Proceedings of the Design Society, Volume 5: ICED25
https://doi.org/10.1017/pds.2025.10247

Map to safety: longitudinal examination of psychological safety in engineering capstone students

Aoran Peng[1,✉], Winnie Fang[1], Jessica Menold[1], Roxanne Moore[2] and Scarlett Miller[1]

[1] Pennsylvania State University, USA, [2] Georgia Institute of Technology, USA

✉ axp5524@psu.edu

ABSTRACT: Teams have been favored due to the diverse knowledge access. However, diversity can also have negative effects, and team outputs can be influenced by many factors, such as psychological safety. While the effects of psychological safety have been studied, its development has received less attention. Prior research in this area has focused either on specific populations or cross-sectional effects. To add to this area, this study examined the longitudinal development of psychological safety in engineering capstone students: how it evolves, and whether this can be influenced by team-related experiences. This study showed that although psychological safety did change meaningfully with time, neither time nor experience alone could capture the change. The results could shed light on the evolution of psychological safety, as well as what factors could potentially influence its development.

KEYWORDS: teamwork, design theory, communication, creativity, education

1. Introduction

A 2024 survey of almost 1,000 U.S. knowledge workers showed that 2/3 of them were unsatisfied with their teamwork experience (Hansen, Brianna, 2024). This is problematic due to the increased utilization of teams as problem-solving units in engineering across academia and the workforce, as teams bring a wider range of knowledge and expertise that aid design and innovation (Glaveanu et al., 2021; Gyory et al., 2019; Kratzer et al., 2010; Singh & Fleming, 2010). However, the statistics provided by Hansen (2024), and perhaps the experiences of many individuals, show that teamwork is often less than ideal. One reason for this perception could be 'social loafing'–an unwillingness to participate in and contribute to group work (De Vreeze & Matschke, 2019; Paulus, 2000; Paulus & Brown, 2003).

Previous works have tried to identify specific factors that might influence team experience and output, one of which is the psychological safety perception of the team members. Psychological safety is *"the shared belief that the team is safe for interpersonal risk-taking"* (A. Edmondson, 1999). Previous studies have found that when psychological safety is high, or when people feel safe, they are more likely to contribute and voice their opinions in the team (A. C. Edmondson & Lei, 2014). In a separate study, psychological safety was also found to be positively linked to information sharing and engagement, creativity, and performance (Frazier et al., 2017). However, compared to the impact of psychological safety, there has been relatively little research on the creation and/or erosion of psychological safety (A. C. Edmondson & Lei, 2014), and most studies in this area focused on cross-sectional research and not longitudinal research (Frazier et al., 2017).

In more recent years, there have been attempts at addressing this gap. For example, the work of Cole et al. (2022) explored ways in which psychological safety can be measured longitudinally in engineering student teams, and what factors can impact the growth and reduction of psychological safety over time. In addition, in a separate paper by Cole et al., (2022), the authors found that in virtual teams, psychological safety takes longer to stabilize. While these previous works have shed some valuable light on how psychological safety changes over time, both papers focused on freshman engineering students.

This study addresses a gap regarding longitudinal change in psychological safety in more advanced grade levels. More specifically, this study sought to examine the trend of psychological safety of senior capstone engineering student teams over the course of five-time points in a semester. In addition, this study also examined how student experiences contribute to the manifestation of psychological safety at different times. The results of this study provide key insight into the development of psychological safety over the course of one engineering design project. This can add to previous research and identify areas of focus for future development of interventions to foster psychological safety in senior engineering students.

2. Related works

Research into how to best foster psychological safety is one area of teamwork research aimed at improving team outputs. However, previous studies have found that psychological safety can change depending on the time and situation. Therefore, this study aims to investigate how psychological safety develops in senior capstone students. This section serves to review relevant literature and lay a foundation to support the current paper.

2.1. Team output can be influenced by team dynamics and interaction

Teams are a complex and dynamic system as its members adapt to the various situations and demands (Kozlowski & Ilgen, 2006). The benefits of the team in a project setting have been proven with numerous previous research (Dennis, 1996; Gyory et al., 2019; Karau & Williams, 1993). For example, good team dynamic can improve team project outcome (DiTullio, 2010), while negative team dynamics and experience can damage performance and outcomes (Buckenmyer, 2000; Jones, 1996). When asked, students often indicated that things like bad interaction and organization, bad communication, and bad collaboration contributed to their bad experiences (Buckenmyer, 2000). To investigate more closely at teamwork, some factors from Salas et al.'s Nine Critical Considerations of Teamwork (9 C's) (2015) as well as the factors of cohesiveness and creativity were used to set the structure of the examination.
The first factor of the model is team composition, which can have a complicated effect on the team performance. For example, more diverse teams have been found to perform better because members are able to bring different contributions (De Cooman et al., 2015; Muchinsky & Monahan, 1987; Neuman et al., 1999; Piasentin & Chapman, 2007). On the other hand, more homogeneous teams have also been found to perform better as similarities between members foster a sense of belonging (Muchinsky & Monahan, 1987; Neuman et al., 1999). Another factor is communication, where effective communication is positively linked to member satisfaction and motivation (Brewer et al., 2015; Mikkelson et al., 2015). This could be due to communication builds trust and connection which can improve teamwork (Zemliansky, 2012). Conflict is another factor. And contrary to popular belief that conflict impedes team cohesion and negatively impacts emotion (Gladstein, 1984; Hackman & Morris, 1975; Saavedra et al., 1993; J. A. Wall Jr & Callister, 1995; V. D. Wall Jr & Nolan, 1986), depending on the circumstances, conflict can have positive effects as well. For example, it was found that some conflicts can stimulate the exchange of information during the decision-making process (Levine et al., 1993), become more creative and boost their problem-solving skills (Levine et al., 1993; Nemeth, 1986; Tjosvold, 1997). Coordination as a factor is tied to communication, where good coordination involves the members of the team sharing a common mental model (Cannon-Bowers et al., 1990; Orasanu, 2022). This shared mental model enables members to make similar decisions and actions without explicit communication (E. E. Entin & Serfaty, 1999). This coordination was then found to be positively related to team performance. (E. Entin et al., 1993), and can help to decrease the cognitive load in stressful situations, as members understand each other and what needs to be done (Orasanu, 2022). The factor of cooperation is when team members emphasize accomplishment as a group and collaboration (Beersma et al., 2003). These traits positively influence team performance and output, as collaboration can build trust and cohesiveness between team members (Ivancevich et al., 1990). Group cohesion, which is defined as the willingness of the group to prefer being with each other to complete the overall goal (Carron et al., 1998), has been found to significantly influence team performance (Carron et al., 2002; Heuzé et al., 2006). This relationship was further supported in sports-related team research, where it was found that teams with higher levels of cohesion had a higher success rate (Martens & Peterson, 1971). Finally, good teamwork work has often been found to promote creativity in the designs they produce. For example, a study found that increased diversity in teams has also been found to increase their innovative rating assessment scores (Ancona &

Caldwell, 1992). And good team environment can also help to decrease anxiety associated with producing and selecting creative ideas and increasing psychological safety (A. C. Edmondson, 2004; Shin & Eom, 2014).

2.2. Evolution of psychological safety in teams

From previous literature, it could be seen that many factors impacting team experience and output are linked to the individual's perception of psychological safety in teams. For example, higher psychological safety can help individuals to more confidently express their opinions and increase their creativity (A. C. Edmondson, 2004). In addition, higher psychological safety can also help individuals to provide feedback to their peers and engage in meaningful discussions (A. C. Edmondson, 2003; West, 1990). And with the presence of conflict in a team related to the task or project problem, the presence and level of psychological safety within the team can influence whether the conflict will have a positive or negative impact on team performance (Bradley et al., 2012).

Because of the importance of psychological safety, it is also important to understand how it can be built. The first area to investigate is the factors impacting team performance, which are related to psychological safety. More specifically, the factors of interpersonal relationships, group interaction and dynamics, the leadership within the group, and the organization's culture and norms are all related to the development of psychological safety (Kahn, 1990).

Time as a factor might also impact psychological safety, as it takes time to develop (Kozlowski et al., 2013). More specifically, in the paper by Cole et al. (2022) examining psychological safety in freshman engineering students, they found that although psychological safety can be reliably measured across a continuous time, it will tend to be more unreliable toward the beginning. In the same paper, the authors found that psychological safety perceptions of the team decreased, which might be an indicator that the variation of psychological safety perception of the individual team members increased (Cole, O'Connell, Gong, et al., 2022). The stage of the design process has also been found to impact psychological safety development in freshman engineering students, with concept generation and selection having the lowest psychological safety scores (Miller et al., 2019). Of the seven C factors (Composition, communication, coordination, cooperation, cohesiveness, creativity, and conflict) they investigated, communication and cohesiveness were found to impact the team's performance and output significantly (Cole, O'Connell, Gong, et al., 2022). However, although these studies have provided valuable findings, the population they focused on was mainly freshman engineering students.

3. Methodology

The study expanded over one academic semester in 2023, and collected data with a total of 26 participants, with 18 participant data used in the final analysis.

3.1. Participants

Participants for this study were all undergraduate students who enrolled in a fourth-year senior capstone engineering design course at a large northeastern university in the Spring semester of 2023 and were 18 years or older. Participants were recruited in their classrooms after consent was initially gathered from their instructors via email. A total of 26 participants were recruited, but those with missing data points were removed from the analysis, resulting in the final 18 that were analyzed for this paper. Due to the focus of the study being the psychological safety scores, no demographic data were collected from the participants.

3.2. Procedure

The data for this study was completed in a fourth-year capstone engineering design course, and the data collection process followed that of the one used in a previous study by Cole et al., (2022), with adaptations for use in a capstone course. Approval was obtained from the Institutional Review Board (IRB) at the Pennsylvania State University. Researchers obtained consent from the instructors to use some class time to collect the data. Data collection occurred at five time points over one spring semester in 2023: 1) start of the project, 2) idea generation, 3) concept selection, 4) prototyping, and 5) final design deliverables. The process of data collection can be better visualized in Figure 1. More specifically, the time between each time point could range from one to five class sessions (equating to 1-3 weeks, given the 2/week class meetings), depending on the curriculum progress and the instructor's needs. Each design session occurred within one class period. Details on what student participants were completing in the course curriculum can be better found in Cole et al. (2022), they will only briefly be explained here. For time point 1, participants were assigned to groups of three to five depending on their project. The

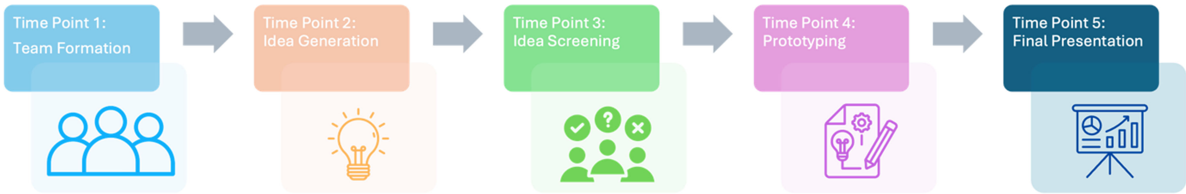

Figure 1. Data collection process

participants then researched the project they were completing. For time point 2, students were guided through the process of idea generation and were asked to generate solutions to the problem they are working on. At time point 3, teams were guided through the process of concept screening and were asked to evaluate the ideas their team members generated. At time point 4, the students were asked to build prototypes of the ideas that passed concept screening. Finally, at time point 5, teams presented their final ideas to the class using PowerPoint presentation and a fully constructed final prototype. For all time points, the survey was completed at the end of the class.

More specifically, at each time point, after the students had completed the assigned activity of interest, a survey was given to them to complete through Qualtrics. At the beginning of the survey, the students were prompted to create a unique code, which they used for all the five time points of data collection. The survey included seven questions developed by Edmondson (1999) to measure the psychological safety of the individual, and two open-ended questions. More specifically, the first seven questions collected 7-point Likert scale responses ranging from 1 (very inaccurate) to 7 (very accurate), and are aimed at examining how comfortable the individuals felt in the team in times when, for example, they made a mistake or if they felt accepted by the team (A. Edmondson, 1999). For the score of psychological safety perception of everyone, it was calculated by first reverse coding any of the negative items in the survey, and then taking the average of all their responses. The two open-ended questions that followed the survey questions asked the participants to write two instances during that measured period where they experienced a positive and a negative interaction in the team.

3.3. Content analysis

To analyse the two open-ended responses captured as part of this survey, a qualitative analysis was conducted using NVIVO. More specifically, the questions asked the students to "Please describe any positive/negative team interactions or activities that impacted the rating." For this part of the analysis, the code book used in the previous work by Cole et al., (2022) was used. This codebook was named the "Seven Critical Considerations (7C's) of Psychological Safety in Engineering Design Teamwork." This structure was established by adding two additional factors (Creativity and Cohesiveness) (Mullen & Copper, 1994) to the five of the Critical Considerations for Teamwork Framework (Composition, Communication, Coordination, Cooperation, and Conflict) (Salas et al., 2015). For this part of the analysis, a total of 104 responses were collected between the students for four-time points, starting at Time Point 2. Of the 104 responses, those who did not give consent were removed, as well as those whose responses were blank, "none", or "n/a." Interrater reliability was established between two raters. One rater was a Post-Doctoral Scholar in Engineering Design and Innovation, and the other rater was a Ph.D. student in Industrial Engineering. The two raters independently coded 20% of the dataset here in NVIVO. They coded at the sentence level, and each sentence was allowed to be coded into multiple categories. Their inter-rater reliability (Cohen's Kappa) was checked at the end. When the reliability was below 0.75, they discussed the coding structure until they reached an agreement. They then went back to the text and re-coded the text. Finally, an inter-rater reliability (Cohen's Kappa) of 0.89% was achieved. Then, one coder coded the rest of this dataset.

For the coding structure, the first code, Cohesiveness, included the instances when participants mentioned anything regarding the team dynamic and team chemistry. For example, one participant mentioned "By this point in the semester, I have grown stronger relationships with certain team members." This was coded to the Cohesiveness factor. For Communication, it was whenever the participants mentioned communication explicitly, or if they mentioned verbal interactions, such as "When we had a problem, all of us could talk about it and tried to figure it out." The Composition factor encompassed descriptions of the team members, for example, "The team seems very knowledgeable in their respective fields." The conflict factor was coded for all the times the participants mentioned things like agreement or disagreement between the teammates, such as "It's difficult to ensure that everyone is on the same page because of their different viewpoints and knowledge on the project thus far." Next, anything that mentioned teamwork or working together as a team was coded to Cooperation, such as "One positive team interaction I had was working on the final report collaboratively." For the factor coordination, it was whenever the participants mentioned anything related to assigning work, managing work or the team, etc. For example, "The team is usually on the same page with what the top priorities are." was coded as Coordination for positive experience. Finally, creativity was when the participants mentioned ideation or selection of ideas, such as "Recently being able to show a "final" device to our sponsor and volunteers. Showed some light at the end of a long and tiring tunnel."

3.4. Change in psychological safety

To investigate the building and waning of psychological safety in engineering design teams studied here, change in psychological safety was used. More specifically, change in psychological safety was calculated based on the psychological safety score of the individual at each time point. Their psychological safety score was calculated as average of their answers for the seven, 7-point Likert Scale type questions of psychological safety. The change in psychological safety was calculated by using this equation (Equation 1):

$$\triangle Psychological\ Safety_{ij} = (Psychological\ Safety)_i - (Psychological\ Safety)_j \quad (1)$$

where i is the current time point, ranging from 2-5 and j is the previous time point, ranging from 1-4. This process produces four changes in psychological safety scores for each participant and is either a negative value, zero, or positive value. A negative value indicates that the current time point score was lower than the previous time point score, and therefore shows that psychological safety decreased; 0 means that there was no change; and a positive value indicates that the psychological safety increased between the two time points.

4. Data analysis and results

In this section, details will be provided regarding the results from the analysis of the longitudinal study. For this current dataset, the average psychological safety score for each time point was 5.603 ± 0.206, 5.333 ± 0.248, 5.258 ± 0.300, 5.381 ± 0.268, 5.278 ± 0.282, for time points 1-5, respectively. This is presented as the average ± standard error. Statistical analysis of the data was carried out using SPSS v.29.0.0.0(241). Prior to analysis, internal consistency between the participant responses in the psychological safety survey questions was checked, and the results showed a high internal consistency, as determined by a Cronbach's alpha of 0.830.

4.1. RQ1: How did the psychological safety perception of the individual change with time?

This research question was established to explore if over the course of the semester the psychological safety perception of the students changed significantly. We hypothesize that time would have a significant impact on the development of psychological safety, as supported by prior research by (Cole, O'Connell, Gong, et al., 2022). However, as this is more explorative, there is no hypothesis regarding the specific relationship between time and psychological safety development for this population.

To answer this research question, a one-way repeated measures ANOVA was conducted, with the continuous dependent variable being the psychological safety score at each time point. Before conducting the actual analysis, assumptions were checked and were found to be satisfied. One outlier was found per inspection of the box plot. Analysis was done with and without the outlier, and no significant differences in results were observed, therefore, the outlier was kept in the final analysis. There were no statistically significant changes in the psychological safety perception of the students over time, $F(4, 68) = 1.209$, $p = 0.315$, partial = 0.066. Psychological safety perception of the students decreased from 5.603 ± 0.206 to 5.333 ± 0.248 at time point 2, then decreasing further to 5.258 ±

0.300 at time point 3, then increasing slightly to 5.381 ± 0.268 at time point 4, and finally decreasing again at time point 5 to 5.278 ± 0.282.
To get a more clear understanding of the scores observed, growth curve modeling was conducted with SPSS. More specifically, a mixed-effect model was conducted to examine the effects of the time points on the scores. The intercept was significant, b = 5.552, SE = 0.228, t(21.873) = 24.309, p < 0.001. This indicates a strong baseline score. However, the effect of the time point was not significant, b = -0.060, SE = 0.050, t(21.873) = -1.205, p = 0.241. The results also revealed significant variability in repeated measures, Variance = 0.225, SE = 0.042, Wald Z = 5.39, p < 0.001, indicating that scores changed meaningfully across time points. Additionally, there was significant variability in intercepts among participants, Variance = 0.691, SE = 0.282, Wald Z = 2.453, p = 0.014. However, the variance associated with time points was not significant, Variance = 0.023, SE = 0.014, Wald Z = 1.641, p = 0.101.

4.2. RQ2: What is the relationship between the positive experiences of the students and their changes in psychological safety?

This research question was established to explore how the positive experiences of the students, categorized into the 7C's metric, can influence the ebbs and flows of psychological safety in individuals in the population studied here. Since this is a more exploratory investigation, no formal hypothesis was formed. However, it was found in a previous study that communication was the most often mentioned code for freshman engineering students (Cole, O'Connell, Gong, et al., 2022). Therefore, it might have some impact on psychological safety changes. For this part of the analysis, because each response was allowed to be coded into multiple categories, the content analysis resulted in 150 entries for positive experiences.
To examine this impact, quantitative responses from the open-ended responses section of the survey were analyzed. we conducted a Kruskal-Wallis H Test, with the dependent variable being the scores of the change in psychological safety, and the independent variable being the categories of positive interaction. The nonparametric test was selected because, for each of the 7 categories of the independent variable, the number of entries is not equal. This makes the parametric test for the relationship of interest, one-way ANOVA, more subject to errors. Therefore, the Kruskal-Wallis H Test was chosen.
More specifically, the Kruskal-Wallis H test was conducted to determine if there were differences in change in psychological safety scores between the 7 groups of experiences for positive experience: Cohesiveness (n = 14), Communication (n = 32), Composition (n = 9), Conflict (n = 2), Cooperation (n = 41), Coordination (n = 46), Creativity (n = 6). The distribution of change in psychological safety scores was the same for all seven categories, as assessed by visual inspection of a boxplot. The change in psychological safety scores increased from Cohesiveness (Mdn = -0.143) to Communication (Mdn = 0.143), to Composition (Mdn = 0.429), and then it decreased slightly to Conflict (Mdn = 0.071), after which it decreased again to Cooperation (Mdn = -0.143), then it remained the same at Coordination (Mdn = -0.143), and increasing a bit to Creativity (Mdn = 0). However, these differences were not statistically significant, (6) = 7.224, p = 0.301.

4.3. RQ3: What is the relationship between the negative experiences of the students and their changes in psychological safety?

This research question was established to explore how the negative experiences of the students, categorized into the 7C's metric, can influence the building and waning of psychological safety in the population studied here. Since this is a more exploratory investigation, no formal hypothesis was formed. However, it was found in a previous study that communication was the most often mentioned code for freshman engineering students (Cole, O'Connell, Gong, et al., 2022). Therefore, it might have some impact on psychological safety changes. For this part of the analysis, because each response was allowed to be coded into multiple categories, the content analysis resulted in 135 entries for positive experiences.
To examine this impact, quantitative responses from the open-ended responses section of the survey were analyzed. we conducted a Kruskal-Wallis H Test, with the dependent variable being the scores of the change in psychological safety, and the independent variable being the categories of negative interaction.
More specifically, the Kruskal-Wallis H Test was conducted to determine if there were differences in change in psychological safety scores between the 7 groups of experiences for positive experience: Cohesiveness (n = 9), Communication (n = 30), Composition (n = 19), Conflict (n = 7), Cooperation (n = 23), Coordination (n = 39), Creativity (n = 8). The distribution of change in psychological safety was

the same for all seven categories, as assessed by visual inspection of a boxplot. The change in psychological safety scores increased from Cohesiveness (Mdn = -0.429) to Communication (Mdn = 0.107), then decreased again to Composition (Mdn = -0.286), and decreased further to Conflict (Mdn = -0.429), where it remained the same to Cooperation (Mdn = -0.429), and then increased at Coordination (Mdn = 0.071), and finally increasing more at Creativity (Mdn = 0.143). However, these differences were not statistically significant, $\chi^2(6) = 7.173$, $p = 0.305$.

5. Discussion, limitations, and future work

This study was designed to investigate the development of psychological safety in senior engineering students in their capstone design teams. The analysis of the data gathered found that for the population studied here:

- First, psychological safety did change meaningfully through time, but there was no trend just by looking at each time point.
- Both positive and negative experiences, mapped using the 7C structure, were not able to capture the change in psychological safety.

The first significant result of this study suggests that while participants start at significantly different baseline scores, the variable time points do not significantly predict changes in scores over time. More specifically, this means that while there is some significance in the progress of psychological safety over time, the underlying trend could not be captured using the five time points studied here. This could suggest that something else might be influencing the development of psychological safety through time, more than time itself. The significant variability in repeated measures highlights the importance of considering time in understanding score dynamics.

The second significant result of this study suggests that the development of psychological safety through time also could not be predicted by the positive and negative experiences of the participants, mapped using the 7Cs structure. However, upon closer examination of the data, it could be seen that there are variations in the median change in psychological safety scores for each of the seven categories for both positive and negative interactions. This could be that the change is too small to be statistically significant. It would also be that the current sample size or the current structure cannot capture the change in psychological safety. For example, the number of entries was different for each of the 7 categories, for both positive and negative interactions. In addition, for positive experience, Coordination, Cooperation, and Communication are the three most mentioned categories, in this respective order with coordination having the highest frequency. On the other hand, for negative interactions, the factors with the highest frequencies are Coordination, Communication, and Cooperation, in this order. These are the same factors as the positive experience, with only the ordering being slightly different. While the differences in the change of psychological safety scores are not significant between them and the other factors, the high frequency could still be a good indicator that instructors need to pay attention to them during teamwork. In addition, it is important to note that all seven factors, for both positive and negative experiences, were considered independently. It could also be that some of these factors work together to truly represent the experience they have since each experience can be coded into multiple categories. Therefore, future research can also investigate how to consider these codes simultaneously to capture the interaction effect. Another factor to consider is that for this analysis, we only investigated the change in psychological safety between two consecutive time points. However, nothing was done in the data to indicate which segment of time each change happened. For example, the change between time point 1 and 2 is indifferent from time point 2 and 3. This could also have influenced the results, and is worth further investigation.

Lastly, although this study was able to produce some notable findings, there are still some limitations that should be noted. The first limitation is the population limitation, the population used here was limited in number and region. Therefore the results found could be applicable to only this population and future studies should expand not only the number but also to other universities as well to see if results persist. In addition, other factors can influence team composition and experience such as various demographics factors like gender, as well as the instructors of the courses who interact with the participants closely. Previous experiences with teamwork, as well as additional training in related aspects such as leadership can also impact participant perceptions and behaviour. These all points to potential areas for future investigations. Lastly, another interesting area for future investigation is the interaction between positive and negative experiences. These two experiences as a whole might provided a more holistic picture of their team experience, and examining them together might provide additional interesting results. This would be another area worth investigating.

6. Conclusion

This study was conducted to investigate the development of psychological safety in engineering capstone students through time and to see if that development can be influenced by the type of experiences students have. To do so, we examined data from 18 senior engineering capstone students. The results of this study showed that there was a meaningful change in psychological safety scores over time, but time itself is not enough to capture this change. In addition, positive and negative experiences, categorized using the 7Cs framework, also cannot capture the change in psychological safety. Based on these findings, it could be inferred that something deeper is at play for the development of psychological safety in capstone engineering students, beyond just time and experience. Examining just the categories themselves, it could be seen that communication, coordination, and cooperation are the three factors noted by most students for both positive and negative experiences, and therefore suggest that these factors should be paid more attention to by researchers and educators. Finally, these findings can all help to serve as empirical evidence to support future investigations into psychological safety and its development, as well as inform researchers and educators on areas of interest in this area.

Acknowledgment

Special thanks to the Learning Factory for their support in the data collection process of this paper. This paper is based on the work supported by the National Science Foundation (Grant No. 2337014).

References

Ancona D. G., & Caldwell D. F. (1992). Demography and design: Predictors of new product team performance. *Organization Science*, 3(3), 321-341.

Beersma B., Hollenbeck J. R., Humphrey S. E., Moon H., Conlon D. E., & Ilgen D. R. (2003). Cooperation, competition, and team performance: Toward a contingency approach. *Academy of Management Journal*, 46(5), 572-590.

Bradley B. H., Postlethwaite B. E., Klotz A. C., Hamdani M. R., & Brown K. G. (2012). Reaping the benefits of task conflict in teams: The critical role of team psychological safety climate. *Journal of Applied Psychology*, 97(1), 151.

Brewer P. E., Mitchell A., Sanders R., Wallace P., & Wood D. D. (2015). Teaching and learning in cross-disciplinary virtual teams. *IEEE Transactions on Professional Communication*, 58(2), 208-229.

Buckenmyer J. A. (2000). Using teams for class activities: Making course/classroom teams work. *Journal of Education for Business*, 76(2), 98-107.

Cannon-Bowers J. A., Salas E., & Converse S. (1990). Cognitive psychology and team training: Shared mental models in complex systems. *Human Factors Bulletin*, 33(1990), 1-4.

Carron A. V., Brawley L. R., & Widmeyer W. N. (1998). *The measurement of cohesiveness in sportgroups*.

Carron A. V., Colman M. M., Wheeler J., & Stevens D. (2002). Cohesion and performance in sport: A meta analysis. *Journal of Sport and Exercise Psychology*, 24(2), 168-188.

Cole C., O'Connell A., Gong Z., Jablokow K., Mohammad S., Ritter S., Heininger K., Marhefka J., & Miller S. R. (2022). What factors impact psychological safety in engineering student teams? A mixed-method longitudinal investigation. *Journal of Mechanical Design*, 144(12), 122302.

Cole C., O'Connell A., Marhefka J., Jablokow K., Mohammad S., Ritter S., & Miller S. (2022). *How Long Until We Are (Psychologically) Safe? Longitudinal A. Investigation of Psychological Safety in Virtual Engineering Design Teams in Education*. 767-784.

De Cooman R., Vantilborgh T., Bal M., & Lub X. (2015). Creating Inclusive Teams Through Perceptions of Supplementary and Complementary Person-Team Fit. *Group & Organization Management*, 41(3), 310-342. https://doi.org/10.1177/1059601115586910

De Vreeze J., & Matschke C. (2019). Don't Put Me in This Group: Assignment to Non-Preferred Groups Increases Disidentification and a Preference for Negative Ingroup Information. *Social Psychology*, 50(2), 80-93. https://doi.org/10.1027/1864-9335/a000363

Dennis A. R. (1996). Information exchange and use in group decision making: You can lead a group to information, but you can't make it think. *MIS Quarterly: Management Information Systems*, 20(4), 433-454.

DiTullio L. (2010). *Project team dynamics: Enhancing performance, improving results*. Berrett-Koehler Publishers.

Edmondson A. (1999). Psychological safety and learning behavior in work teams. *Administrative Science Quarterly*, 44(2), 350-383. https://doi.org/10.2307/2666999
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), 1419-1452.
Edmondson A. C. (2004). Psychological safety, trust, and learning in organizations: A group-level lens. In KramerR. M. & CookK. S. (Eds.), *Trust and distrust in organizations: Dilemmas and approaches* (pp. 239-272). Russell Sage Foundation.
Edmondson A. C., & Lei Z. (2014). Psychological safety: The history, renaissance, and future of an interpersonal construct. *Annu. Rev. Organ. Psychol. Organ. Behav.*, 1(1), 23-43.
Entin E. E., & Serfaty D. (1999). Adaptive team coordination. *Human Factors*, 41(2), 312-325.
Entin E., Serfaty D., Entin J., & Deckert J. (1993). *CHIPS: Coordination in hierarchical information processing structures*. TR-598). Burlington, MA: ALPHATECH.
Frazier M. L., Fainshmidt S., Klinger R. L., Pezeshkan A., & Vracheva V. (2017). Psychological Safety: Meta A.-Analytic Review and Extension. *Personnel Psychology*, 70(1), 113-165. https://doi.org/10.1111/peps.12183
Gladstein D. L. (1984). Groups in context: A model of task group effectiveness. *Administrative Science Quarterly*, 499-517.
Glaveanu V. P., Ness I. J., & Rasmussen L. J. T. (2021). Chapter 2—Creative success in collaboration: A sociocultural perspective. In McKayA. S. Reiter-Palmon R., & Kaufman J. C. (Eds.), *Creative Success in Teams* (pp. 19-32). Academic Press. https://doi.org/10.1016/B978-0-12-819993-0.00002-3
Gyory J. T., Cagan J., & Kotovsky K. (2019). Are you better off alone? Mitigating the underperformance of engineering teams during conceptual design through adaptive process management. *Research in Engineering Design*, 30(1), 85-102. https://doi.org/10.1007/s00163-018-00303-3
Hackman J. R., & Morris C. G. (1975). Group tasks, group interaction process, and group performance effectiveness: A review and proposed integration. *Advances in Experimental Social Psychology*, 8, 45-99.
Hansen Brianna. (2024, February 8). Teamwork 2024 Report: This isn't working. *Mural*. https://www.mural.co/blog/insights-into-teamwork-effectiveness
Heuzé J.-P., Sarrazin P., Masiero M., Raimbault N., & Thomas J.-P. (2006). The relationships of perceived motivational climate to cohesion and collective efficacy in elite female teams. *Journal of Applied Sport Psychology*, 18(3), 201-218.
Ivancevich J. M., Matteson M. T., & Konopaske R. (1990). *Organizational behavior and management*.
Jones D. W. (1996). Empowered teams in the classroom can work. *The Journal for Quality and Participation*, 19(1), 80.
Kahn W. A. (1990). Psychological conditions of personal engagement and disengagement at work. *Academy of Management Journal*, 33(4), 692-724.
Karau S. J., & Williams K. D. (1993). Social Loafing: Meta A.-Analytic Review and Theoretical Integration. *Journal of Personality and Social Psychology*, 65(4), 681-706. https://doi.org/10.1037/0022-3514.65.4.681
Kozlowski S. W., Chao G. T., Grand J. A., Braun M. T., & Kuljanin G. (2013). Advancing multilevel research design: Capturing the dynamics of emergence. *Organizational Research Methods*, 16(4), 581-615.
Kozlowski S. W., & Ilgen D. R. (2006). Enhancing the effectiveness of work groups and teams. *Psychological Science in the Public Interest*, 7(3), 77-124.
Kratzer J., Leenders R. T. A. J., & Van Engelen J. M. L. (2010). The social network among engineering design teams and their creativity: A case study among teams in two product development programs. *International Journal of Project Management*, 28(5), 428-436. https://doi.org/10.1016/j.ijproman.2009.09.007
Levine J. M., Resnick L. B., & Higgins E. T. (1993). Social foundations of cognition. *Annual Review of Psychology*, 44(1), 585-612.
Martens R., & Peterson J. A. (1971). Group cohesiveness as a determinant of success and member satisfaction in team performance. *International Review of Sport Sociology*, 6(1), 49-61.
Mikkelson A. C., York J. A., & Arritola J. (2015). Communication competence, leadership behaviors, and employee outcomes in supervisor-employee relationships. *Business and Professional Communication Quarterly*, 78(3), 336-354.
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. 59278, V007T06A026.
Muchinsky P. M., & Monahan C. J. (1987). What is person-environment congruence? Supplementary versus complementary models of fit. *Journal of Vocational Behavior*, 31(3), 268-277. https://doi.org/10.1016/0001-8791(87)90043-1
Mullen B., & Copper C. (1994). The Relation Between Group Cohesiveness and Performance: An Integration. *Psychological Bulletin*, 115(2), 210-227.
Nemeth C. J. (1986). Differential contributions of majority and minority influence. *Psychological Review*, 93(1), 23.
Neuman G. A., Wagner S. H., & Christiansen N. D. (1999). *The Relationship Between Work-Team Personality Composition and the Job Performance of Teams*. 24(1).

Orasanu J. (2022). *Shared mental models and crew decision making*. 1066-1066.

Paulus P. B. (2000). Groups, teams, and creativity: The creative potential of idea-generating groups. *Applied Psychology*, 49(2), 237-262. https://doi.org/10.1111/1464-0597.00013

Paulus P. B., & Brown V. R. (2003). *Enhancing Ideational Creativity in Groups* (1-Book, Section). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195147308.003.0006

Piasentin K. A., & Chapman D. S. (2007). Perceived similarity and complementarity as predictors of subjective person-organization fit. *Journal of Occupational and Organizational Psychology*, 80(2), 341-354. https://doi.org/10.1348/096317906x115453

Saavedra R., Earley P. C., & Van Dyne L. (1993). Complex interdependence in task-performing groups. *Journal of Applied Psychology*, 78(1), 61.

Salas E., Shuffler M. L., Thayer A. L., Bedwell W. L., & Lazzara E. H. (2015). Understanding and improving teamwork in organizations: A scientifically based practical guide. *Human Resource Management*, 54(4), 599-622.

Shin Y., & Eom C. (2014). Team Proactivity as a Linking Mechanism between Team Creative Efficacy, Transformational Leadership, and Risk-Taking Norms and Team Creative Performance. *The Journal of Creative Behavior*, 48(2), 89-114. https://doi.org/10.1002/jocb.42

Singh J., & Fleming L. (2010). Lone Inventors as Sources of Breakthroughs: Myth or Reality? *Management Science*, 56(1), 41-56. https://doi.org/10.1287/mnsc.1090.1072

Tjosvold D. (1997). Conflict within interdependence: Its value for productivity and individuality. *Using Conflict in Organizations*, 23.

Wall J. A. Jr, & Callister R. R. (1995). Conflict and its management. *Journal of Management*, 21(3), 515-558.

Wall V. D. Jr, & Nolan L. L. (1986). Perceptions of inequity, satisfaction, and conflict in task-oriented groups. *Human Relations*, 39(11), 1033-1051.

West M. A. (1990). *The social psychology of innovation in groups*.

Zemliansky P. (2012). Achieving experiential cross-cultural training through a virtual teams project. *IEEE Transactions on Professional Communication*, 55(3), 275-286.