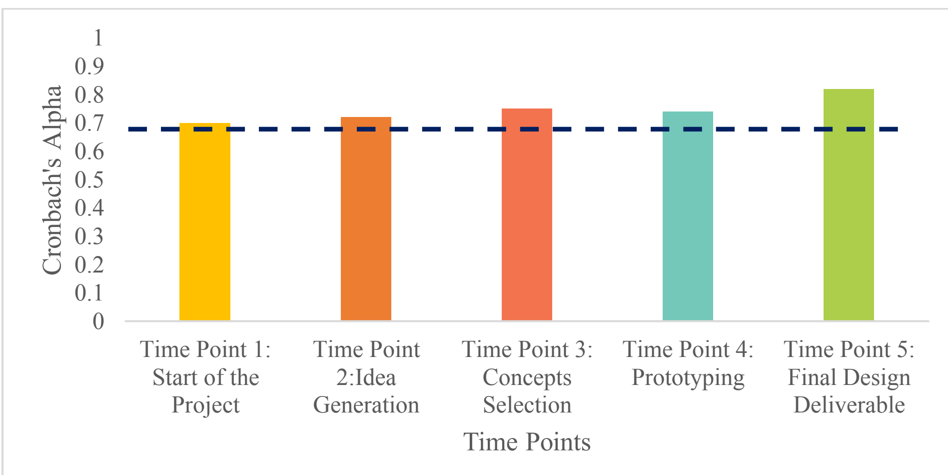

Figure 4: The internal consistency of the psychological safety scale over the five time points as measured by Cronbach's alpha. The dashed line shows the acceptable level of reliability (0.7).

After assessing the internal consistency of the scale's reliability, we also examined consistency of scores among the team members. This is because psychological safety is a shared group property that manifests at the team level [15]. It commences with individual feelings but as more interaction takes place, psychological safety turns into a collective team phenomenon [23, 27]. Therefore, it is necessary to justify aggregating the scale to the team level because psychological safety describes team perceptions instead of individual [15]. When all the individual members in the team agree with each other on the overall psychological safety level, psychological safety can be considered a shared team level construct [27]. As a result, it is important to justify aggregation to the team level through interrater reliability and agreement within the teams to establish psychometric integrity in assessing team psychological safety [25].

The first and more dominant method to justify aggregation is using the measure of interrater agreement with r_{WG} *indices* [22]. The interrater agreement is assessed using r_{WG} because there are multiple members rating the same questions while using the same internal scale [25]. This index compares the actual variance to the expected variance to check whether the team members are answering the psychological safety survey randomly. Therefore, when the members reach a perfect agreement on the questions, the observed variance will equal 0 and will result in an r_{WG} value of 1 [25]. Specifically, the r_{WG} ratio is calculated as such:

$$r_{wg} = 1 - \frac{observed\ group\ variance}{expected\ random\ variance} \quad (1)$$

The commonly accepted value for r_{WG} is 0.70 and above [58]. While the r_{WG} accounts for the agreement of one question, the $r_{WG(J)}$ has been computed to represent a multi-item index since a total of 7 questions are included in the psychological safety survey [15]. Because an $r_{WG(J)}$ value has been calculated for each team, an average and median $r_{WG(J)}$ value have been reported for each of the time points. The average of the $r_{WG(J)}$ ranged from 0.88 to 0.91, and the median of the $r_{WG(J)}$ ranged from 0.91 to 0.94. These results demonstrate that the interrater agreement of the psychological safety level within teams is acceptable. See Table 2 for mean and median $r_{WG(J)}$ values for each time point.

Table 2: Average Team Psychological Safety Descriptive Statistics and Psychometric Properties Across Time

Time	Mean	SD	α	Mean $r_{wg(j)}$	Median $r_{wg(j)}$	ICC(1)	ICC(2)
1: Start of the Project	5.95	.54	.70	.91	.91	.26	.57
2: Idea Generation	6.02	.50	.72	.88	.93	.18	.43
3: Concept Selection	6.10	.49	.75	.89	.94	.13	.35
4: Prototyping	6.12	.46	.74	.88	.92	.08	.24
5: Final Design Deliverable	6.07	.56	.82	.89	.93	.11	.31

The second commonly used measure to justify aggregation that captures both interrater agreement (consensus) and reliability (consistency) are the intraclass correlations (ICCs) [22]. Particularly, ICC values are typically used to represent effect size due to group measurement, whereas r_{WG} tends to be the stronger measure for suggesting aggregation [25]. Two types of intraclass correlations, ICC(1) and ICC(2), were calculated for this study (see Table 2). ICC(1) uses team membership to estimate the total proportion of variability [58], which can be explained by the extent to which one team member could represent every member who took the survey in the group. A larger ICC(1) value indicates a larger degree of interchangeability between the team members—or the extent to which one team member's rating may represent all ratings within the group [58]. In the study, ICC(1)s ranged from 0.08 to 0.26 throughout five different time points. Results from Time Point 1 showed that the psychological safety variance due to team membership was greatest at 26%. While surprising, as we would expect lower ICC(1) values earlier on due to limited team interactions, the low α may play a role in this distinction from other time points. As time progressed, results from Time Point 2 and 3 revealed decreases in ICC(1) values at 18% and 13%, respectively. Then, ICC(1) was lowest at Time Point 4 (8%), suggesting relatively little variability in psychological safety due to team membership, pointing to lack of a shared model of psychological safety. This could be because Time Point 4 represents a later point in the semester, where team members were likely to be working on different tasks related to the main project, such as building prototypes or working on individual assignments (such as CAD homework) [44]. When team members were not expecting to work together on tasks and engaged in "divide and conquer" tactics, the team developed less of a shared mental model over time, which resulted in a small percentage of variability in psychological safety due to team membership [44]. The ICC(1) then recovered at Time Point 5 with an 11% variability in psychological safety scores due to team membership. This increased value exhibited that when the team worked together on the same tasks for the final project, the shared mental model [44] promoted the increase in amount of variance explained by team membership.

While ICC(1) captures the interrater agreement, ICC(2) captures the reliability of the mean ratings [25]. In this case, ICC(2) takes into account how many people there are in a group to denote how reliable the team means would be based on the consensus and consistency of member ratings [58]. Estimation of ICC(2) on the reliability of mean ratings are generally higher in magnitude than ICC(1) estimates because it is adjusted based on group size [22]. Similar to the pattern of ICC(1) results, ICC(2) showed the highest value at Time Point 1, where ICC(2)=0.57. After Time Point 1, the ICC(2) decreased at Time Point 2, 3, and 4 with values of 0.43, 0.35, and 0.24, respectively. Similarly to ICC(1), the ICC(2) also recovered a little at Time Point 5, where ICC(2)= 0.31. Overall, the ICC(1) and ICC(2) values resulted in a decreasing pattern from Time Point 1 to 4, which dropped the most at Time Point 4, where students worked on different tasks, weakening the shared mental model. However, the ICC values slightly recovered at Time Point 5, demonstrating that the students were back to working with each other more so than in Time Point 4.

RQ2: What factors impact the building or waning of psychological safety in an engineering design student team?

To understand the factors that impact the building or waning of psychological safety in student engineering teams, qualitative responses from the open-ended survey questions were analyzed in an exploratory study. Specifically, this research question sought to discover what factors were correlated with higher or lower psychological safety across all teams using our Seven Critical Considerations (7 C's) of Engineering Design Teamwork. The open-ended questions from the survey were analyzed to identify major themes that may have impacted psychological safety. While 2,742 responses were received from participants, 650 were discarded due to a lack of detailed response by the participants (e.g., "Not Applicable", "Nothing to report"), and the 62 "Other" responses that were irrelevant to the questions asked were discarded as well (i.e., "second guessed a lot") totaling 2,030 open-ended responses. Importantly, these responses came from individuals across all teams, therefore they represent individual perceptions of psychological safety.

In all, 7 main discussion topics and 18 subtopics were defined after the investigation of the data. As detailed in Figure 5, "Communication" was referred to the most for positive team interactions while "Coordination" contained the highest frequency count for negative team interactions. The following section presents detailed descriptions and examples of the discussion topics that may have impacted the formation or decline of psychological safety.

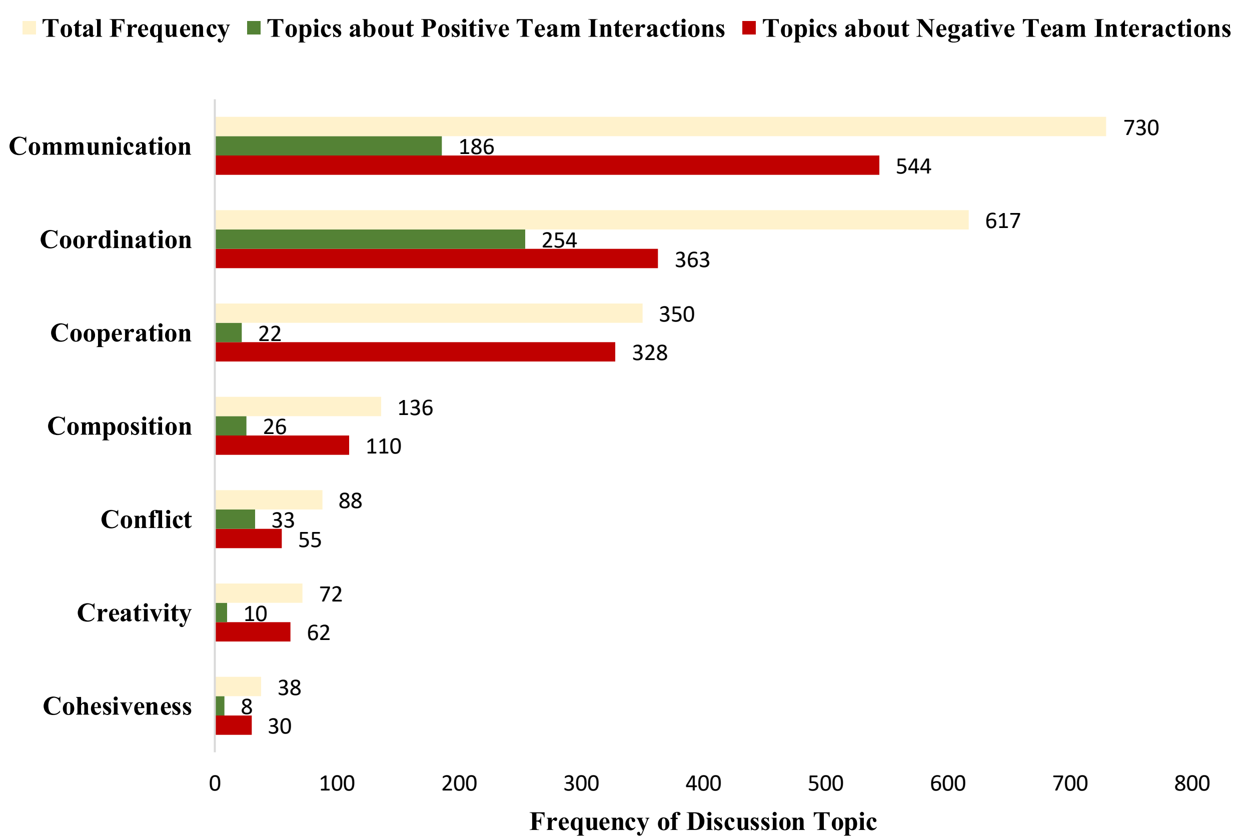

Figure 5: Discussion topics, the total frequency of occurrence, and the number of times the topic contains positive or negative team interactions.

5.1 Communication

The discussion topic that was the most frequently cited interaction or activity that impacted participants' ratings of psychological safety during the five time points was the communication (f = 730) within the team. Three sub-topics in this area were also identified, including: *being respectful of other's ideas* (f = 136), *listening* (f = 63), and *general communication* (f = 531). Specifically, under the *general communication* subtopic, where members provided and received information which led to the attitudinal, behavioral, and cognitional change of a team, teams indicated whether they were able to communicate well throughout the process. There were usually no specific details to explain the communication as presented in the survey responses, but rather it was stated in a general way under this subtopic. For example, a participant in Team 32 expressed that, "We communicated very well without hurting each other's feelings when [we] chose not to consider an idea." Along with general communication, *being respectful to other's ideas* is a topic that participants frequently mentioned. Particularly during *concept screening*, this respectfulness was mentioned frequently. For example, a participant from Team 38 commented that, "Everything we do is respectful and considerate of the feelings and ideas of others. Supportive language is used almost always." Participants also mentioned *listening* as another subtopic. Under *listening*, team members paid attention and did not interrupt others. A negative example of this came from Team 40 in the quote, "Sometimes I had to repeat myself a few times to be listened to." Communication was the most-often occurring positive interaction in our sample (f = 544).

5.2 Coordination

The second most frequently discussed topic was coordination or team behavioral mechanism that can cause team recourses to result [44] (f = 617). Discussion about coordination has been categorized into six subtopics, including *efficiency* (f = 158), *work contribution & participation* (f = 185), *task completion* (f = 155), *time management* (f = 57), *absence & lateness* (f = 32), and *not explicitly stated* (f = 30). Specifically, under coordination, one of the most frequent topics was when participants discussed they had accomplished tasks in an efficient manner. For example, a participant in Team 30 indicated that "We did work very efficiently today…" Another topic that participants brought up many times was whether group members actively participated in the activities or whether everybody contributed the same effort without having an idle team member not doing anything. For example, a participant in Team 26 commented, "Not evenly split up the jobs." Meanwhile, a participant in Team 31 indicated that "Everyone contributes and everyone puts in their work." Additionally, students commented about how their teams accomplished their tasks in a timely manner. For example, a participant from Team 13 stated that, "We waited last minute for everything."

These results indicate that task performance was positively related to psychological safety [13] in an engineering context. Also, our qualitative results agree with the notion that problem-solving efficacy also correlates with psychological safety in group-level research [12]. More specifically, the results suggest that coordination including team efficiency, members' contribution and participation, task accomplishment, time management, and the absence and punctuality of team members were likely to be linked to the formation or decline of psychological safety.

5.3 Cooperation

The third topic participants discussed the most involved cooperation. During team activities, cooperation is a motivation driver for a team, which the team builds upon with attitude and beliefs [44]. Cooperation had been discussed in a high frequency (f = 350). More specifically, there was a higher frequency that a participant mentions cooperation when asked about positive interaction (f = 328) than negative interaction (f = 22). For example, a participant in Team 27 mentioned that, "We [are] all able to collaborate our ideas to make a great product." Our results show that collaborating well, working together, and helping each other could impact the formation of psychological safety. These results also demonstrate the existence of a relationship between cooperation and psychological safety [51].

5.4 Composition

The fourth most frequent topic involved team composition (f = 136). Discussions about the composition included participants commenting on the characteristics of their team members and the individual factors that are related to the outcome of the team performance [44]. These were broken down into two categories: *personal traits* (f = 94) and *not explicitly stated* (f = 42). Usually, participants discussed two of the personality traits out of the "Big Five," [45] which are openness (f = 59) and extraversion (f = 35). During the team activities, especially for Time Points 2 and 3, these activities required participants to share their own and evaluate others' ideas. For example, a participant from Team 49 discussed that, "Everybody is open-minded to new ideas." These results suggest that having open and extroverted team members (f = 85) could promote the building of psychological safety. In comparison to the frequency of positive composition interactions (f = 110), negative interactions were relatively small (f = 26).

5.5 Conflict

The fifth most frequent topic that participants discussed was conflict. Conflict describes the incompatibility of ideas and perspectives regarding the tasks and assignments [44]. Throughout the data, teams mentioned reaching consensus in making decisions when providing answers for positive interactions (f = 55). They also mentioned that conflict occurred when team members possessed different views (f = 33) could also impact their ratings that decreased their psychological safety scores. A positive response from participant in Team 37 was stated, "We could all agree on what ideas were good and what ideas were bad." In contrast with other topics which had relatively huge differences between the frequency of positive and negative interactions, frequency counts for "conflict" for both cases were relatively close. The results indicate that there may be a connection between having team conflict and the climate of psychological safety, and conflict could act like a double-edged sword when it comes to psychological safety. This is similar with the relevant research which stated that conflict is an important factor when it comes to improving team performance [34], as psychological safety is also positively related to team performance [15].

5.6 Creativity

The sixth most frequent topic involved discussions on the creativity of the ideas. Starting at Time Point 2 (Idea Generation), students were told by the instructors not to hold back from generating any ideas that came to mind in order to inspire more creativity during the process. Some participants mentioned team members generating creative ideas as positive interactions (f = 62) for building psychological safety. For example, a participant from Team 28 commented, "We came up with many unique and good ideas." While infrequent, participants complained about lack of creative ideas as a negative interaction (f = 10).

5.7 Cohesiveness

The final, least frequently discussed topic was the cohesiveness in the group (f = 38); interpersonal attraction, commitment to task, and/or group pride [52] when team members worked towards a common goal. Specifically, there were some participants who talked about having good or poor connections with their teammates when discussing positive and negative interactions. For instance, a participant from Team 26 commented, "We are able to laugh and get closer to each other." Prior research demonstrates that group cohesiveness is distinct from psychological safety, as team cohesiveness could reduce the likelihood of interpersonal risk-taking [53], as team psychological safety focuses on whether individuals consider themselves safe for risk-taking [15].

RQ3: What factors impact the building or waning of psychological safety in an engineering design student team at each of the five time points?

In addition to reviewing the prevalence of these factors throughout the project, we also wanted to explore how these factors impacted psychological safety at each Time Point. To achieve this, 14 linear regressions were applied to understand the effect of the number of positive and negative comments for each of the 7 C's at each of the five time points, see Table 3 for summary. This allowed us to investigate what factors contribute to lower or higher psychological safety during an engineering design project. Prior to conducting this analysis, assumptions were checked. Specifically, scatterplots were used to verify the linearity of the variables. In addition, homoscedasticity and normality of the residuals was verified, and no outliers were found. To account for the potential inflation of Type I error, a Bonferroni correction was applied [59]. Specifically, the typical p=0.05 value was divided by 14 (number of linear regressions), giving an adjusted p-value of p=0.00357 to define statistical significance. Importantly, team-level aggregations of psychological safety were used, with the degrees of freedom reported in the following paragraph. Importantly, Figure 6 shows the breakdown of these topics across all time points for all individuals, however the frequencies were normalized within each team and related back to team psychological safety scores for the analysis.

Table 3: Statistically Significant Factors that Contributed to the Building or Waning of Psychological Safety over Time

Time Point	Factor	*Constant*	*B*	*SE B*	*β*	*t*	*p*	*df*	*Adj. R^2*
1. Start of Project	Communication (-)	6.076	-1.166	.301	-.470	-3.876	<.001	53	20.6%
2. Idea Generation	Communication (+)	5.730	.553	.159	.428	3.485	<.001	54	16.8%
	Cohesiveness (-)	6.069	-5.717	1.815	-.394	-3.150	.003	54	14.0%

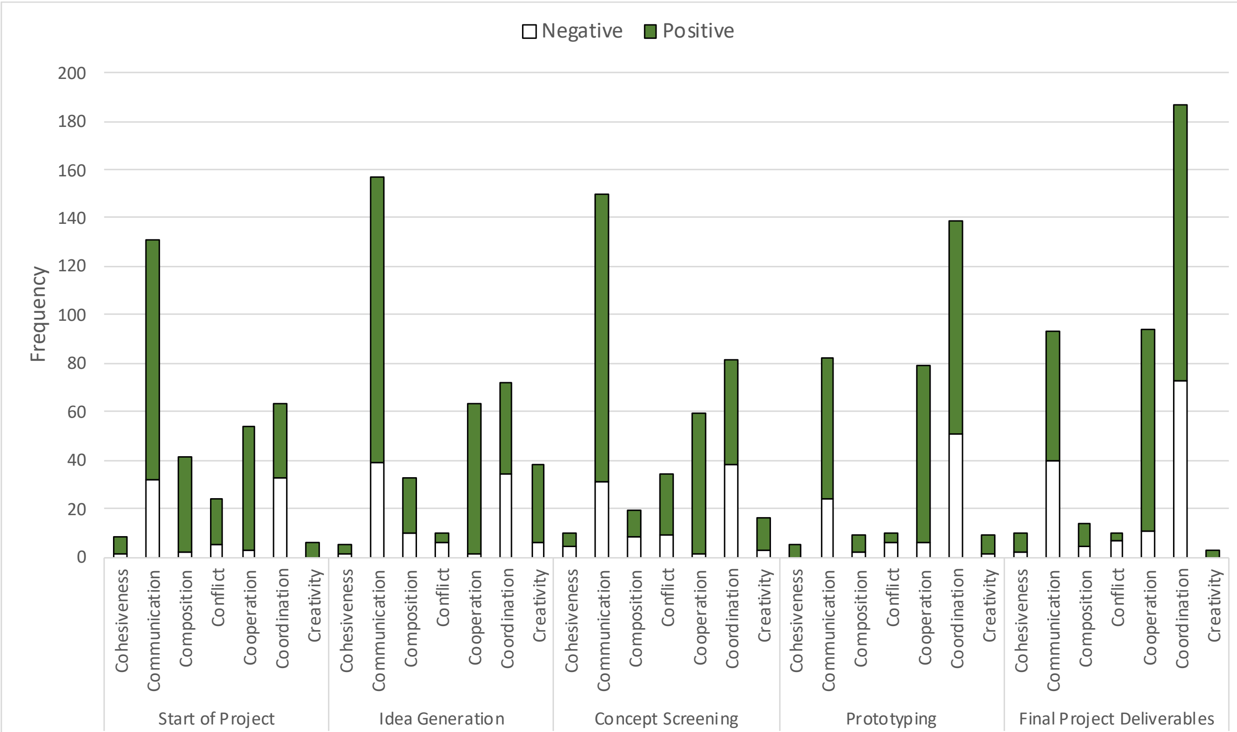

Figure 6: Frequencies of topics (positive and negative) that impacted ratings of psychological safety throughout the trajectory of the project.

The results of the regression analyses at Time Point 1 (start of the project) revealed one significant prediction variables. Specifically, the normalized negative conflict comments during the start of the project statistically significantly predicted team psychological safety at Time Point 1, $F(1,53) = 15.020$, $p < 0.0005$, accounting for 22.1% of the variation in team psychological safety with adjusted $R^2 = 20.6\%$; a medium effect size. The results of the regression analyses at Time Point 2 (concept generation) revealed two significant prediction variables. Specifically, the normalized positive communication comments a team documented statistically significantly predicted team psychological safety, $F(1,54) = 12.143$, $p=0.001$, accounting for 18.4% of the variation in team psychological safety with adjusted $R^2 = 16.8\%$; a medium effect size. In addition, the normalized cohesiveness negative comments statistically significantly predicted team psychological safety, $F(1,54) = 9.924$, $p=0.003$, accounting for 15.5% of the variation in team psychological safety with adjusted $R^2 = 14.0\%$; a medium effect size. Finally, the results of the regression analyses at Time Point 3 (concept screening), Time Point 4 (prototyping), and Time Point 5 (final design deliverables) failed to reveal any significant prediction variables.

These results are important because they demonstrate what factors could affect team psychological safety during each of the five time points, such as negative communication and conflict comments being linked to lower psychological safety during prototyping. The number of negative communication comments was significant during Time Point 1 (start of project), with positive communication comments significant during Time Point 2 (concept generation). Additionally, negative cohesiveness comments at Time Point 2 were also significant. Based on these results, communication had the largest impact on psychological safety within this study, particularly at the earlier stages. This demonstrates the importance of promoting communication early in a team.

6.0 DISCUSSION, LIMITATIONS, AND FUTURE WORK

The main objective of this study was to understand how applicable and reliable the measure of psychological safety was in *engineering* student teams *over time.* In addition, we sought to identify potential factors that impact the building and waning of psychological safety in these teams over time. The main findings from this study were as follows:

- The psychological safety scale was reliable for all time points, where Cronbach's alpha values ranged from 0.70 to 0.82, with the lowest at Time Point 1.
- Psychological safety scores among teams were consistent with ICC(1) and ICC(2) being highest at Time Point 1 at 0.26 and 0.57, respectively.
- In total, *communication* contained the highest number of negative comments, whereas *coordination* contained the highest number of positive comments.
- Negative communication was significantly negatively related to psychological safety at Time Point 1.
- At Time Point 2, positive communication was significantly positively related to psychological safety, and negative cohesiveness was significantly negatively related to psychological safety.

Cronbach's alpha values met the threshold for reliability at all time points and the intraclass correlations, ICC(1) and ICC(2), were highest at Time Point 1 at 0.26 and 0.57, respectively. This indicates that the amount of variance explained by team membership (ICC(1)) [25], and thus the extent to which mean ratings can be distinguished reliably among groups (ICC(2)) [25], was highest at the first time point. Following Time Point 1, ICC(1) and ICC(2) gradually decreased over time until Time Point 4 (prototyping), and then recovered at Time Point 5 (final design deliverables). This demonstrates that the shared team model of psychological safety weakened over time, but recovered slightly at the last time point. This may be related to teams making an effort to "get on the same page" as Time Point 5 mainly focuses on the processes of turning in the remaining the final project deliverables. While we may not expect such a decrease at Time Point 4 due to dominant personalities potentially taking over and teams coming together prior to sharing low-fidelity prototypes, all prototyping activities were condensed as one time point, so we lack the granularity of analysis to make any statement on its effects. Conversely, *coordination* work had the highest number of positive and negative comments at Time Point 5, so coordination may contribute increase in the shared team model. This aligns with prior work, which showed that team members working on different assignments separately can weaken the shared mental model of a team [44], decreasing interrater agreement and reliability. Findings also indicate that sometimes the survey may not be reliable throughout the design process, limiting to what extent psychological safety exists as a team construct.

In addition to the IRR and IRA analyses, the qualitative analyses helped to identify some interactions occurring within the teams. Specifically, positive coordination and negative communication comments were the highest overall, emphasizing the importance of coordination for being successful in an engineering design project, and how poor communication can hurt a team's chances of success. As problem-solving and evaluation activities rely on good *communication*, whether that be through words, diagrams or numbers [18]. Conversely, coordination is critical to producing design outputs in engineering both efficiently and successfully [49, 50]. Our findings are further substantiated by the statistical analyses of the normalized comments per team, where normalized positive communication comments were correlated with an increase in psychological safety at Time Point 2 (concept generation). This aligns with prior work, as *communication* during the generation stage is necessary for being able to share new ideas or alternatives to tackle a design problem [18]. However, normalized negative communication comments at Time Point 1 (start of project) were correlated with lower psychological safety, implying that lack of discussion within teams made it difficult for team members to feel psychologically safe with one another. This can be detrimental to both performance and developing a strong, shared model of psychological safety, as good communication is necessary discussing problems effectively [12, 15]. Furthermore, it helps to mitigate the negative impacts of diversity that are inconducive to a safe team atmosphere [12]. Therefore, future work should look to developing ways of fostering more communication and discussions within teams at the beginning of a team project. Finally, the third analysis showed that the normalized number of negative cohesiveness comments were correlated with lower psychological safety at Time Point 2. While contradictory to some prior work [53], our findings suggest that lack of interpersonal attraction, lack of task commitment, and not doing things as a team can lower team psychological safety. Thus, these findings compliment prior work that demonstrated the importance of cohesiveness in overcoming the fear of risk in sharing individual opinions [54].

While this study was exploratory in nature, the results help us begin to answer some difficult fundamental questions on supporting team performance in engineering education, such as; *when* do teams need interventions to support team performance, *what* types of interventions might be successful, and *how* can we successfully intervene in these scenarios? These are complex questions, however, that the current study only begins to shed light on. Other factors contributing to team performance, such as team expertise, personality, and interaction behaviors also need to be considered.

Although the current study presents some of the first evidence on the reliability of psychological safety in engineering teams and is among the first to examine the longitudinal trajectory of psychological safety, there are several limitations that suggest directions for future work. First, while actual psychological safety scores were applied to the analyses for RQ2 and RQ3, RQ1 just focuses on the validation process due to the varying trajectories of teams' scores, limiting our ability to make any generalizations on psychological safety score trends over time. In fact, preliminary graphs showed lack of any trend in the building/waning of psychological safety among teams. Importantly, while the ICC values seem low on the surface, these values focus on the effect size due to team membership [25], meaning that less team interaction due to dividing tasks is the most probable cause for low ICC values. Furthermore, the study does not include important drivers of the building or waning of psychological safety—such as the composition of the team in the form of cognitive style (deep-level diversity) [40, 41], and gender and racial background (surface-level diversity) [39], for example. In addition, the study was conducted in a cornerstone engineering design class. Thus, the results may only be applicable to this population. Finally, qualitative studies tend to contain selection biases based on the sample [60], but we feel that impacts on the data would be minimal, as the survey items did not explicitly contain words from the 7 C's, such as communication, in the questions. However, we acknowledge that the word "interaction" in the open-ended question may be biasing, eliciting more responses about communication, for example. Further biasing could come from teaching styles, although the focus is on team interactions, therefore we would expect minimal influence from this. Importantly, these findings are limited to engineering design student teams, as industry teams tend to differ in terms of motives [42], consequently affecting perceptions of positive and negative interactions. Finally, students tend to work outside the classroom on tasks, however how teams divided tasks was unreported and could impact psychological safety. Future work will be geared toward exploring the trajectory of psychological safety on a larger scale, including a larger range of design problems and working preferences.

Along with expanding the study to a larger scale, it is also important to consider how psychological safety impacts team design outputs to justify investigation. Particularly, while prior work has identified the importance of psychological safety in enabling creativity and providing critical feedback [16, 30], the current study did not study the impact of psychological safety on design process outputs. Therefore, future work will also examine the role of psychological safety on engineering design team performance throughout the course of a design project. Taken together, these new investigations will lead us even closer to understanding how to support students within team design activities through intervention methods that promote performance.

7.0 CONCLUSION

The aim of this study was to understand the longitudinal reliability of the psychological safety construct in engineering student teams. In addition, we tried to seek insights on what factors could impact the building or waning of the psychological safety in the 68 teams. The results of this study leave several implications for engineering classroom instruction and engineering design research. First, our results show that psychological safety can be reliably measured throughout the trajectory of an engineering design project, however it is more likely to be unreliable towards the beginning. Particularly, the shared team model of psychological safety had a tendency to decrease over time, which implies that individual perceptions of psychological safety tend to differ more within the teams. Based on these findings, we suggest implementing methods to encourage strong communication within teams to develop an understanding of and promote a shared model of psychological safety. Second, our results show that communication and coordination play an important role in the perceived success and failure of teams' abilities in working. This complements prior work that emphasizes the importance of communication [18] and coordination [49, 50] in engineering design. Finally, statistical analyses showed the impact of communication on psychological safety at the start of the project and concept generation, as well as the impact of cohesiveness on psychological safety during concept generation. These results emphasize the need to ensure that good communication can be maintained to solve problems effectively [12], and demonstrate the importance of cohesiveness in feeling psychologically safe to share individual opinions, aligning with prior findings [54]. As such, these findings provide a framework for future intervention methods in promoting psychological safety in engineering design teams over a design project trajectory.

ACKNOWLEDGMENTS

Special thanks are given to Zibing (Janice) Gong in helping with the analysis of the reliability of the psychological safety scale and Katie Heininger in collecting data for Summer 2018. This study is based upon work supported by the National Science Foundation under Grant No. 1825830.

REFERENCES

[1] Cohen, S. G., and Bailey, D. E., 1997, "What makes teams work: Group effectiveness research from the shop floor to the executive suite," Journal of management, 23(3), pp. 239-290.

[2] Mathieu, J., Maynard, M. T., Rapp, T., and Gilson, L., 2008, "Team effectiveness 1997-2007: A review of recent advancements and a glimpse into the future," Journal of management, 34(3), pp. 410-476.
[3] Taggar, S., 2002, "Individual creativity and group ability to utilize individual creative resources: A multilevel model," Academy of Management Journal, 45(2), pp. 315-330.
[4] Salas, E., Cannon-Bowers, J. A., and Johnston, J. H., 1997, "How can you turn a team of experts into an expert team?: Emerging training strategies," Naturalistic decision making, pp. 359-370.
[5] Marks, M. A., Zaccaro, S. J., and Mathieu, J. E., 2000, "Performance implications of leader briefings and team-interaction training for team adaptation to novel environments," Journal of Applied Psychology, 85(6), p. 971.
[6] Hargadon, A. B., 2002, "Brokering knowledge: Linking learning and innovation," Research in Organizational behavior, 24, pp. 41-85.
[7] Ayag, Z., and Ozdemir, R. G., 2009, "A hybrid approach to concept selection through fuzzy analytic network process," Computers & Industrial Engineering, 56(1), pp. 368-379.
[8] Dym, C. W., JW; Winner, L, 2003, "Social Dimensions of Engineering Designs: Observations from Mudd Design Workshop III," Journal of Engineering Education, 92(1), pp. 105-107.
[9] Dutson, A. J., Todd, R. H., Magleby, S. P., and Sorensen, C. D., 1997, "A Review of Literature on Teaching Engineering Design Through Project ‐ Oriented Capstone Courses," Journal of Engineering Education, 86(1), pp. 17-28.
[10] Froyd, J. E., 2005, "The engineering education coalitions program," Educating the engineer of 2020: Adapting engineering education to the new century, pp. 82-97.
[11] McGourty, J., Shuman, L., Besterfield-Sacre, M., Atman, C., Miller, R., Olds, B., Rogers, G., and Wolfe, H., 2002, "Preparing for ABET EC 2000: Research-based assessment methods and processes," International Journal of Engineering Education, 18(2), pp. 157-167.
[12] Edmondson, A. C., and Lei, Z., 2014, "Psychological Safety: The History, Renaissance, and Future of an Interpersonal Construct," Annual Review of Organizational Psychology and Organizational Behavior, 1(1), pp. 23-43.
[13] Frazier, M. L., Fainshmidt, S., Klinger, R. L., Pezeshkan, A., and Vracheva, V., 2017, "Psychological safety: A meta ‐ analytic review and extension," Personnel Psychology, 70(1), pp. 113-165.
[14] Salas, E., Cooke, N. J., and Rosen, M. A., 2008, "On teams, teamwork, and team performance: Discoveries and developments," Human factors, 50(3), pp. 540-547.
[15] Edmondson, A., 1999, "Psychological safety and learning behavior in work teams," Administrative science quarterly, 44(2), pp. 350-383.
[16] Edmondson, A. C., 2003, "Speaking Up in the Operating Room: How Team Leaders Promote Learning in Interdisciplinary Action Teams," Journal of Management Studies, 40(6), pp. 1419-1452.
[17] Burningham, C., and West, M. A., 1995, "Individual, Climate, and Group Interaction Processes as Predictors of Work Team Innovation," Small Group Research, 26(1), pp. 106-117.
[18] Dym, C. L., and Little, P., 2014, Engineering design: A project-based introduction, John Wiley and sons.
[19] Duhigg, C., 2016, "What Google learned from its quest to build the perfect team," The New York Times Magazine.
[20] Kessel, M., Kratzer, J., and Schultz, C., 2012, "Psychological Safety, Knowledge Sharing, and Creative Performance in Healthcare Teams," Creativity and Innovation Management, 21(2), pp. 147-157.
[21] Kozlowski, S. W., Chao, G. T., Grand, J. A., Braun, M. T., and Kuljanin, G., 2013, "Advancing multilevel research design: Capturing the dynamics of emergence," Organizational Research Methods, 16(4), pp. 581-615.
[22] Levesque, L. L., Wilson, J. M., and Wholey, D. R., 2001, "Cognitive Divergence and Shared Mental Models in Software Development Project Teams," Journal of Organizational Behavior, 22(2), pp. 135-144.
[23] Mohammed, S., Ferzandi, L., and Hamilton, K., 2010, "Metaphor No More: A 15-Year Review of the Team Mental Model Construct," Journal of Management, 36(4), pp. 876-910.

[24] Tavakol, M., and Dennick, R., 2011, "Making sense of Cronbach's alpha," Int J Med Educ, 2, pp. 53-55.
[25] LeBreton, J. M., and Senter, J. L., 2008, "Answers to 20 questions about interrater reliability and interrater agreement," Organizational research methods, 11(4), pp. 815-852.
[26] Cronbach, L. J., 1951, "Coefficient alpha and the internal structure of tests," Psychometrika, 16(3), pp. 297-334.
[27] Kozlowski, S. W., and Klein, K. J., 2000, "A multilevel approach to theory and research in organizations: Contextual, temporal, and emergent processes."
[28] Beigpourian, B., Luchini, F., Ohland, M., and Ferguson, D., "Psychological Safety as an Effective Measurement in Engineering Classrooms," Proc. 2019 ASEE Annual Conference and Exposition.
[29] Edmondson, A. C., Kramer, R. M., and Cook, K. S., 2004, "Psychological safety, trust, and learning in organizations: A group-level lens," Trust and distrust in organizations: Dilemmas and approaches, 12, pp. 239-272.
[30] West, M. A., 1990, "The social psychology of innovation in groups."
[31] Dylla, N., 1991, "Thinking Methods and Procedures in Mechanical Design,"Dissertation, Technical University of Munich.
[32] Nikander, J. B., Liikkanen, L. A., and Laakso, M., 2014, "The preference effect in design concept evaluation," Design Studies, 35(5), pp. 473-499.
[33] King, A. M., and Sivaloganathan, S., 1999, "Development of a Methodology for Concept Selection in Flexible Design Strategies," Journal of Engineering Design, 10(4), pp. 329-349.
[34] Bradley, B. H., Postlethwaite, B. E., Klotz, A. C., Hamdani, M. R., and Brown, K. G., 2012, "Reaping the benefits of task conflict in teams: the critical role of team psychological safety climate," The Journal of applied psychology, 97(1), pp. 151-158.
[35] Jonassen, D. H., and Cho, Y. H., 2011, "Fostering Argumentation While Solving Engineering Ethics Problems," Journal of Engineering Education, 100(4), pp. 680-702.
[36] Kahn, W. A., 1990, "Psychological Conditions of Personal Engagement and Disengagement at Work," The Academy of Management Journal, 33(4), pp. 692-724.
[37] Castro, D. R., Anseel, F., Kluger, A. N., Lloyd, K. J., and Turjeman-Levi, Y., 2018, "Mere listening effect on creativity and the mediating role of psychological safety," Psychology of Aesthetics, Creativity, and the Arts, 12(4), pp. 489-502.
[38] Siemsen, E., Roth, A. V., Balasubramanian, S., and Anand, G., 2009, "The Influence of Psychological Safety and Confidence in Knowledge on Employee Knowledge Sharing," Manufacturing & Service Operations Management, 11(3), pp. 429-447.
[39] Mohammed, S., and Angell, L. C., 2004, "Surface- and deep-level diversity in workgroups: examining the moderating effects of team orientation and team process on relationship conflict," Journal of Organizational Behavior, 25(8), pp. 1015-1039.
[40] Kirton, M., 1976, "Adaptors and innovators: A description and measure," Journal of Applied Psychology, 61(5), pp. 622-629.
[41] Kirton, M. J., 2011, Adaption-Innovation in the Context of Diversity and Change, Routledge, London, UK.
[42] Passow, H. J., and Passow, C. H., 2017, "What Competencies Should Undergraduate Engineering Programs Emphasize? A Systematic Review," Journal of Engineering Education, 106(3), pp. 475-526.
[43] Zwikael, O., and Unger-Aviram, E., 2010, "HRM in project groups: The effect of project duration on team development effectiveness," International Journal of Project Management, 28(5), pp. 413-421.
[44] Salas, E., Shuffler, M. L., Thayer, A. L., Bedwell, W. L., and Lazzara, E. H., 2015, "Understanding and Improving Teamwork in Organizations: A Scientifically Based Practical Guide," Human Resource Management, 54(4), pp. 599-622.
[45] Judge, T. A., Higgins, C. A., Thoresen, C. J., and Barrick, M. R., 1999, "The Big Five Personality Traits, General Mental Ability, and Career Success Across the Life Span," Personnel Psychology, 52(3), pp. 621-652.

[46] Edmondson, A. C., and Mogelof, J. P., 2006, "Explaining psychological safety in innovation teams: organizational culture, team dynamics, or personality?," Creativity and innovation in organizational teams, Psychology Press, pp. 129-156.
[47] Carmeli, A., and Gittell, J. H., 2009, "High-quality relationships, psychological safety, and learning from failures in work organizations," Journal of Organizational Behavior, 30(6), pp. 709-729.
[48] Kostopoulos, K. C., and Bozionelos, N., 2011, "Team Exploratory and Exploitative Learning: Psychological Safety, Task Conflict, and Team Performance," Group & Organization Management, 36(3), pp. 385-415.
[49] Girard, P., Robin, V., and Barandiaran, D., "Analysis of collaboration for design coordination," Proc. 10th ISPE International Conference on Concurrent Engineering : Research and Applications - CE2003, J. Cha, R. Goncalves, A. Steiger - Garção, pp. pp. 607 -614.
[50] Coates, G., Duffy, A. H. B., Whitfield, I., and Hills, W., 2004, "Engineering management: operational design coordination," Journal of Engineering Design, 15(5), pp. 433-446.
[51] Tjosvold, D., Yu, Z. y., and Hui, C., 2004, "Team Learning from Mistakes: The Contribution of Cooperative Goals and Problem ‐ Solving," Journal of Management Studies, 41(7), pp. 1223-1245.
[52] Mullen, B., and Copper, C., 1994, "The relation between group cohesiveness and performance: An integration," Psychological Bulletin, 115(2), pp. 210-227.
[53] Edmondson, A. C., 2002, Managing the risk of learning: Psychological safety in work teams, Cambridge, MA: Division of Research, Harvard Business School.
[54] Üçok, D. I., and Torun, A. A., 2016, "The Relationship of Group Cohesiveness, Psychological Safety, Control Over Work, and Competitive Work Environment with Organizational Silence: The Mediating Role of Motives of Silence," International Journal of Business, Economics and Management Perspectives, 1(4), pp. 62-80.
[55] Thompson, G., and Lordan, M., 1999, "A review of creativity principles applied to engineering design," Proceedings of the Institution of Mechanical Engineers, Part E: Journal of Process Mechanical Engineering, 213(1), pp. 17-31.
[56] Zhang, Y., Fang, Y., Wei, K.-K., and Chen, H., 2010, "Exploring the role of psychological safety in promoting the intention to continue sharing knowledge in virtual communities," International Journal of Information Management, 30(5), pp. 425-436.
[57] Hsieh, H.-F., and Shannon, S. E., 2005, "Three Approaches to Qualitative Content Analysis," Qualitative Health Research, 15(9), pp. 1277-1288.
[58] Klein, K. J., and Kozlowski, S. W. J., 2000, "From Micro to Meso: Critical Steps in Conceptualizing and Conducting Multilevel Research," Organizational Research Methods, 3(3), pp. 211-236.
[59] Armstrong, R. A., 2014, "When to use the Bonferroni correction," Ophthalmic and Physiological Optics, 34(5), pp. 502-508.
[60] Collier, D., and Mahoney, J., 1996, "Insights and Pitfalls: Selection Bias in Qualitative Research," World Politics, 49(1), pp. 56-91.

Figure Captions List

Table Captions List

Supplementary Material: Image of Electronically Delivered Psychological Safety Survey Items and Open-Response Questions

Please select a response to each of the items below on how you feel it represents your feelings:

	1 very inaccurate	2	3	4	5	6	7 Very accurate
If you make a mistake on this team, it is often held against you.	O	O	O	O	O	O	O
Members of this team are able to bring up problems and tough issues.	O	O	O	O	O	O	O
People on this team sometimes reject others for being different.	O	O	O	O	O	O	O
It is safe to take a risk on this team.	O	O	O	O	O	O	O
It is difficult to ask other members of this team for help.	O	O	O	O	O	O	O
No one on this team would deliberately act in a way that would undermine my efforts.	O	O	O	O	O	O	O
Working with the members of this team, my unique skills and talents are valued and utilized.	O	O	O	O	O	O	O

Please describe any POSITIVE team interactions or activities that impacted your ratings. Be specific.

Please describe any NEGATIVE team interactions or activities that impacted your ratings. Be specific.

Supplementary Material: Description of the Seven Critical Considerations (7 C's) of Psychological Safety in Engineering Design Teamwork

Critical Consideration	Description	Example
Composition	"The individual factors relevant to team performance; what constitutes a good team member; what is the best configuration of team member knowledge, skills, and attitudes (KSAs); and what role diversity plays in team effectiveness." ([44] p. 5).	"My teammates were all kind and nice." "We all are a little stubborn with our own ideas, myself included. For our group it's hard to admit your idea is wrong."
Communication	"A reciprocal process of team members' sending and receiving information that forms and re-forms a team's attitudes, behaviors, and cognitions." ([44]. p. 5).	"Today our team was a little more outspoken. We are getting more comfortable interacting and sharing our opinions." "Didn't feel as if ideas were heard as much."
Coordination	"The enactment of behavioral and cognitive mechanisms necessary to perform a task and transform team resources into outcomes." ([44] p. 5).	"We have equally shared the work load …when researching info on our topic." "We waited last minute for everything."
Creativity	The response refers to the quality and/or of quantity of ideas generated in a team.	"We came up with many unique and good ideas." "Some of the team are still unaware if their ideas are good or bad or not some keep to themselves. Not enough of expanded idea thinking here."
Cooperation	"The motivational drivers of teamwork. In essence, this is the attitudes, beliefs, and feelings of the team that drive behavioral action." ([44]. p. 5).	"We all collaborated in way that successfully helped the design process of our project." "We were doing our own thing not collaborating."
Conflict	"The perceived incompatibilities in the interests, beliefs, or views held by one or more team members." ([44] p. 5).	"We came to a consensus on what we should do." "It seems we had a split decision on ideas..."
Cohesiveness	The response refers to interpersonal attraction, commitment to task, and/or group pride [52] when team members work towards a common goal.	"We are able to joke around and have good chemistry. I can sense that we will perform well in the project and am excited to begin the project with this team." "It was very hard to explain what we were actually trying to achieve to other members of the group. I want this project to be successful to my standards and everyone elses, and I feel that sometimes our standards are not high enough as a group."
Not Applicable	The response indicates that neither a positive nor negative interaction/event occurred that impacted their ratings.	"None" "N/A"