How Long Until We Are (Psychologically) Safe? A Longitudinal Investigation of Psychological Safety in Virtual Engineering Design Teams in Education

Courtney Cole, Abigail O'Connell, Jacqueline Marhefka, Kathryn Jablokow, Susan Mohammad, Sarah Ritter, and Scarlett Miller
The Pennsylvania State University, USA
Corresponding author email address: scarlettmiller@psu.edu

This paper investigates team psychological safety (N=34 teams) in a synchronous online engineering design class spanning 4 weeks. While work in this field has suggested that psychological safety in virtual teams can facilitate knowledge-sharing, trust among teams, and overall performance, there have been limited investigations of the longitudinal trajectory of psychological safety, when the construct stabilizes in a virtual environment, and what factors impact the building of psychological safety in virtual teams. The results of this study identified that the construct of psychological safety took more time to become a reliable construct in virtual design teams, but once it stabilized, it did not change. Additionally, qualitative findings point to issues with communication and conflict across various stages of the design process in the development of psychological safety. Finally, we identify potential interventions to enhance team mental model development in the early phases of virtual teaming to support team psychological safety.

Introduction

What helps teams to remain effective during a worldwide pandemic? The COVID-19 pandemic has forced us to explore this question as education had to shift to remote formats, relying on conference call applications for events from classroom lectures to proctored exams [1]. Mixed-methods research has shown that this adjustment to online learning has negatively impacted students due to increased stress that can harm class performance [2] and

© Springer Nature 2022

induce hesitance when using tools for conveying social cues [3]. While online learning is not a new concept, e.g., Massive Open Online Courses (MOOCs) that have been in use for several years [4], prior work showed that student-to-student interaction in these methods of online learning is typically low [5]. These reduced interactions are especially problematic when it comes to project-based courses that require greater student interaction, such as in engineering design classes. This could in turn prevent the development of team psychological safety, or the "shared belief that the team is safe for interpersonal risk taking" ([6] p. 123). This paper uses a mixed-methods approach to illustrate a framework for measuring psychological safety and analyzing qualitative data to understand how psychological safety influences team interactions in online project-based courses like engineering design. We focus our analysis on 1) understanding when psychological safety becomes reliable and established as a measure, 2) how psychological safety varies between virtual and traditional teams, and 3) the factors that teams perceive to influence the building or waning of psychological safety.

Measuring and Validating Psychological Safety in Teams Research

Psychological safety plays a critical role in how teams coordinate and carry out tasks. Specifically, research has shown that this "safety" in teams is established through deep interactions and conversations that facilitate how team members feel treated and viewed by others [7], building from perceptions at the individual level, and then emerging as a collective team phenomenon [8, 9]. These feelings of psychological safety have also been shown to build and wane over time [10], and have been shown to translate to online contexts [11-13]. However, there are several challenges to supporting psychological safety in teams, including identifying when a shared understanding of team psychological safety is established and how it is affected over the course of a team's lifecycle. Measuring this establishment can show *to what extent* team members can converge to a strong Team Mental Model (TMM), i.e., a measure of how similar team members' organized knowledge representations are with each other [14]. Because psychological safety is dependent on whether team members are in agreement about how they feel about their team [6, 10], the effects of a stronger TMM demonstrate the potential for high psychological safety to drive team performance.

The first challenge of supporting psychological safety is understanding if there is a shared understanding of psychological safety in a team. This shared understanding is important because psychological safety is a *team-level* construct [6], referring to a *shared* belief that a team is safe for interpersonal risk taking. As such, the first step in computing psychological safety is justifying aggregations of individual team members' ratings of psychological safety to the team level and validating the team construct. This justification is established via two methods: (1) calculating interrater reliability (IRR, using Cronbach's alpha), which indicates when perceptions of psychological safety are established, and (2) computing interrater agreement indices (r_{wg} and ICCs), which justifies aggregating the scale to the team-level [15]. Cronbach's alphas that do not meet the acceptable threshold ($\alpha > 0.70$) may occur due to poor inter-relatedness between items or because heterogeneous constructs are present in the scale [16]. In addition, meta-analysis has shown that greater social presence, such as projecting one's self and the ability to perceive one another in an online environment [17], is related to greater persistence, retention, motivation, and success in an online course [18]. These factors could impact how participants complete course requirements, such as the psychological safety surveys. Therefore, understanding these aspects can help with understanding the challenges of extracting such team-level metrics, especially when some team members do not engage in the course. In contrast with interrater reliability, agreement indices (such as r_{wg}) and team member consistency measures (such as ICC, or intraclass correlation coefficient) allow us to investigate Team Mental Models (TMMs) [19]. Low ICC or r_{wg} values would bring to concern the validity of the TMM, indicating a lack of team consensus on team perceptions of psychological safety [6], and thus raise concern for aggregating team member responses to the team level [20].

The second challenge in supporting psychological safety is understanding when a shared perception of psychological safety is established in a team. While psychological safety has been shown to be a consistent, generalizable, and multilevel predictor of numerous outcomes important to individuals, teams, and organizations [21], these studies tend to implement "snap-shot" methods that capture the psychological safety of a team at the *end* of a project rather than over the course of a project. This is problematic because while these studies show that psychological safety is important factor in a team's performance, we do not know *when* to intervene, or *what type* of intervention would promote psychological safety. While recent research has shown that the TMM on psychological safety may stabilize early on in a team's life cycle in traditional teams [22, 23], psychological safety may be more difficult to establish in virtual teams as task interdependence increases [24]. This can arise from barriers in sensing team members' contexts and motives [25], lack of social and visual cues [26], and social loafing [27]. Furthermore, such obstacles can impact team performance negatively if individuals are "not on the same page" [28]. Thus, the first step is to understand the establishment of psychological safety in the environment that most students and schools can access.

Potential Differences with Virtual Teams in the Engineering Design Process

This section highlights why there may be differences in psychological safety between traditional (in-person) and virtual teams' engineering design processes and how we might be able to foster psychological safety during these stages. Particularly, meta-analytic evidence has shown that psychological safety influences tasks that are complex, knowledge-intensive, and involve creativity and sense-making [29]. These tasks make up the engineering design process [22, 23]. However, moving engineering design to an online environment may negatively impact team psychological safety because engineering teams rely on *knowledge-sharing* [22] to develop design solutions, which can suffer in an online environment [11]. Even more problematic, psychological safety is typically measured using snapshot methods where only one measure is obtained [30], emphasizing the importance of questioning if psychological safety manifests itself differently throughout the engineering design process [23]. However, there has been limited evidence exploring the role of psychological safety throughout the engineering design process, particularly in a virtual setting. While the design process is categorized into three phases including generation, evaluation, and communication [31, 32], the cornerstone of this process is team formation.

The beginning of most engineering design projects begins with team formation, where teams first meet and establish team culture. This early engagement is critical to the establishment psychological safety in a team, but research on traditional teams has shown that teams often vary in terms of formation, leadership, culture, norms, and accountability [33, 34] and that developing *trust* is a critical component of psychological safety [35]. This is further complicated in virtual teams where a lack of trust and free-riding team members is more prevalent, decreasing the likelihood of knowledge-sharing between individuals [11, 26, 36, 37]. In addition, trust can be harder to establish in virtual teams due to lower social presence and slower communication, which can disrupt performance outputs [28], as well as a lack of social cues in the online environment [12, 25] which can limit a teams' abilities to communicate naturally [13, 38]. Leadership is also related to trust such that team leaders can set the tone to create a psychologically safe environment [10, 29]. In online environments, leaders can build trust through "technological cues," such as performing kind gestures and maintaining constant team communication [39]. Stemming from a lack of trust, meta-analytic evidence has shown that the more virtual the team, the greater the opportunity for conflict over tasks [28]. However, conflict is not inherently bad [40], as psychological safety can allow teams to leverage conflict by encouraging team members to speak up and problem-solve through the issue [6, 10]. Otherwise, failing to control conflict can threaten team effectiveness and increase time for task completion [41]. To address trust issues, prior work used icebreakers and social games as interventions to build trust, which have been successful [42]. Additionally, structuring distributed synchronous peer-learning interactions improved performance and participation [43]. However, how we could even begin to apply such tools to foster team psychological safety at the start requires further preliminary work.

Branching after the start of a project, the concept generation stage of the design process relies on teams to develop creative ideas to be evaluated at subsequent stages [44]. Psychological safety is important during this process because low psychological safety can impair the ability to communicate ideas and knowledge [11, 45], as well as provide teams with the freedom to take risks by offering creative solutions [6]. Psychological safety also plays a vital role in concept screening when teams make go/no-go decisions when moving forward with concepts, as teams with high levels of psychological safety are more likely to be open to providing feedback can benefit teams and feeling safe for risk-taking [6], particularly when selecting creative ideas. Additionally, leader agreeableness can promote psychological safety [46], helping teams to engage in the aforementioned behaviors. During prototyping, students try to convey their designs [47]. In the final stage of the design process, teams compile their work during the final deliverables stage. This stage can be affected by poor communication, which can promote interpersonal tension [10], and lack of time management [23]. In the case of low psychological safety, such issues can fester if team members do not feel safe to question the status quo [6]. This implies that low psychological safety can promote insufficiencies in coordinating together, substantiating its importance until a project's end.

Across the engineering design stages, other factors can also play a role in the development of psychological safety. For example, The Nine Critical Considerations of Teamwork (9 C's) [48] (adapted to the Seven Critical Considerations of Engineering Design (7 C's) in [49]), contain factors highlighted in meta-analytic research [10, 29]. In virtual teams, such issues can be amplified due to reliance on technology, which could limit the ability to communicate and coordinate work [12]. Therefore, how such factors relate to psychological safety should be investigated as well.

Research Design and Methodology

Based on the previous literature, the goal of this work was to investigate the role of psychological safety in virtual engineering design teams. Specifically, this paper was developed to answer the following research questions:

- **RQ1:** How long does it take teams' psychological safety to become a reliable and established measure in virtual teams? Once established, how does psychological safety change over the design process?
- **RQ2:** To what extent is team psychological safety different between in-person (traditional) and virtual teams throughout the design process?
- **RQ3:** What factors impact team psychological safety in virtual engineering design teams throughout the course of a design project?

To answer the research questions presented above, an empirical study was conducted at a large northeastern university in the US during the first project of a cornerstone engineering design course over the Summer 2020 semester. The course provides students with the opportunity to go through an in-depth 4-week design project. The time points in this study represent the milestones in the engineering design process for a team [31](see Fig. 1).

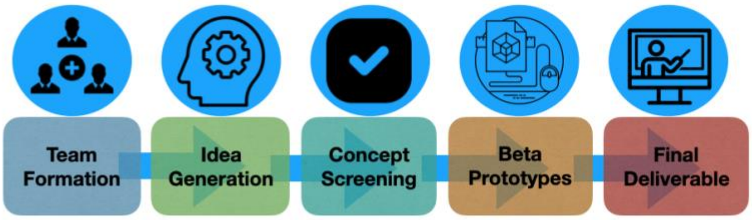

Fig. 1. Study timeline – psychological safety was captured at the end of each design stage, lasting approximately 3 hours each (total time period: 4 weeks)

Participants

Thirty-four engineering design student teams across six sections (i.e., classrooms) composed of 127 participants (93 males and 34 females) participated in the study. All participants were enrolled in a first-year engineering design course at a large northeastern university. The study was integrated into the curriculum through a series of surveys at the end of stages to represent the time points, where Miller et al.'s [23] work looked at traditional teams in engineering design. However, the current investigation examines six sections of a first-year engineering design course, and both studies occurred as a condensed summer session course. The traditional teams focused on designing for the developing world, and the virtual teams focused on designing for the developed world. Students were graded based on their participation.

Procedure

The online course followed a course schedule similar to the in-person study [23], where the psychological safety of the teams was analyzed over the same five time points (see Figure 1). Each design stage at their respective time point lasted approximately three hours, with approximately five days to a week in between each stage, and teams were grouped together for about half the time. In the online course, Zoom was used to conduct all lectures, which were broadcasted as live PowerPoint presentations. All students were able to use webcams and microphones to interact within the class, both in the main "classroom" with everyone present, and in breakout rooms. Importantly, at the end of each design stage, students completed an electronically distributed seven-question psychological safety survey developed by Edmondson [6], which is computed as an average of team members' scores for the team measure. These survey questions investigate important aspects of a team's relationship, such as team members feeling comfortable making mistakes without criticism, bringing up issues to overcome obstacles, and feeling valued as a team member [6]. At the end of each design milestone, students completed the psychological safety survey (five times). Participants were also required to provide positive and negative comments to support their rating in the psychological safety survey.

All participants consented at the beginning of the study based on the Institutional Review Board guidelines established at the university. Table 1 shows a summary of the number of participants, their instructors, and the design tasks in this study, along with the sample from [23].

Table 1 Descriptions of Design Challenges Based on Instructor and Semester

Project	Team Type	Semester	Instructor	Number of Teams	Number of Students
Tackle food insecurity in developing countries as a result of climate, conflict, unstable markets, food waste, and lack of investment in agriculture.	Traditional [23]	SU 2018	A	12	46
Based on the changes to the shopping experience imposed by COVID-19 restrictions, address the needs of grocery store shoppers for the world of today and tomorrow.	Virtual	SU 2020	B	12	46
			C	11	40
			D	6	23
			E	5	18

At the ***team formation*** stage, instructors grouped students into teams of 3 and 4 people randomly. Next, students were presented with a design challenge which varied depending on the section and instructor of the course. The teams then conducted in-depth context research on their design challenge, which guided their focus for their design project.

During the ***concept generation*** stage, students attended a lecture delivered via Zoom on customer needs and developed their problem statements within their teams in breakout rooms. After this, an innovation lesson that focused on the importance of creativity in engineering design was conducted. Next, the participants were guided through a series of idea generation exercises where they were asked to individually sketch as many ideas as possible in a 15-minute session in nominal brainstorming groups. After generating ideas, students shared them on Stormboard; a web-based application for sharing sketches of ideas.

During the ***concept screening*** stage, participants were led through a concept screening activity where they individually assessed all the ideas generated by their design team and rated them as “Consider” or “Do Not Consider.” This was conducted in the main Zoom “classroom” to limit teams from interacting. Ideas in the "Consider" category were concepts that the participant felt would most likely fulfill the goals for their design challenge, while ideas in the "Do Not Consider" category were concepts that the participants felt were not satisfactory for achieving their goals for the design challenge. This was continued until all ideas from the group were assessed. The students then discussed the ideas they screened and formed two piles as a group – “Consider” and “Do Not Consider.” They were tasked with picking out four distinct ideas to prototype in the next design stage.

During the ***beta prototypes*** stage, instructors held a lecture on prototyping methods, discussing the benefits of creating physical and virtual prototypes, as well as mockups of user experiences. From there, the students were placed in their breakout rooms to discuss the best prototyping options (at least 2 different methods) for their beta prototypes, depending on the information they were seeking from testing with users. From there, they started working on their prototypes as a team.

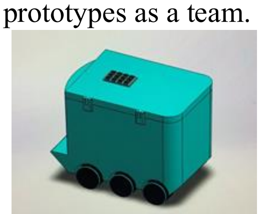

Fig. 2. Example of CAD rendering from one of the virtual teams – this represents a cart that travels to customers and stores refrigerated and unrefrigerated groceries

The project ended at the ***final deliverables*** stage, in which the final deliverables were completed including a formal PowerPoint presentation, a final design report, and a high-fidelity prototype including a CAD rendering of the design; an example is shown in Figure 2.

Results

Our results are presented in this section with relation to our research question. The statistical data were analyzed via the SPSS v.26. A value of $p < .05$ was used to define statistical significance. For all research questions, Cronbach's alpha and ICCs were used to establish reliability, whereas team psychological safety measures were used as a dependent variable.

Our first research question was developed to investigate how long it takes for psychological safety to become established as a reliable measure and whether the reliability changes over the course of the project. Our hypothesis was that psychological safety would not become a reliable measure until at least the idea generation stage due to a lack of social cues [12, 25] and social loafing while online [27]. Furthermore, lack of social presence can cause issues with motivation to complete tasks [18], therefore we hypothesized there may be decreased scale validity in the online environment. Finally, we hypothesized that ICC(1) and ICC(2) would tend to be lower in virtual teams, demonstrating lack of homogeneity in individuals' TMMs.

The results of our analysis reveal that the scale did not meet this internal reliability threshold during *team formation* and *idea generation*, shown in Figure 3. The results of this part of the research question support our hypothesis such that individuals' perceptions of psychological safety may take longer to become established in virtual teams measured based on the scale's reliability score, because team members may not have a sufficient understanding of their psychological safety. This could be attributed to participants having interpreted the scale items as heterogeneous and thus they rated each item differently [16]. Importantly, when we compare these findings to those captured during in-person instruction [23], we see that the construct of psychological safety is developed only after substantial time spent in the team setting, see Figure 3. This may be because the online environment makes it more difficult for team members to converge to a similar team mental model, due to lack of social cues [12, 25] that may limit teams' abilities to communicate naturally [13, 38]. This in turn can lead to a breakdown in coordination and trust [12], ultimately hurting team performance in multiple facets, such as biased interpretations of team members' behaviors.

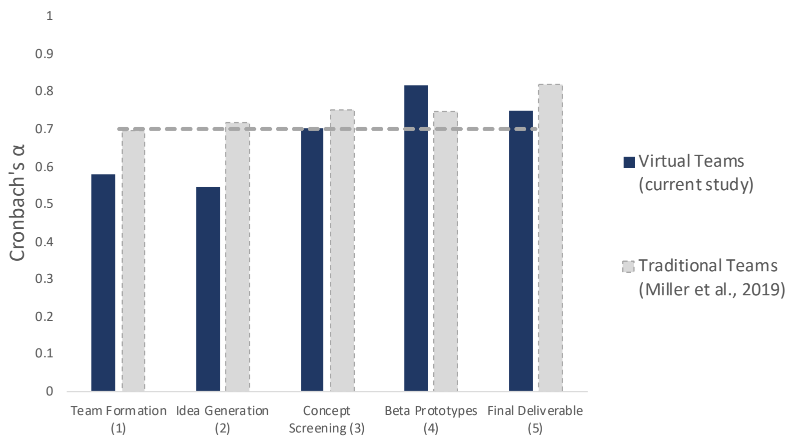

Fig. 3. The internal consistency of the psychological safety scale over the five design stages as measured by Cronbach's alpha (α) between traditional [23] and virtual engineering design student teams. The dashed line shows the acceptable level reliability (0.7).

To identify when the construct of psychological safety was established in the teams, the consistency of scores among team members was analyzed. In other words, team members must have a shared agreement regarding the overall level of psychological safety of the team [6] at each design stage because psychological safety describes the team rather than individual perceptions. As such, if there were disagreements between team members about the level of psychological safety at any design stage it would mean the team did not have a shared view of this construct and thus it would not be considered a shared team level construct [15]. The results revealed that the mean $r_{wg}(j)$ for remaining valid design stages ranged from .85 to .91, indicating acceptable agreement on psychological safety level within teams. Interestingly, these measures were not drastically different from the traditional teams, as mean $r_{wg}(j)$ values were between 0.86 and 0.92 [23]. Similar findings were reflected in ICC values, where ICC(1) was highest for final deliverables in the traditional teams (0.32) [23], and highest for beta prototypes in the virtual teams (0.34). It was also lowest at concept screening in both team types. Additionally, ICC(1)s revealed a range of variance in psychological safety due to team membership, ranging from 9% (small effect) to 34% (large effect) [15]. The larger value at beta prototypes reflects greater team interaction compared to concept screening and final deliverables. ICC(2) values followed a similar trend, indicating greater reliability of group mean ratings of psychological safety over time [15].

Finally, results from the one-way repeated measures ANOVA failed to show any statistically significant changes in psychological safety, over time, $F(1.591, 38.183) = 2.046$, $p = .151$. These findings support our hypothesis, as prior research in engineering design teams showed that the trajectory of psychological safety varied by team [23]. This implies that it is not the design activity that contributes to any rises or dips in psychological safety, but rather the interpersonal interactions that occur within each team. This can be generalized to various fields where teamwork is necessary, particularly in an online setting. Importantly, psychological safety may not change much across all teams within an organization until a significant amount of time passes, allowing the construct to manifest [10]. This effect of limited time spent as a team is evident in these results, see Table 2.

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

Design Stage	Mean	SD	α	Mean rwg(j)	Median rwg(j)	ICC(1)	ICC(2)
Concept Screening	6.04	.517	.70	.91	.92	.09	0.26
Beta Prototypes	6.22	.651	.82	.90	.95	.34	0.64
Final Deliverables	6.17	.565	.75	.85	.94	0.18	0.42

While the first RQ looked at how psychological safety did not become a reliable measure across teams until concept screening and there was not a significant difference across the design stage that were valid for analysis, the second RQ looked at how psychological safety compares between virtual and traditional teams throughout the design process. The data from the 12 teams in [23] were obtained from the authors and included in this analysis. Our hypothesis was that psychological safety would be lower in virtual teams, particularly during the earlier stages (e.g., team formation and concept generation) of the design project. The results show that at and beyond the concept screening stage in an engineering design project, differences in team psychological safety do not vary significantly between traditional and virtual teams (Table 3). These findings refute our hypothesis somewhat, as we would expect differences to be more prominent towards the beginning of the project, as prior research shows that it can be more difficult to foster psychological safety in online settings due to lack of social interactions that help form relationships and build trust [12, 36, 37]. These results imply that psychological safety does not vary between team types when comparing teams at each one of the design stages analyzed. However, results also point to something deeper occurring within the team.

Table 3 Independent t-test Results from the Virtual and Traditional Teams [23]

Design Stage	Mean (virtual teams)	Mean (traditional teams)	SD (virtual teams)	SD (traditional teams)	t-statistic	p-value
Concept Screening	6.04	5.96	.517	.563	t(43) = -.574	.569
Beta Prototypes	6.22	6.11	.651	.587	t(38) = -.509	.614
Final Deliverables	6.17	6.18	.565	.668	t(41) = .031	.975

While RQ1 and RQ2 established when psychological safety was established in design teams and how psychological safety varied between traditional and virtual teams, the goal of RQ3 was to identify the factors that impacted psychological safety within each design stage of the design process. To analyze the open-ended questions required at the end of the psychological safety survey, "Please describe any positive/negative team interactions or activities that impacted the rating," qualitative analysis was conducted. Using directed content analysis based on a codebook developed by Gong [49], we coded the 1,027 qualitative responses collected in the study at the end of each survey. Our hypothesis was that different factors impact the development of psychological safety due to different tasks and skills involved in each stage of the engineering design process [31]. Specifically, psychological safety can help teams feel safe to share ideas, feel open to receiving feedback, and feel safe for risk-taking [6, 10]. In all, seven main factors were coded based on [49]; see Figure 4 for the frequencies.

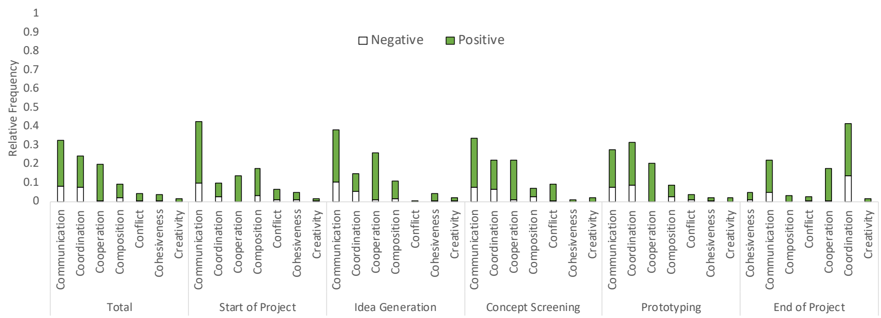

Fig. 4. Relative frequencies of discussion topics (total frequency of topic divided by total number of comments for each stage) spanning the project trajectory.

Throughout the design process, the most to least frequently mentioned topics were coded in this order: communication, coordination, cooperation, composition, conflict, cohesiveness, and creativity. Based on Cronbach's alpha (RQ1), we began the analysis from concept screening. A summary of the statistical findings is shown in Table 4.

Table 4 Linear Regression Results from the Qualitative Frequency and Psychological Safety Comparisons

Concept Screening	*B*	*SE B*	*β*	*t*	*p*
Negative Conflict	-5.204	1.302	-.583	-3.996	<.001
Final Deliverables	*B*	*SE B*	*β*	*t*	*p*
Total Negative Comments	-.558	.189	-.481	-2.958	<.001
Negative Communication	-1.403	.389	-.557	-3.609	<.001

The results of the regression analyses revealed certain factors to be significantly influential on psychological safety. Specifically, these occurred at both the concept screening and final deliverables stages. There were no significant findings for the beta prototypes stage. Specifically, results of the regression analyses from *concept screening* revealed one significant prediction variable. Specifically, the frequency of negative conflict comments on a team from *concept screening* significantly predicted team psychological safety, $F(1,31) = 15.66$, $p < 0.001$, accounting for 34% of the variation in team psychological safety with adjusted R^2=31.9%, a large effect size. The resulting prediction equation was: team psychological safety = 6.115 - 5.204*conflict (negative). Some participants cited lack of agreeableness as a negative aspect within their teams, for example: *If someone does not agree they make sure they make it known [P378].* This aligns with the notion that teams with lower psychological safety tend to be less agreeable, particularly when someone who is less agreeable assumes a leadership role [46]. This can discourage teams from considering a broader spectrum of ideas during the screening process. Additionally, while conflict can be beneficial throughout the design process [40], lack of agreeability in this case does not stimulate productive discussion for selecting ideas.

Contrasting with significant findings for negative conflict from *concept screening*, the *final deliverables* stage revealed different findings. The results of the regression analyses at *final deliverables* revealed two significant prediction variables. Specifically, the total negative comments a team documented statistically significantly predicted team psychological safety, $F(1,29) = 8.752$, $p = 0.006$, accounting for 23.2% of the variation in team psychological safety with adjusted R^2=20.5%, a medium effect size. The resulting prediction equation was: team psychological safety = 6.435 - 0.558* total comments (negative). Particularly, negative coordination comments (f=31) focused on topics such as being absent or late, lack of efficiency, working remotely, finishing on time, and work contributions. For example, participants cited examples about efficiency: *We wouldn't immediately address the confusion, instead of trying to figure out the problem for ourselves [P302]*. Such comments imply that the virtual teams did not always understand the instructions, nor did they have the self-control to stay on task. Particularly, lack of social and visual cues [12] from classroom instructors, and social loafing [27] can detrimental to virtual teams.

Additionally, several other participants complained about team members being either late or absent, for example: *[Student A] left class without telling us multiple times, then claimed he/she had doctor's appointments multiple times. When we ask for help [he/she] doesn't respond or does underwhelming work [P379]*. These kinds of comments imply a lack of trust in others, which can be detrimental to psychological safety [35], particularly online [11, 26]. Additionally, it implies that the team member may be free-riding, another issue in virtual teams [28, 37]. Issues related to communication were prevalent even more so. Specifically, the total communication negative comments statistically significantly predicted team psychological safety, $F(1,29) = 13.027$, $p < 0.001$, accounting for 31.0% of the variation in team psychological safety with adjusted R^2=28.6%, a large effect size [50]. The resulting prediction equation for final deliverables was team psychological safety = 6.336 - 1.403* communication (negative). Many participants lamented about the lack of communication from team members along having discussions. For example, a participant said: *Not everyone talks very often there could be more feedback on the topics that are brought up throughout the discussions [P376].* This comment hints at the need for addressing the common issue of lack of social cues in the online environment [12, 25], which is blamed for limiting teams' abilities to communicate naturally [13, 38]. Additionally, other team members voiced issues about contributing honest feedback: *Members did not criticize one another ideas too harshly, in fear of hurting others feelings [P351].* This is part of low psychological safety where teams may be afraid to challenge the status quo [10]; in this case, not feeling safe to challenge others' ideas during the design process.

While statistical analyses cannot be completed for team formation and idea generation, there are some qualitative insights that may aid in understanding why this is, as well as potential plans for technological intervention development. Specifically, at *team formation*, the category with the highest negative frequency was communication (f=22). Many complaints centered around team members not talking much, for example: *One person on my team didn't talk the entire time. They were talking in the background, texting, and overall seemed very uninterested in what we were doing [P379].* Other comments were very similar, and some participants thought that the lack of conversation was due to being in a new team: *I do not think that there were any negative interactions besides the fact that it can be a little quiet the first time you with work new people [P354].* This could be why psychological safety was unreliable, as lack of communication could make it difficult to gauge various facets of psychological safety, as well as make it difficult to build a TMM through getting to know one another.

Similar to team formation, the frequency of similar issues about communication (f=21) were presented by the participants during idea generation. Specifically, some participants were still thought that awkward silences were due to still being a new team, for example: *Occasional stops where there was silence, possible cause we were out of ideas and still new to the group [P271].* Such comments hint at the need for an intervention method that could encourage more participation in group conversations.

Conclusion and Discussion

In this paper, we investigated how the construct of psychological safety stabilizes in a virtual environment, the role of technology in the development of psychological safety, and potential interventions for supporting psychological safety development through a mixed-methods study of a first-year engineering design students over the trajectory of a course project. The key findings from our study are as follows:

- Individual perceptions of psychological safety take longer to establish and may be less reliable earlier on in virtual teams than traditional teams.
- There was no statistical difference in the psychological safety of virtual teams during concept screening, beta prototypes, and final deliverables compared with traditional teams.
- Negative comments about team conflict and communication were linked to lower psychological safety during concept screening and final deliverables, respectively.

At the first two design stage, *team formation* and *concept generation*, the data were unusable for statistical analyses at RQ2 and RQ3 because they did not meet the criteria set by Cronbach's alpha. Scale reliability can deteriorate due the presence of heterogeneous constructs in the scale [16], and because psychological safety covers various dimensions [6], that could be interpreted as heterogeneous constructs. However, psychological safety is valid across various settings [10], and has been successfully applied to traditional teams in engineering design [23]. This prompted further investigation of the comments from these design stages. The majority of comments from these design stages were categorized under *communication*, where most participants voiced concerns over the lack of discussion in their breakout rooms. This lack of communication early on signals that teams simply do not interact enough to understand how safe they feel, as lack of social cues can limit teams' abilities to communicate naturally [13, 38]. Thus, teams are prevented from building a Team Mental Model that promotes understanding the team's motives [14]. This calls for instructors to implement strategies to encourage communication earlier on in a project with virtual teams.

Because it is usually more difficult for communication to occur organically in a virtual team [12, 25, 33, 36], we recommend that instructors address this challenge before a project begins. To help students get to know one another, prior studies found that icebreakers and social games can be implemented to foster trust (which is conducive to psychological safety [35]) within teams [42] and encourage socializing to build psychological safety organically [39]. Similarly, prior work suggests utilizing informal exchanges and text chats to build a sense of community and avoid anxiety when using webcams [3], helping students to feel more comfortable.

While lack of participation can hinder the establishment of psychological safety, conflict can disrupt its building. Although conflict has its benefits for promoting higher team performance in engineering design [40], in this study, conflict was more likely to be associated with lower psychological safety during *concept screening*. This calls for encouraging students to tactfully confront issues and listen to others to increase psychological safety in the long run and prevent issues from festering [6]. Thus, teams can use conflict combined with higher psychological safety to their advantage [10].

Additionally, lack of communication hinders the development of psychological safety early in the design process. Although we suggest team-building activities to allow team members to get to know one another, such activities may not fully solve the communication problem. From research in MOOCs, structuring distributed synchronous peer-learning interactions and adding incentives for participation [43] could be applied to encourage more input from members within each team. However, this would require the development of a plugin that can be applied to various communication tools, such as video conference software. In addition to a lack of participation and interactions within teams, qualitative results point to issues with social loafing, which is a notable deficiency in virtual teams [27]. Thus, we suggest developing a tool to detect and discourage social loafing (e.g., when the room is quiet) within breakout rooms without being too intrusive.

For virtual teams, there was a lack of responses to the psychological safety survey at many design stages, so data points from teams with at least 50% participation were used in the analysis. However, r_{wg} values show that the team psychological safety values tend to become stabilized as the time progresses, aligning with the notion that psychological safety takes time to manifest and stabilizes over time [10]. Thus, we can assume most teams would be in agreement regarding their psychological safety in the later points of the project.

Additionally, findings from RQ2 show that because psychological safety does not vary between the team types (virtual versus traditional), psychological safety is dependent on more than whether a team is virtual or not. Thus, further investigation as to what causes differences within each team is necessary, as we currently lack an understanding of what specifically causes psychological safety to increase or decrease within a virtual team. For example, undetected interactions, such as students communicating via text-based messaging during activities in the main Zoom "classroom" could have impacted psychological safety in various ways in the virtual teams. However, we do not believe these interactions occurred frequently, if at all, as students were not motivated in any way to covertly communicate. Other confounding factors could come from stress induced by the pandemic; however, this was not studied here, nor were individual personality factors and gender. Furthermore, while design task, task duration, and task complexity may be impactful on design outputs, its impact on psychological safety was not investigated due to the scope of the paper. Similarly, prototyping may have been affected as well, but this was also outside the scope. Finally, while the guidance in this first project of the course may have impacted whether psychological safety changed, such changes are expected to be minimal, as team interactions were not manipulated directly.

Acknowledgements

This material is based upon work supported by the National Science Foundation under Grant No. 1825830. Any opinions, findings, and conclusions/recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

References

1. Grudin J (2020) After the iron horse: COVID-19 responses in education. interactions. 4:32–34.
2. Son C, Hegde S, Smith A, Wang X, Sasangohar F (2020) Effects of COVID-19 on college students' mental health in the United States: Interview survey study. J Med Internet Res. 9.
3. Yarmand M, Solyst J, Klemmer S, Weibel N (2021) "It feels like I am talking into a void": Understanding interaction gaps in synchronous online classrooms. Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. Association for Computing Machinery. p. 1-9.

4. Al-Rahmi W, Aldraiweesh A, Yahaya N, Bin Kamin Y, Zeki AM (2019) Massive Open Online Courses (MOOCs): Data on higher education. Data in Brief.118-125.
5. Kulkarni C, Cambre J, Kotturi Y, Bernstein MS, Klemmer SR (2015) "Talkabout: Making distance matter with small groups in massive classes," Proc. 18th ACM Conference on Computer Supported Cooperative Work and Social Computing, Association for Computing Machinery, 1116–1128.
6. Edmondson AC (1999) Psychological safety and learning behavior in work teams. Administrative science quarterly. 2:350-383.
7. Duhigg C (2016) What Google learned from its quest to build the perfect team The New York Times Magazine 26.
8. Kozlowski SWJ, Chao GT, Grand JA, Braun MT, Kuljanin G (2013) Advancing Multilevel Research Design:Capturing the Dynamics of Emergence. Organizational Research Methods. 4:581-615.
9. Mohammed S, Ferzandi L, Hamilton K (2010) Metaphor no more: A 15-Year review of the team mental model construct. Journal of Management. 4:876-910.
10. Edmondson AC, Lei Z (2014) Psychological Safety: The History, Renaissance, and Future of an Interpersonal Construct. Annual Review of Organizational Psychology and Organizational Behavior. 1:23-43.
11. Zhang Y, Fang Y, Wei K-K, 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. 5:425-436.
12. McLeod PL (2013) Distributed people and distributed information: Vigilant decision-making in virtual teams. Small Group Research. 6:627-657.
13. Olson J, Grinnell L, McAllister C, Appunn F, Walters K (2012) Towards a theoretical model of the impacts of incorporating webcams in virtual teams. Review of Business Information Systems (RBIS). 2:73-88.
14. Mohammed S, Klimoski R, Rentsch JR (2000) The measurement of team mental models: We have no shared schema. Organizational Research Methods. 2:123-165.
15. LeBreton JM, Senter JL (2008) Answers to 20 questions about interrater reliability and interrater agreement. Organizational Research Methods. 4:815-852.
16. Tavakol M, Dennick R (2011) Making sense of Cronbach's alpha. International journal of medical education.53-55.
17. Swan K, Shih LF (2005) On the nature and development of social presence in online course discussions. Journal of Asynchronous learning networks. 3:115-136.
18. Richardson JC, Maeda Y, Lv J, Caskurlu S (2017) Social presence in relation to students' satisfaction and learning in the online environment: A meta-analysis. Computers in Human Behavior.402-417.
19. DeChurch LA, Mesmer-Magnus JR (2010) Measuring shared team mental models: A meta-analysis. Group Dynamics: Theory, Research, and Practice. 1:1-14.

20. Klein KJ, Kozlowski SWJ (2000) From Micro to Meso: Critical Steps in Conceptualizing and Conducting Multilevel Research. Organizational Research Methods. 3:211-236.
21. Edmondson AC, Lei Z (2014) Psychological safety: The history, renaissance, and future of an interpersonal construct. Annu Rev Organ Psychol Organ Behav. 1:23-43.
22. Dylla N. Thinking Methods and Procedures in Mechanical Design. Mechanical Design. Munich, Germany: Technical University of Munich; (1991).
23. Miller S, Marhefka J, Heininger K, Jablokow K, Mohammed S, Ritter S (2019) "The trajectory of psychological safety in engineering teams: A longitudinal exploration in engineering design education," Proc. ASME 2019 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, 1-11.
24. Maynard MT, Gilson LL (2014) The role of shared mental model development in understanding virtual team effectiveness. Group & Organization Management. 1:3-32.
25. Bazarova NN, Walther JB (2009) Attributions in virtual groups: Distances and behavioral variations in computer-mediated discussions. Small Group Research. 2:138-162.
26. Hacker JV, Johnson M, Saunders C, Thayer AL (2019) Trust in virtual teams: A multidisciplinary review and integration. Australasian Journal of Information Systems. 0.
27. Piezon SL. Social Loafing and Free Riding in Online Learning Groups. College of Communication and Information Studies: Florida State University; (2011). p. 157.
28. Ortiz de Guinea A, Webster J, Staples DS (2012) A meta-analysis of the consequences of virtualness on team functioning. Information & Management. 6:301-308.
29. Frazier ML, Fainshmidt S, Klinger RL, Pezeshkan A, Vracheva V (2017) Psychological Safety: A Meta-Analytic Review and Extension. Personnel Psychology. 1:113-165.
30. Lenberg P, Feldt R (2018) "Psychological safety and norm clarity in software engineering teams," Proc. Proceedings of the 11th International Workshop on Cooperative and Human Aspects of Software Engineering, Association for Computing Machinery, 79-86.
31. Dym CL, Little P (2014) Engineering design: A project-based introduction. 4th ed. John Wiley and sons.
32. Pugh S (1991) Total Design: Integrated Methods for Successful Product Engineering. Addison-Wesley.
33. Fredrick TA (2008) Facilitating better teamwork: Analyzing the challenges and strategies of classroom-based collaboration. Business Communication Quarterly. 4:439-455.
34. Blumenfeld PC, Marx RW, Soloway E, Krajcik J (1996) Learning with peers: From small group cooperation to collaborative communities. Educational researcher. 8:37-39.

35. Conchie SM, Donald IJ, Taylor PJ (2006) Trust: Missing Piece(s) in the Safety Puzzle. Risk Analysis. 5:1097-1104.
36. de Pillis E, Furumo K (2007) Counting the cost of virtual teams. Communications of the ACM. 12:93–95.
37. de Pillis E, Furumo K. Virtual vs. face-to-face teams: Deadbeats, deserters, and other considerations. 2006 ACM SIGMIS CPR Conference on Computer Personnel Research. Claremont, California, USA: Association for Computing Machinery; (2006). p. 318–320.
38. Olson JD, Appunn FD, McAllister CA, Walters KK, Grinnell L (2014) Webcams and virtual teams: An impact model. Team Performance Management. 3/4:148-177.
39. Ford RC, Piccolo RF, Ford LR (2017) Strategies for building effective virtual teams: Trust is key. Business Horizons. 1:25-34.
40. Jonassen DH, Cho YH (2011) Fostering argumentation while solving engineering ethics problems. Journal of Engineering Education. 4:680-702.
41. Baltes BB, Dickson MW, Sherman MP, Bauer CC, LaGanke JS (2002) Computer-mediated communication and group decision making: A meta-analysis. Organizational Behavior and Human Decision Processes. 1:156-179.
42. Depping AE, Mandryk RL, Johanson C, Bowey JT, Thomson SC (2016) "Trust me: Social games are better than social icebreakers at building trust," Proc. 2016 Annual Symposium on Computer-Human Interaction in Play, 116-129.
43. Coetzee D, Lim S, Fox A, Hartmann B, Hearst MA (2015) "Structuring interactions for large-scale synchronous peer learning," Proc. 18th ACM Conference on Computer Supported Cooperative Work & Social Computing, 1139-1152.
44. Thompson G, 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. 1:17-31.
45. Kessel M, Kratzer J, Schultz C (2012) Psychological Safety, Knowledge Sharing, and Creative Performance in Healthcare Teams. Creativity and Innovation Management. 2:147-157.
46. Walumbwa FO, Schaubroeck J (2009) Leader personality traits and employee voice behavior: Mediating roles of ethical leadership and work group psychological safety. Journal of Applied Psychology. 5:1275-1286.
47. Kolodner JL, Wills LM (1996) Powers of observation in creative design. Design Studies. 4:385-416.
48. Salas E, Shuffler ML, Thayer AL, Bedwell WL, Lazzara EH (2015) Understanding and improving teamwork in organizations: A scientifically based practical guide. Human Resource Management. 4:599-622.
49. Gong J. The trajectory of psychological safety in engineering teams: A longitudinal exploration in engineering design education. Industrial Engineering: The Pennsylvania State University; (2020).
50. Cohen J (2013) Statistical power analysis for the behavioral sciences. 2nd ed. Routledge.