7-12-2024 12:30 PM

Examining the Role of Core Self Evaluations, Core Group Evaluations, and Individual and Team Referent Psychological Safety on Team Member Behaviours and Performance

Helen H. Lee, *The University of Western Ontario*

Supervisor: Allen, N., *The University of Western Ontario*

A thesis submitted in partial fulfillment of the requirements for the Doctor of Philosophy degree in Psychology

Abstract

Psychological safety has been defined by Edmondson (1999) as "a shared belief that the team is safe for interpersonal risk taking" (p.354). Psychological safety has been found to predict a host of beneficial outcomes for teams and organizations. However, the items in the most used measure of psychological safety developed by Edmondson (1999) demonstrates inconsistency in terms of mixing referents at both the individual and team level. Individual referent (IR) psychological safety appears to be more likely to reflect individuals' perceptions of psychological safety rather than encompassing the psychological safety of a team as a whole. On the other hand, team referent (TR) psychological safety appears to be an appropriate conceptualization of team psychological safety that aligns more closely with Edmondson's (1999) definition of team psychological safety as a shared climate of the team.

I examined the validity of an IR measure and TR measure of psychological safety by modifying the items used in Edmondson's (1999) measure. Overall, the findings suggested that although IR psychological safety and TR psychological safety are related, they demonstrate distinctiveness in their level of within-group agreement, in how they relate to other variables, and in evidence of incremental validity when the referent is aligned with the level of analysis. Study 1 and 2 both revealed that consistent across both studies, CSE appears to be an antecedent of IR psychological safety at the individual level, but not for TR psychological safety. At the group level, the findings relating to the relationship between CSE and TR psychological safety were unexpected but interesting. The findings suggested that higher CSE predicts lower TR psychological safety when accounting for the effect of CGE at the team level. CGE appears to be a robust predictor of IR and TR psychological safety at both the team and individual level.

In general, IR psychological safety appears to be a stronger predictor of individual-level team members' behaviour (task performance, citizenship behaviour) whereas TR psychological safety appears to be a more potent predictor of behaviours at the team level (task performance, citizenship behaviour, counterproductive work behaviours). There were two team outcomes examined: perceived team performance (which was reported by team members) and graded team effectiveness (which was the final grade on a team project). For

perceived team performance, TR psychological safety did not demonstrate incremental validity in perceived team performance above and beyond IR psychological safety as predicted. No relation was found between IR psychological safety and TR psychological safety and graded team effectiveness.

The roles of IR and TR psychological safety as a mediator were also examined. At the individual level, IR psychological safety plays a mediating role in the relationship between CSE and both task performance and citizenship behaviors. TR psychological safety plays a mediating role in the relationship between CGE and both task performance and citizenship behaviors at the individual level. Additionally, TR psychological safety mediates the relationship between CGE and both counterproductive behaviours and perceived team performance ratings at the team level.

Keywords

psychological safety, core self-evaluations, core group evaluations, team composition models, teamwork

Summary for Lay Audience

Psychological safety describes a team environment where members feel safe taking risks, such as sharing dissenting opinions, proposing a new idea, asking for help, or making suggestions, without fear of harsh judgment, rejection, or negative repercussions. When team members feel psychologically safe, it often leads to better outcomes for both individuals, teams, and the organization. However, how we measure psychological safety has been a bit thorny, as it is typically measured by mixing questionnaire items that refer to the self with those that refer to team members as a whole.

In my research, I explored two different ways to measure psychological safety: one that references the self (IR - Individual Referent) and one that references the team as a whole (TR - Team Referent). The individual measure (IR) tends to reflect personal feelings, meaning it might not fully capture the team's shared sense of safety. On the other hand, the team measure (TR) better aligns with the idea of psychological safety as something the entire team experiences together.

The findings suggest that individuals who view themselves more highly (CSE - Core Self-Evaluations) were more likely to report higher IR and TR psychological safety. It was also found that the team's shared views about their value or capabilities as a team (CGE - Core Group Evaluations) are crucial in predicting both IR and TR psychological safety, regardless of whether the data was examined individually or by groups. When looking at the group as a whole, teams with members who had higher CSE (sense of worth and confidence in themselves and their abilities) tended to rate their team as being less psychologically safe when we also took into account the influence of CGE. This could mean that when there are very confident people on a team, it might create pressure for others to conform or hold back their opinions, making the team less safe.

When it comes to team behaviours and performance, the individual-focused measure (IR) is a better predictor of how team members perform tasks and engage in helpful behaviors when the data is examined at the individual level, while the team-focused measure (TR) is more predictive of higher incidences of the same behaviours and lower incidences of unhelpful behaviours that harm the team. While psychological safety influenced how teams

perceived their performance, it didn't significantly impact the actual grades teams received on their projects. Overall, my research showed that IR and TR safety, while connected, are distinct concepts that need to be considered separately.

I also looked at how feeling safe (psychological safety) acts as a link that explains how CSE and CGE influences team member behaviours and performance. For individuals, IR psychological safety helps explain how their view of themselves (CSE) leads to better performance on tasks and the likelihood they will engage in more helping behaviors. In other words, CSE predicts higher levels of IR psychological safety, which in turn, predicts better task performance and helping behaviours. Moreover, TR psychological safety helps explain how CGE leads to better performance and helpful behaviors among individuals. That is, CGE predicts higher levels of TR psychological safety, which in turn predicts better task performance and helping behaviours. For teams, TR psychological safety helps explain how the level of CGE predicts unhelpful behaviours and team performance. That is, CGE predicts higher levels of TR psychological safety, which in turn, predicts lower unhelpful behaviours and higher team performance.

Co-Authorship Statement

Study 1 in this thesis includes data that I presented at a conference. The paper that was submitted for the conference presentation was completed in collaboration with Natalie Allen, Kyle Cameron, and Hayden Woodley as co-authors. The citation for this conference paper is as follows:

Lee, H., Allen, N., Cameron, K., & Woodley, H. (2013, June 8-11). *Rethinking the measurement of psychological safety: referent use validity and the role of core self-evaluations* [Paper Presentation]. Administrative Sciences Association of Canada 2013 Convention, Calgary, AB, Canada.

Acknowledgments

First and foremost, I want to thank Jesus Christ for without whom I would not be here today. I thank God for carrying my burdens, giving me His strength and comfort, and guiding me through this journey. There were times when it felt like I was walking around the same mountain again and again, facing a journey of wilderness. Yet, I am grateful to God that through this refining process, I emerged with a stronger belief that He never left my side.

I am deeply grateful to my family for their unwavering prayers and constant encouragement. Thank you, dad, for teaching me how to write and for always encouraging me to question my assumptions. Thank you, mom, for modeling strength and perseverance. To my sister, Grace, thank you for loving me unconditionally, never failing to make me laugh, and always reminding me to trust God.

My heartfelt thanks also go to my Calvary Chapel Oakville family, especially Pastor Paul Fernando and Gabriela Fernando, for their ceaseless prayers and for being the steadfast support I needed to cross the finish line.

To my supervisor, Natalie Allen, your support, genuine care, and acceptance of me has meant more to me than words can express. Thank you for believing in my ability to finish my PhD, even when I struggled to believe it myself.

A special thanks to John Meyer, who was not only on my PhD committee but also on my master's committee. Having you as an instructor sparked my deep interest in leadership and motivation which has not only inspired my career aspirations but principles and core values to live by. I am sincerely grateful for the opportunity to have learned from you.

Blair Evans, thank you for your help during my departmental defense and for assisting me in organizing my thoughts and the necessary steps I needed to complete when I needed it most. Your guidance was truly crucial.

Rick Goffin, I am grateful for the extensive feedback you provided on the dissertation, which was integral to my work. As a professor, you helped me develop a deeper appreciation for logic, structure, and critical thinking—skills that have greatly shaped my approach to my research and work.

I also extend my gratitude to my senate exam committee members—John Meyer, Douglas Brown, Victoria Esses, and Shannon Sibbald—for their insightful feedback and the engaging and thoughtful way that they delivered it. It truly made the dissertation defense an enjoyable experience.

I am incredibly honored to graduate as an Industrial/Organizational Psychology PhD from Western University, given the caliber of the faculty. I am thankful to all the instructors I had who taught me and inspired my interest in understanding what makes people thrive at work.

Thank you to all my fellow Western psychology colleagues over the years who I came to grow very fond of. A special thank you to my lab and officemate, Hayden Woodley, who kept me accountable and invested so much time in helping me prepare. Thank you, Eva Yeung, for your time and giving me such helpful feedback on my dissertation presentation. Julia McMenamin, thank you for helping to organize my defense celebration, making the completion of this journey all the more meaningful. I appreciate how we have kept in touch despite our busy schedules. Kyle Cameron and Natasha Ouslis, thank you for your friendship and encouragement throughout the years.

I am also deeply grateful to Josh Lee, a fellow Western PhD graduate and someone I consider a brother, for helping me stay grounded through our many meaningful conversations about navigating this world with an eternal purpose, and for helping me find strength and hope through difficult challenges.

I also want to thank all my colleagues at HRSG, especially Evanya Musolino and Twiladawn Stonefish, who have been my cheerleaders throughout this journey and have always been a joy to work with.

Thank you to all my best friends, whom I consider sisters. Thank you, Archana Narayan, whose seen every single version of me, yet has been there for me for so many years. Your steadfast support and friendship are invaluable, and you never once faltered in your belief that I would complete this milestone. I also want to thank the Ault family and Alanna, Alicia, and Amanda—my set of chosen sisters—for the love you pour onto me. Thank you, also, to Kim Lee, for having the biggest heart and your consistent presence in my life, and Talisha Matheson, for always being a safe confidante where I can be my most vulnerable. I can

always count on you to be an empathetic voice of reason whenever I need inspiration and a reality check.

I want to express my deep appreciation to my counselor, Marie Allen, who helped me cultivate a godly mindset toward this achievement and encouraged me to maintain an eternal perspective—by doing my best while surrendering the outcome to God's hands and drawing strength from Him.

And last, but certainly not least, to Sherwyn Simmons, my good and perfect gift from heaven—it is no coincidence that we met closer to the end of this journey. I truly believe that I needed your God-sent love and support to finally bring this chapter to a close.

Table of Contents

Abstract ii

Summary for Lay Audience iv

Co-Authorship Statement vi

Acknowledgments vii

Table of Contents x

List of Tables xiii

List of Figures xv

List of Appendices xvi

Chapter 1 1

1 Introduction to Psychological Safety 1

1.1 Core self-evaluation (CSE) as Antecedents of Psychological Safety 3

1.2 Conceptualization and Measurement of Psychological Safety 7

1.3 Aggregation-based Operationalizations of Psychological Safety 8

1.4 Level of Measurement Issue: Target Referent Inconsistency 9

Chapter 2 11

2 Study 1 Introduction 11

2.1 Study 1 Objectives and Hypotheses 11

2.2 Study 1 Method 12

2.2.1 Participants and Procedure 12

2.2.2 Study 1 Measures 12

2.2.3 Student Participation in the Study 15

2.2.4 Study 1 Participant Characteristics 15

2.2.5 Study 1 Interclass Correlation Coefficients (ICC1 and ICC2) 15

2.2.6 Study 1 Confirmatory Factor Analysis 17

2.2.7 Study 1 Multilevel Confirmatory Factor Analysis 19
2.2.8 Study 1 Descriptive Statistics 20
2.2.9 Relations Between CSE and IR and TR Psychological Safety 22
2.2.10 Examining the Relation Between CSE and TR Psychological Safety using Multilevel Linear Modeling (MLM) 22
2.3 Study 1 Discussion 23
2.3.1 Study Limitations 24
2.3.2 Implications and Future Research Directions 24
Chapter 3 26
3 Study 2 26
3.1 Study 2 Hypotheses and Rationale 26
3.1.1 Aggregation-based Operationalization of TR Psychological Safety 26
3.1.2 Factor Structure of IR and TR Psychological safety 27
3.1.3 CSE as an Antecedent of IR and TR Psychological Safety 27
3.1.4 CGEs as Antecedents of IR and TR Psychological Safety 28
3.1.5 IR and TR Psychological Safety and Team Member Behaviours 29
3.1.6 IR and TR Psychological Safety and Team Performance 32
3.1.7 Incremental Validity 34
3.1.8 Psychological safety as a mediator 34
3.2 Study 2 Method 36
3.2.1 Participants and Procedure 36
3.2.2 Study 2 Measures 38
3.3 Study 2 Results 42
3.3.1 Study 2 Participant Characteristics 42
3.3.2 Test of Questionnaire Order Effect 42
3.3.3 Study 2 Descriptive Statistics 43

3.3.4 ICCs for IR and TR Psychological Safety 46

3.3.5 Factor Structure of IR and TR Psychological Safety 46

3.3.6 Examining CSE as an Antecedent of IR and TR Psychological Safety Using Correlational Analysis 49

3.3.7 Examining CGE as an Antecedent of IR and TR Psychological Safety Using Correlational Analysis 49

3.3.8 Examining CGE and CSE as Predictors of IR and TR Psychological Safety Using Multilevel Modeling 49

3.3.9 IR and TR psychological Safety and Team Member Behaviours 53

3.3.10 IR and TR Psychological Safety, Perceived Team Performance and Graded Team Effectiveness 55

3.3.11 Incremental Validity of IR and TR psychological safety 56

3.3.12 Psychological Safety as a Mediator 60

3.4 Study 2 Discussion 65

4 General Discussion 68

4.1 Contributions of the Current Research 68

4.2 Implications of Findings 70

4.3 Limitations and Implications for Future Research 70

4.3.1 Further Research Directions 73

4.4 Implications for Practice 75

References 77

Appendices 85

Curriculum Vitae 90

List of Tables

Table 1. *Study 1 Original Edmondson (1999), Individual-referent and Team-referent Psychological Safety Items* 14

Table 2. *Intraclass Correlations for IR and TR Psychological Safety (Study 1)* 17

Table 3. *CFA for IR Psychological Safety with Selected MIs Modeled (Study 1)* 18

Table 4. *MLCFA for TR Psychological Safety with Selected MIs Modeled (Study 1)* 20

Table 5. *Means, Standard Deviations, and Correlations for Individual-Level Variables* 21

Table 6. *Means, Standard Deviations, and Correlations for Aggregated Team-Level Variables* 21

Table 7. S*tudy 2 Original Edmondson (1999), Individual-referent and Team-referent Psychological Safety Items* 40

Table 8. t-tests Comparing Questionnaire Versions on IR and TR Psychological Safety (Study 2) 42

Table 9. *Means, Standard Deviations, and Correlations for Individual-Level Variables (Study 2)* 44

Table 10. *Means, Standard Deviations, and Correlations for Team-Level Variables (Study 2)* 45

Table 11. *Intraclass Correlations for IR and TR Psychological Safety (Study 2)* 46

Table 12. *CFA Model Fit Statistics for Measurement Models Tested in Study 2* 48

Table 13. *MLM Regressions Examining CSE on IR and TR Psychological Safety* 50

Table 14. *MLM Regressions Examining CGE on IR and TR Psychological Safety* 51

Table 15. *MLM Regressions Examining CSE and CGE as Predictors of IR and TR Psychological Safety* 52

Table 16. *Summary of the Hierarchical Linear Model Results for the Incremental Contribution of IR Psychological Safety Beyond TR Psychological Safety at the Individual Level* 58

Table 17. *Summary of the Hierarchical Linear Model Results for the Incremental Contribution of TR Psychological Safety Beyond IR Psychological Safety at the Team Level* 59

Table 18. *Summary of Hypotheses and Findings for Study 2* 62

List of Figures

Figure 1. *Study 2 Proposed Model* 36

Figure 2. *Study 2 Participant Inclusion and Exclusion* 37

List of Appendices

Appendix A: Core Self Evaluations Measure 85

Appendix B: Core Group Evaluations Measure 86

Appendix C: Team Member Behaviour Measures 87

Appendix D: Perceived Team Performance Measure 88

Appendix E: Ethics Approval Letter 89

Chapter 1

1 Introduction to Psychological Safety

Psychological safety was initially introduced by Schein and Bennis (1965) who suggested that a psychologically safe environment is a context which "encourages provisional tries and which tolerates failure without retaliation, renunciation, or guilt" (p. 45). Kahn (1990) defined psychological safety as "feeling able to show and employ one's self without fear of negative consequences to self-image, status, or career" (p. 708). In Edmondson's (1999) seminal paper, psychological safety is defined as "a shared belief that the team is safe for interpersonal risk taking" (p.354).

In a psychologically safe team climate, members feel comfortable engaging in behaviours that may carry interpersonal risks, such as sharing their thoughts, their ideas, and concerns. Such behaviours may involve risks of potential threats to one's social standing, reputation, relationships, or self-image (Edmondson, 1999). These risks may stem from the possibility of facing harsh criticism, rejection, or negative consequences as a result of being perceived as incompetent, ignorant, negative or disruptive (Edmondson, 1999; Edmondson, 2012; Edmondson & Bransby, 2023). Psychological safety, described by Edmondson and Bransby (2023) as "a state of reduced interpersonal risk" (p. 55), is essential in today's complex, dynamic workplace for enabling behaviours that facilitate learning, change, and innovation (Edmondson & Lei, 2014; Frazier et al., 2017).

Edmondson (1999) has argued that understanding psychological safety is crucial when considering how modern team dynamics have shifted towards more fluid and adaptive forms of collaboration. A team has been defined in academic literature as a collection of individuals who work together towards a common goal, sharing responsibilities, resources, and expertise to achieve a shared outcome (e.g., Cohen & Bailey, 1997; Sims, Salas, & Burke, 2005). Traditionally, teams in organizations have been stable entities with clear boundaries, often focused on producing an identifiable product, service, or decision (Hackman, 2012). However, given that modern teams are typically more dynamic and fluid, with team members often belonging to multiple teams

and handling a wider variety of tasks with high levels of complexity and uncertainty, Edmondson introduced the term “teaming” to describe the active process of working together interdependently, characterized by learning, adaptability, and problem-solving in real-time (Edmondson, 2012). In knowledge-intensive environments where rapid change and complex challenges require flexibility and continuous learning, fostering a psychologically safe climate is considered essential for effective teaming by enabling open communication, clear thinking, productive conflict, innovation, accountability, and resilience in the face of challenges (Edmondson, 2012).

Corroborating this perspective, research has demonstrated that psychological safety enhances team performance by encouraging team members to engage in activities that help them acquire, share, and combine knowledge—collectively referred to as team learning behaviours (e.g., Edmondson, 1999; Kim et al., 2020). Specific team learning behaviours include information sharing, feedback seeking, asking for help, discussing errors, and experimentation which involves team members trying out new approaches, taking calculated risks, and learning from their mistakes (Edmondson, 1999).

Psychological safety has garnered the attention of researchers in both organizational and academic contexts. High profile organizations such as Google and IDEO have adopted practices intended to build psychological safety in teamwork (IDEO, 2019). In a research initiative known as *Project Aristotle*, Google’s People Operations investigated what characteristics differentiated high-performing teams from low-performing ones (Rozovsky, 2015). Data was collected from over 200 double-blind interviews and surveys assessing more than 250 group attributes of over 180 teams active at Google. Five group variables were identified as significantly affecting multiple outcome metrics across different types of teams (Google, n.d.; Rozovsky, 2015). These five variables were psychological safety, dependability, structure/clarity, meaning, and impact (Rozovsky, 2015). Psychological safety was found to be the most important group dynamic for the success of teams at Google. Teams with higher psychological safety were more likely to bring in higher revenue and were twice as more likely to be rated as effective by executives (Google, n.d.). As a result of these findings, Google has employed measures of psychological safety and has implemented practices to enhance

psychological safety, such as kicking off team meetings by sharing risks that team members have taken in a previous week (Google, n.d.).

Psychological safety has also garnered considerable attention in academic research and has been linked to several beneficial team-related outcomes. In a meta-analysis conducted by Frazier et al. (2017), psychological safety was found to be positively and significantly related to engagement ($\hat{\rho}$ = .44), task performance ($\hat{\rho}$ = .29), and satisfaction ($\hat{\rho}$ = .69) at the group level. Psychological safety has also been found to facilitate the sharing of ideas and actions such as suggestions for making organizational improvements (e.g., Collins & Smith, 2006; Siemsen et al., 2009) and the development of new products and services (Baer & Frese, 2003). In the Frazier et al. (2017) meta-analysis, strong relationships were also found between psychological safety and group information sharing ($\hat{\rho}$ = .50) and learning behaviour ($\hat{\rho}$ =.52). In addition, a weaker, but significant, positive relation was found between psychological safety and creativity ($\hat{\rho}$ = .29) at the group level of analysis.

With respect to antecedents, positive leader relations ($\hat{\rho}$ = .39) (operationalized as transformational leadership and trust in one's leader) were found to relate positively to psychological safety in groups (Frazier et al., 2017). In addition, work context supportiveness ($\hat{\rho}$ = .51) was also positively related to psychological safety at the group level. In terms of work design characteristics, the authors found positive group-level relations between psychological safety and each of autonomy ($\hat{\rho}$ = .35), interdependence ($\hat{\rho}$ = .40), and role clarity ($\hat{\rho}$ = .51).

1.1 Core self-evaluation (CSE) as Antecedents of Psychological Safety

A gap in the literature pertains to how team member characteristics contribute to the emergence of team-level psychological safety. Indeed, it is surprising how little attention has been paid to this, given that *team composition*, or the varying levels of an attribute across team members, is one of the most frequently studied characteristics of teams (Hollenbeck et al., 2004). Research on team composition is typically focused on understanding which attributes of individual team members facilitate team performance

and, accordingly, such research can have implications for team-member selection and team composition. In the meta-analysis conducted by Frazier et al. (2017), three personality traits (proactive personality, emotional stability, learning orientation) were found to be positively related to psychological safety at the individual level of analysis. However, at the group level, there were only enough studies to meta-analytically examine the relation between learning orientation and psychological safety ($\hat{\rho}$ = .40; Frazier et al., 2017).

Given that individual team member traits may contribute to the characteristics of a team, applying our knowledge of team composition to the study of psychological safety represents a meaningful avenue of research. Thus, a goal of Study 1 was to examine the relation between core self-evaluations and psychological safety at the group level of analysis. *Core self-evaluations* (CSE) reflect beliefs about one's competence and one's ability to control his or her life, as well as a general feeling that life will turn out well for oneself (Judge, Locke, & Durham, 1997). CSE represent a broad, higher-order personality composite of generalized self-efficacy, self-esteem, low neuroticism, and locus of control (Judge, 2009). CSE have been found to predict variance in several positive outcomes, such as motivation and job performance, over and above narrower traits such as self-esteem (Judge, 2009). Although the composition of CSE in teams has been examined as a predictor of team performance (Haynie, 2012), the role of CSE in the emergence of psychological safety has not yet been investigated.

CSE has been found to predict a number of outcomes (e.g., Judge, 2009) better than any of the core narrow traits which comprise CSE (i.e., self-esteem, locus of control, generalized self-efficacy, and emotional stability; Chang et al., 2012; Judge, 2009). There has been some speculation about *how* CSE might affect these outcomes. For example, Judge et al. (1997) suggested that people with high CSE, who tend to perceive themselves as possessing both high self-worth and the ability to cope with the demands of life, may experience a spillover effect of this positive frame of mind to situations that they encounter. Additionally, CSE may indirectly influence outcomes by affecting the cognitions people have about their environment (Judge et al., 1997). In previous research, Elliot and Thrash (2002) used an approach/avoidance framework to explain the basic

personality dimension of extraversion. In their framework, an *approach temperament* is a biologically-based sensitivity to positive or desirable stimuli accompanied by a behavioural predisposition toward such stimuli. In contrast, an *avoidance temperament* refers to a biologically-based sensitivity to negative or undesirable stimuli accompanied by a behavioural predisposition away from such stimuli.

Ferris et al. (2011) applied this framework to CSE theorizing and found that high CSE scores were related to a strong approach temperament and a weak avoidance temperament among a sample of students and a sample of full-time working adults. Their findings suggest that individuals with high CSE may be characterized as being more sensitive to positive stimuli and less sensitive to negative stimuli whereas low CSE individuals may be more sensitive to negative stimuli and less sensitive to positive stimuli. Thus, CSE may influence the way in which individuals perceive attributes in their environment, including how psychologically safe the team environment is. In a team setting, team members with higher CSE may be less likely to worry about receiving negative feedback from other team members because they are less sensitive to negative or undesirable information compared to individuals with low CSE. As a result, team members with high CSE may be more likely to take more interpersonal risks (e.g., soliciting feedback). On the other hand, individuals with low CSE may be more likely to avoid interpersonal risks in an effort to avoid negative feedback from other team members due to their heightened sensitivity to undesirable stimuli (e.g., negative feedback).

Research by Fast et al. (2014) has shown a link between lower self-efficacy and the avoidance of feedback. Their research showed that compared to managers with high managerial self-efficacy, managers with low managerial self-efficacy were less likely to solicit employee input and more likely to minimize voice as a tactic to compensate for a threatened ego (Fast et al., 2014).

Research has also examined a link between two of the narrow traits of CSE (i.e., locus of control and emotional stability) and psychological safety. Triplett and Loh (2018), for example, showed that an external locus of control at work (i.e., a perceived

lack of personal locus of control with respect to work life) was negatively related to psychological safety among workers in mining, oil, and gas organizations. The authors argued that individuals with an internal work life locus of control are more likely to engage in proactive behaviours and open communication than those with an external work life locus of control, resulting in higher psychological safety. With respect to emotional stability, Seyda and Tabancali (2020) reported a positive correlation between emotional stability and psychological safety in a sample of 475 public primary school teachers in Istanbul. This finding confirmed previous research by Edmondson and Mogelof (2005) which demonstrated that neuroticism was linked to lower psychological safety among individuals in innovation teams.

Based on the body of research described above, and my interest in group phenomena, Study 1 was designed to examine CSE as an antecedent of team psychological safety at the group level of analysis. Group-level CSE was operationalized by the average CSE of individual members in a team. For team composition variables such as the personality traits of team members, the team-level form of the construct does not require within-group consensus on scores. This is because the effects of the overall "amount" of CSE, within each team, are of interest, and CSE are not theorized as a shared construct that converges within a team. This approach to conceptualizing team CSE using the average team score is similar to that used in previous research examining team composition effects (e.g., Bell, 2007) and reflects an additive model of a group-level construct (Chan, 1998).

At the group level, the effect of high CSE among team members may be manifested in the group context through consistent patterns of group behaviours that promote psychological safety. Team members who have greater mean CSE may be more likely to engage in interpersonal risk-taking behaviours due to their lowered sensitivity to negative stimuli (Ferris et al., 2011). For example, team members who are higher in CSE may frequently engage in feedback seeking behaviours. This may demonstrate to other team members that it is acceptable to engage in similar behaviours and encourage them to follow suit, regardless of their individual level of CSE. As a result, teams with higher

CSE overall may develop group norms which reinforce interpersonal risk-taking behaviours and promote perceptions of psychological safety at the team level of analysis.

1.2 Conceptualization and Measurement of Psychological Safety

Although initial definitions were focused on psychological safety as an individual-level perception, Edmondson and Lei (2014) have suggested that psychological safety "can best be considered a phenomenon that lives at the group level" (p. 37). In her paper, Edmondson (1999) emphasized that team psychological safety "must characterize the team rather than individual members of the team, and team members must hold similar perceptions of it" (p. 354-355) in order to be considered a group-level construct. The majority of organizational psychology research examining psychological safety uses Edmonson's (1999) definition, which refers to psychological safety as the extent to which a team's climate encourages interpersonal risk-taking behavior such as bringing up problems, reporting mistakes, or seeking help (Edmondson, 1999).

Although psychological safety has received substantial attention linking it to a host of antecedents and outcomes, there appears to be an issue with regard to its measurement. The items contained in the most widely used measure of psychological safety (Edmondson, 1999) are inconsistent in terms of their target referent (i.e., individual or team). This is a potential limitation of the measure as the referent used may alter the conceptualization of psychological safety as a construct. For example, it is possible that one may feel a high level of psychological safety as *an individual within a team* but may simultaneously observe that *team members as a unit* generally do not perceive the team climate as safe. It is conceivable that an individual's own feeling of psychological safety may contribute to, and positively relate to, the team's overall perceptions of psychological safety. However, if psychological safety is to be viewed as a perception that is shared by team members as proposed by Edmondson (1999), it may be more appropriate to assess team members' perceptions of whether *team members* generally feel psychologically safe.

Given that the interpretation of the construct of psychological safety may be altered at different levels of the referent, it is necessary to investigate psychological safety using a consistent referent. Another goal of this dissertation, therefore, was to determine the appropriate referent for the construct of psychological safety. To achieve this aim, I conducted a comparative examination of two operationalizations of psychological safety: (a) the *direct consensus model* in which team psychological safety is conceptualized as the average level of perceptions of psychological safety in a team using an *individual referent* and (b) the *referent-shift model* in which team psychological safety is conceptualized as the average level of perceptions of psychological safety in a team using a *team referent* (Chan, 1998). In the following sections, more detailed discussions of aggregation-based operationalizations of psychological safety and the issue of mixed referents are presented.

1.3 Aggregation-based Operationalizations of Psychological Safety

As noted above, *team* psychological safety was introduced into the organizational literature by Edmondson (1999) as a shared, group-level construct. Psychological team constructs such as team psychological safety are typically measured by collecting data from individual team members and then aggregating across members to define the team-level construct. In light of the fact that the aggregated form of the individual-level data is an imperfect analogue of the team-level construct, Chan (1998) proposed an organizing framework outlining a typology of team composition models to guide researchers in the process of validating various operationalizations of focal group-level constructs. The two operationalizations in Chan's (1998) framework that are most closely aligned to Edmondson's (1999) definition of team psychological safety are the: (a) direct consensus model and (b) referent-shift consensus model.

When conceptualized using the direct consensus model, psychological safety is operationalized by aggregating individual team members' appraisals of how psychologically safe they feel as individuals in their team. In this case, lower-level individual attributes are conceptualized as "functionally isomorphic to another form of the construct at the higher level" (Chan, 1998 p. 237). Thus, if the direct consensus model

is the appropriate conceptualization of psychological safety, this would mean that individual-level measurements of psychological safety would represent shared perceptions of psychological safety at the team level. In other words, individual team members' perceptions of psychological safety in the team would be considered as a functional analogue of the team's perceptions of psychological safety. In order to justify representing psychological safety at the team level as an aggregate of individual-level scores, evidence of within-group consensus (as indexed by within-group agreement of individual-level scores) on the measure of individual-level psychological safety is required (Chan, 1998). Moreover, adequate between-group variability in scores on individual-level psychological safety also should be demonstrated to justify aggregating the individual-level data to the team level (Chan, 1998).

On the other hand, in the referent-shift consensus model, team psychological safety is operationalized as the average of each team member's individual ratings of the *team's* perception of psychological safety. Thus, rather than respondents reporting their own perceptions about psychological safety in the team using an individual referent, the referent in the referent-shift model is the team. Respondents report how they believe team members, *as a group*, perceive the psychological safety climate. In contrast to the direct consensus model, the referent-shift model conceptualizes the lower-level individual attributes (e.g., individual perceptions of psychological safety) that are being assessed for consensus as conceptually distinct from the higher-level team construct (e.g., team perceptions of psychological safety). Similar to the criteria for establishing the validity of a direct consensus model, the referent-shift operationalization requires agreement within the group on individual-level ratings and sufficient between-group variability on measures of psychological safety using a team referent.

1.4 Level of Measurement Issue: Target Referent Inconsistency

Edmondson's (1999) measure of psychological safety does not account for the distinction between the direct consensus and referent-shift models; that is, items in this measure use referents appropriate for both models. There are seven items in total in Edmondson's (1999) measure which require respondents to report their perceptions of

psychological safety using: (a) an individual referent (i.e., "Working with members of this team, *my* unique skills and talents are valued") or (b) a team referent (e.g., "*Members of this team* are able to bring up problems and tough issues"). Three of the items include an individual referent and two of the items include a team referent. Two of the remaining items do not include an explicit referent but are likely to be treated as individual-referent items since one's own perception of how safe they feel is likely to be more salient and easier to access (e.g., *"It is safe to take a risk on this team."*, *"It is difficult to ask other members of this team for help."*).

Using multiple referents for items in a single scale may be problematic because it is possible that an individual's experience of psychological safety is not isomorphic to the shared psychological safety that characterizes the group. Given Edmondson's conceptualization of psychological safety as a *shared* belief about the team, it would seem that the appropriate target referent is either the *team* or *the members of a team*, considered collectively. This would ensure that individuals are responding with regard to the team as a whole, rather than their individual experiences within the team. When the 'team' is used as the target referent, it would then seem reasonable to aggregate the responses of individual team members and expect that they reflect the team's psychological safety climate. However, because of the inconsistency of the referent used in the measure, it is not clear what conceptualization the measure actually captures. At a minimum, it would appear that "mixing" referents, within a single scale, places some doubt on the construct validity of the psychological safety measure.

Chapter 2

2 Study 1 Introduction

2.1 Study 1 Objectives and Hypotheses

In Study 1, I examined the team-level relation between CSE and psychological safety when the latter is measured with both individual referent and team referent items. I propose that teams with higher levels of CSE are more likely to exhibit higher levels of psychological safety. Thus, the following hypothesis was proposed:

> *H1: CSE scores will be positively related to team psychological safety at the team level, regardless of the referent used.*

To examine the issue regarding target referent inconsistency, the original psychological safety measure developed by Edmondson (1999) was modified in this study so that one version of the measure contained items referring to individual respondents' perceptions of psychological safety which reflects a direct consensus conceptualization, while another measure contained items addressing the team as a whole, thus reflecting a referent-shift conceptualization. This is an exploratory investigation of a measurement issue that has not yet been studied in relation to psychological safety. Thus, rather than testing specific hypotheses, statistical estimates of within-group agreement and between-group variability were evaluated to assess the appropriateness of the operationalizations. If the direct consensus model is the appropriate operationalization, adequate within-group agreement, and between-group variability on scores on the *individual referent measure* would be observed. If a referent-shift consensus model is more appropriate, adequate within-group agreement and between-group variability on the *team referent measure of psychological safety* most likely would be observed. Ultimately, by examining the within-group agreement and between-group variance, the validity of each model for operationalizing team psychological safety was assessed.

2.2 Study 1 Method

2.2.1 Participants and Procedure

The participants in this study comprised of undergraduate engineering course during the 2012-2013 academic year. During the course, students were required to work in teams in order to complete two major engineering design projects. Data were collected at one time point as part of a larger data collection initiative that involved multiple time points. Teams were formed in the second week of September. At the time of data collection, teams had been working together for approximately 10 weeks and had received a grade for their first team project. The first project required that teams submit a written report describing a design solution that would increase the usability of an object. Paper/pencil questionnaires were administered to participants during the weekly laboratory component of the course. Participation in the study was voluntary and students who participated received extra credit towards their final grade in the course.

2.2.2 Study 1 Measures

2.2.2.1 Core self-evaluations

A 12-item measure developed and validated by Judge et al. (2003) was used to assess CSE. Sample CSE items include, “I am confident I get the success I deserve in life” and “I am filled with doubts about my competence” (reverse-scored). Responses were based on a five-point scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*) scale (α = .85).

2.2.2.2 Individual and Team Referent Measures of Psychological Safety

Psychological safety was measured using adapted measures of Edmondson’s (1999) original 7-item scale. One subset of the sample filled out the individual referent measure. The other subset of the sample filled out the team referent measure. Each student and their team were assigned to a studio where they attended their laboratory sessions for the engineering course. The version of the psychological safety questionnaire that each team received was randomly assigned to studios.

The items for each measure are included in Table 1. In the *individual referent (IR) psychological safety measure*, two items in the scale that were originally worded as "*Members of this team* are able to bring up problems and tough issues" and "People on this team sometimes reject *others* for being different" were adapted in order to keep the items relevant to the individual rather than the team throughout the scale. The items were changed to: "In this team, *I* am able to bring up problems and tough issues" and "People on this team sometimes reject *me* for being different." Given the exploratory nature of this study, I kept two of the items that consisted of an ambiguous referent (i.e., "It is safe to take a risk on this team," and "It is difficult to ask other members of this team for help") as part of the IR psychological safety scale under the assumption that participants are likely to interpret these items in relation to themselves. The Cronbach's alpha for the IR psychological safety scale was .78.

The *team referent (TR) psychological safety measure* was adapted so that items only referred to the team. Sample items on this measure include: "It is difficult for *team members* to ask other members of the team for help." and "If *team members* make a mistake on this team, it is often held against *them*." The Cronbach's alpha for the team referent scale was .75. For both the individual and team referent measures, item responses were based on a scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*).

Table 1. *Study 1 Original Edmondson (1999), Individual-referent and Team-referent Psychological Safety Items*

Original Edmondson (1999) Psychological Safety Scale	Modified Individual-Referent Psychological Safety Scale	Modified Team-Referent Psychological Safety Scale
1. If **you** make a mistake on this team, it is often held against **you**.	1. If **you** make a mistake on this team, it is often held against **you**.	1. If **team members** make a mistake on this team, it is often held against **them.**
2. **Members of this team** are able to bring up problems and tough issues.	2. In this team, **I** am able to bring up problems and tough issues.	2. **Members of this team** are able to bring up problems and tough issues.
3. **People on this team** sometimes reject others for being different.	3. People on this team sometimes reject **me** for being different.	3. **Members of this team** sometimes reject others for being different.
4. It is safe to take a risk on this team.	4. It is safe to take a risk on this team.	4. It is safe for **team members** to take a risk on this team.
5. It is difficult to ask other members of this team for help.	5. It is difficult to ask other members of this team for help.	5. It is difficult for **team members** to ask other members of this team for help.
6. **No one on this team** would deliberately act in a way that undermines my efforts.	6. No one on this team would deliberately act in a way that undermines **my** efforts.	6. **Members of this team** would not deliberately act in a way that undermines **other team members'** efforts.
7. Working with members of this team, **my** unique skills and talents are valued and utilized.	7. Working with members of this team, **my** unique skills and talents are valued and utilized.	7. Working with members of this team, **team members'** unique skills and talents are valued and utilized.

2.2.3 Student Participation in the Study

A total of 410 students participated in at least one of three time points of the larger data collection. The measures in the current study were collected at the second time point. Of the 410 students who participated in at least one time point of the larger data collection effort, seven students dropped out of the course prior to the second time point and another 29 individuals also did not complete the questionnaire at the second time point. A few teams had only one or two respondents who complete the questionnaires. These teams were dropped from the sample, resulting in a final sample of 374 participants in 94 teams. 179 individuals completed the IR psychological safety measure, and 193 individuals completed the TR psychological safety measure. One individual completed the CSE measure, but not the IR psychological safety measure and another individual completed the CSE, but not the TR psychological safety measure. Teams were made up of three to five members each. The average group size in the sample was 4 members.

2.2.4 Study 1 Participant Characteristics

The average age of participants was 18 years, ranging from 16 to 30 years. The majority of the sample was male (71.3%) and Caucasian (55.7%). 12.6% of participants identified as East Asian, 11.1% of participants identified as Arabic or East Indian, 3.2% identified as Southeast Asian, 2.0% identified as Black, .2% identified as Native American, .2% identified as Hispanic, and the remaining 15% identified as Other.

2.2.5 Study 1 Interclass Correlation Coefficients (ICC1 and ICC2)

The goal of the current study was to address the construct validity of the individual referent and team referent measures. As previously discussed, to demonstrate the construct validity of shared-unit team variables, adequate within-team agreement and between-team differences is required. The conventional method of assessing within-group agreement and between-group variability in the team literature is the calculation of *intraclass correlation coefficients* (ICCs). ICC(1) indicates the proportion of variance explained by team membership (Bliese, 2000). ICC(2) is an estimate of the reliability of the group means and is a function of both ICC(1) and the number of raters in a team (Bliese, 1998). An ANOVA framework is used to compute ICCs where the independent variable is the grouping variable (e.g., team ID), and the

dependent variables is the construct being examined (e.g., psychological safety). For team constructs that are theorized to be *shared* among members within each team aggregation of individual scores to the team level is considered to be justified if acceptable ICC values are yielded (Bliese, 2000; James, 1982). Although there is no unanimous agreement on the cut-off for acceptable ICC values for supporting the aggregation of individual scores, ICC(1) has been reported to generally range from 0 to 0.50 with a median of .12 across various studies on climate variables (Bliese, 2000; James, 1982). ICC(1) has been recommended as the primary basis for determining whether to aggregate climate variables (James, 1982). Given that team psychological safety has been defined as a team climate by Edmondson (1999), ICC(1) was used in Study 1 to evaluate support for the validity of the referent-shift and direct consensus models for conceptualizing psychological safety. The benchmark reported by James (1982) of .12 was used as a cutoff score for ICC(1) to indicate a sufficient value for justifying aggregating individual scores to the team level. For ICC(2), researchers often use a cutoff score of .60 to support aggregation of individual scores to the team level (Glick, 1985). The ICC(2) values that were found for both IR and TR psychological safety in this study were lower than .60. However, the ICC(2) value for TR psychological safety was higher than some other studies for psychological safety with similar group sizes that have detected significant relations in group-level data (e.g., Bradley et al., 2012; Greenbaum et al., 2020; Koopman et al., 2016). It is possible that ICC(2) may have been affected by the relatively small team size compared to group sizes typically used in research recommending ICC(2) cutoffs (Bliese & Halverson, 1998; O'Neill et al., 2018). Thus, I did not use a cutoff for ICC(2) and relied on the ICC(1) cutoff to compare the validity of the referent-shift and direct consensus models of psychological safety (Bliese, 1998; Bliese & Halverson, 1998). ICC(1) and ICC(2) values were examined to compare evidence to support aggregation for IR and TR psychological safety scores.

Table 2 contains the computed ICC*s* for the aggregated measures of psychological safety. ICC(1) for IR psychological safety was lower than .12, suggesting insufficient justification for the aggregation of individual team members' scores to the team level [ICC(1) = .04]. On the other hand, ICC(1) for TR psychological safety was sufficiently high, indicating that it is acceptable to aggregate individual responses on the team referent measure to the team level [ICC(1) = .19]. Thus, the results demonstrated that a sufficiently high ICC(1) was yielded for TR psychological safety but not for IR psychological safety. This suggests that a referent-shift

consensus model may be the appropriate operationalization of *shared* psychological safety in a team.

Table 2. ***Intraclass Correlations for IR and TR Psychological Safety (Study 1)***

	ICC(1)	ICC(2)
1. Mean IR Psychological Safety	0.04	0.13
2. Mean TR Psychological Safety	0.19	0.49

Note. n = 43 teams for *Individual Referent PS*, n = 46 teams for *Team Referent PS*; PS = Psychological Safety.

2.2.6 Study 1 Confirmatory Factor Analysis

CFAs were conducted to assess the construct validity of the IR psychological safety and TR psychological safety measures. The CFAs were conducted using *MPlus 8.6* (Muthén & Muthén, 2012, 2021) using the maximum-likelihood estimator. For the IR psychological safety measure, all seven individual-referent items were specified to load onto a single factor. CFA results indicated that this model did not provide an adequate fit to the data, $\chi^2(14) = 46.09$, comparative fit index (CFI) = .89, Tucker-Lewis fit index (TLI) = .83, root mean square error of approximation (RMSEA) = .12, and standardized root mean square residual (SRMR) = .06. An exploratory approach was used to modify this model after examining modification indices (MIs) to assess whether the fit of the model may be improved (Byrne, 2012). MI values indicate the amount by which the χ^2 value would decrease if a cited parameter were to be specified and freely estimated (Byrne, 2012). Based on the MIs, the residual variances between items 5 and 6 were allowed to correlate. It is important to note that correlating residuals between items can increase the risk of overinflating the model fit indices. As Byrne (2012) notes, modifications involving residual correlations are acceptable when items share similar content and theoretical underpinnings. Item 5 (“It is difficult to ask other members of this team for help.”) and item 6 (“Members of this team would not deliberately act in a way that undermines my efforts.”) both tap into the concept of trust in team members, which is a fundamental aspect of team dynamics (Mayer et al., 1995; Rousseau et al., 1998). Trust refers to a psychological state in which one is

willing to be vulnerable based on the positive expectations of another party (e.g., Mayer et al., 1995; Rousseau et al., 1998). Specifically, item 5 assesses the difficulty of asking other team members for help, which requires vulnerability and trust in others' willingness to assist. Item 6 evaluates the expectation of team members' supportive behavior, which influences one's willingness to be vulnerable and trust others. In the modified model where items 5 and 6 were allowed to correlate, CFA results indicated a good fit, $\chi^2(13) = 21.691$, TLI = .95. CFI = .97, RMSEA = .06, and SRMR = .04. Modeling these MIs improved the fit of the model significantly, $\chi^2(1)_{\text{difftest}} = 24.40$, $p = .017$). Table 3 contains the standardized factor loadings, residual variances, and *AIC* estimates of each IR psychological safety item.

Table 3. *CFA for IR Psychological Safety with Selected MIs Modeled (Study 1)*

Item	λ *(SE)*	σ^2	*AIC (SE)*
1. If you make a mistake on this team, it is often held against you.	.57 (.06)	.68 (.07)	.32 (.07)
2. In this team, I am able to bring up problems and tough issues.	.66 (.05)	.57 (.07)	.43 (.07)
3. People on this team sometimes reject me for being different.	.61 (.06)	.63 (.07)	.38 (.07)
4. It is safe to take a risk on this team.	.64 (.06)	.59 (.07)	.41 (.07)
5. It is difficult to ask other members of this team for help.	.59 (.06)	.66 (.07)	.34 (.07)
6. No one on this team would deliberately act in a way that undermines my efforts.	.44 (.07)	.81 (.06)	.19 (.06)
7. Working with members of this team, my unique skills and talents are valued and utilized.	.68 (.05)	.54 (.07)	.46 (.07)

Note. $n = 174$. Estimates are from the completely standardized solution. λ *= factor loading;* σ^2 *= residual variance; AIC* = variance accounted for (by factor); *SE* = standard error. All estimates are $p < .01$.

For the TR psychological safety measure, all seven team-referent items were specified to load onto a single factor. CFA results indicated that a one-factor model did not provide an adequate fit to the data, $\chi^2(14) = 42.177$, TLI = .83. CFI = .89, RMSEA = .10, and SRMR = .06. Given that the ICCs indicated significant within-group (ICC1) and between-group (ICC2) variability in TR psychological safety, I conducted a multilevel confirmatory factor analysis (MCFA) to account for the nested structure of the data and estimate the variance components, with results presented in the subsequent subsection. By accounting for the nesting of individuals

within groups, MLCFA can lead to more accurate estimates of model parameters and improved model fit.

2.2.7 Study 1 Multilevel Confirmatory Factor Analysis

2.2.7.1 TR Psychological Safety MLCFA

A MLCFA was conducted on the TR psychological safety items using *MPlus 8.6* (Muthén & Muthén, 2012, 2021) using the maximum-likelihood estimator to account for the nested structure of the data (i.e., individuals within teams). All seven team-referent items were specified to load onto a single factor at both the within and between levels. The results indicated that this initial model did not provide an adequate fit to the data, $\chi^2(28) = 73.06$, CFI = .84, TLI= .75, RMSEA = .09, standardized root mean square residual at the within level (SRMR-W) = .06, standardized root mean square residual at the between level (SRMR-B) = .30. Based on the MIs, the residual variances between items 1, 3, and 5 were allowed to correlate. The three items are as follows: Item 1 (*"If team members make a mistake on this team, it is often held against them."*), item 3 ("*Members of this team sometimes reject others for being different.*"), and item 5 ("*It is difficult for team members to ask other members of this team for help"*). These items all tap into negative team dynamics (i.e., being blamed, being rejected for being different, and difficulty ask for help).

In the modified model where items 1, 3, and 5 were allowed to correlate, CFA results indicated a good fit, $\chi^2(25) = 32.86$, TLI = .95. CFI = .97, RMSEA = .04, SRMR-W = .04, SRMR-B = .27. Modeling these MIs improved the fit of the model significantly, $\chi^2(3)_{\text{difftest}} = 40.20$, $p = .004$). Table 2 contains the standardized factor loadings, residual variances, and *AIC* estimates of each IR psychological safety item.

Although the SRMR-B value of .27 exceeds the commonly cited cutoff of .08, Ene (2020) recommends interpreting SRMR-B values with caution. The results of a simulation study of model fit indices examined in multilevel confirmatory factor analysis (MLCFA) settings revealed SRMR-B values often surpass the traditional .08 threshold that is typically used in a single-level CFA framework, especially under certain conditions (i.e., higher level ICCs, smaller sample sizes and group sizes) (Ene, 2020). Moreover, all other model fit indices indicated acceptable model fit. Thus, it is reasonable to interpret the results with caution but not discount

them entirely based on the SRMR-B value. Table 4 contains the standardized factor loadings, residual variances, and *AIC* estimates of each TR psychological safety item at the within and between level.

Table 4. *MLCFA for TR Psychological Safety with Selected MIs Modeled (Study 1)*

	Within Level			*Between Level*		
Item	*λ (SE)*	σ^2	*AIC (SE)*	*λ (SE)*	σ^2	*AIC (SE)*
1. If team members make a mistake on this team, it is often held against them.	.24** (.10)	.94** (.05)	.06 (.05)	.65* (.32)	.58 (.42)	.42 (.42)
2. Members of this team are able to bring up problems and tough issues.	.60** (.09)	.64** (.10)	.36* (.10)	.93* (.45)	.14 (.82)	.86 (.82)
3. Members of this team sometimes reject others for being different.	.41** (.10)	.83** (.09)	.17* (.09)	.99** (.18)	.01 (.36)	.99** (.36)
4. It is safe for team members to take a risk on this team.	.74** (.08)	.45** (.45)	.55** (.13)	.71 (.55)	.50 (.77)	.50 (.77)
5. It is difficult for team members of this team for help.	.37** (.09)	.86** (.07)	.14* (.07)	.99** (.11)	.03 (.21)	.97** (.21)
6. Members of this team would deliberately act in a way that undermines other team members' efforts.	.46** (.09)	.79** (.08)	.22** (.08)	.93 (2.70)	.14 (5.02)	.86 (5.02)
7. Working with members of this team, team members' unique skills and talents are valued and utilized.	.51** (.10)	.75** (.10)	.26**(.10)	.99** (.27)	.01 (.54)	.99 (.54)

2.2.8 Study 1 Descriptive Statistics

The means, standard deviations, Cronbach's alphas, and zero-order correlations for the study variables at the individual level are presented in Table 5. Table 6 contains the means, standard deviations and intercorrelations for the aggregated (team-level) study variables.

Table 5. *Means, Standard Deviations, and Correlations for Individual-Level Variables*

Variable	M	SD	1.	2.	3.
1. CSE	3.61	.56	(.85)		
2. IR Psychological Safety	3.93	.60	.33*	(.78)	
3. TR Psychological Safety	3.85	.60	.30*	*a.*	(.75)

Note. N = 364 for CSE, *n* = 174 for IR psychological safety, *n* = 190 for TR psychological safety; IR = individual referent; TR = team referent*;* The response scales for measures were on a 7-point scale. *a.* No correlation was computed between IR psychological safety and TR psychological safety because participants filled out only one of the two measures. **p < .01.* Cronbach's alpha reliabilities are in the parentheses along the diagonals.

Table 6. *Means, Standard Deviations, and Correlations for Aggregated Team-Level Variables*

Variable	M	SD	1	2
1. Mean CSE	3.61	.29		
2. Mean IR Psychological Safety	3.93	.31	.46*	
3. Mean TR Psychological Safety	3.86	.37	.24	*b.*

Note. n = 94 teams for *CSE*, *n* = 43 teams for *Individual Referent PS*, *n* = 46 teams for *Team Referent PS,* PS = Psychological Safety. *b.* No correlation was computed between Individual Referent PS and Team Referent PS because participants filled out only one of the two measures. **p < .01*

2.2.9 Relations Between CSE and IR and TR Psychological Safety

Next, I examined the correlations between team CSE and TR psychological safety. At the team level, the correlation between average team CSE scores and average team scores on the team referent measure of psychological safety was examined and found to be non-significant. Thus, the hypothesis that team average CSE would positively predict team psychological safety was not supported. However, the correlation between CSE scores and TR psychological safety scores was significant at the individual level of analysis.

Given that the individual referent measure did not appear viable as a measure of psychological safety as conceptualized by a direct consensus model, I decided to examine the correlation between CSE and psychological safety at the individual level using the IR psychological safety measure. Individuals' CSE scores were significantly related to their scores on the individual referent psychological safety measure, $r = .34$, $p < .01$ (Table 4). I also examined the correlation between average CSE and individual referent psychological safety at the team level and the correlation was $r = .53$, $p < .01$ (Table 5).

Given that CSE was related to both IR psychological safety and TR psychological safety at the individual level, I examined whether the two correlations significantly differed. A Fisher's z test showed that the correlations did not differ significantly ($z = .32$, $p = .75$). Implications of these findings are described in the Discussion.

2.2.10 Examining the Relation Between CSE and TR Psychological Safety using Multilevel Linear Modeling (MLM)

Given that TRPS demonstrated adequate reliability at both the individual and team levels (as evidenced by sufficient ICC values), I examined the relationship between CSE and TRPS using multilevel modeling. The multilevel modeling results yielded a similar pattern to the correlational analysis. Specifically, the results revealed a significant positive relationship between CSE and TRPS at the within-level ($b = 0.340$, $SE = 0.088$, $p < 0.001$), indicating that individuals with higher CSE tend to perceive higher TRPS. However, the between-level results did not show a significant relationship between CSE and TRPS ($b = -0.834$, $SE = 10.284$, $p = 0.935$).

2.3 Study 1 Discussion

To my knowledge, this study represents the first empirical examination of the validity of using individual and team referents in the measurement of team psychological safety. Moreover, the role of CSE in contributing to team psychological safety was investigated.

ICC indices demonstrated adequate team-member agreement and between-group variance on the TR psychological safety measure, suggesting that team psychological safety may be appropriately operationalized using the referent-shift consensus model. However, regarding the direct consensus operationalization of psychological safety, measured using an individual referent, the ICC indices yielded were low suggesting that team-member agreement and between-group variance *did not* justify aggregation of individual scores to the team level. This finding suggests that the direct consensus model may not be a viable operationalization of team psychological safety. Based on this, it seems reasonable to recommend that future studies use a team referent approach when intending to measure team psychological safety as a belief that is shared among individual members of a team.

Using the TR psychological safety measure, the relation between average team members' scores on CSE and average TR psychological safety was non-significant. However, the positive relation between CSE scores and TR psychological safety scores was significant at the individual level of analysis. Thus, it appears that individuals with higher CSE are likely to perceive higher TR psychological safety. When using the individual referent, a positive correlation was found between CSE and IR psychological safety at both the individual and group level of analysis. Thus, both individuals and teams with higher CSE are likely to report higher IR psychological safety.

It is important to note that TR and IR psychological safety items should not be used interchangeably in one scale. While both constructs share some similarities, they appear to have distinct meanings. The results of this study highlight the importance of maintaining this distinction, as the relationship between CSE and TRPS differed significantly from the relationship between CSE and IRPS at the team level. Thus, using TR and IR items interchangeably in one scale can mask important differences between team-level and individual-level psychological safety, and lead to inaccurate conclusions.

Furthermore, the finding that CSE only related to TRPS at the individual level suggests that CSE may primarily influence individual team members' perceptions of psychological safety, rather than shaping the team's shared climate of psychological safety. By maintaining the distinction between TR and IR items, researchers can gain a more nuanced understanding of the construct of psychological safety.

The results of this study also provide evidence that calls into question the conceptualization of IR psychological safety as a shared construct. Moreover, IR psychological safety appears to be a distinct construct from TR psychological safety given that both exhibit differential relations with CSE at the group level of analysis as mentioned. The construct of IR psychological safety may be better conceptualized using an alternative composition model -- the additive model -- which does not require within-group agreement (Chan, 1998). In additive composition models, the meaning of a group construct at the group level is simply a summation of the lower individual-level units regardless of the variance among units (Chan, 1998). This conceptualization is plausible as scores on the individual referent measure reflect the level of psychological safety that individuals in the team perceive and does not require consensus among team members.

2.3.1 Study Limitations

A limitation of the current study was that a cross-sectional design was used. Thus, causation with respect to the relation between CSE and IR psychological safety and the directionality of this relation cannot be established. Future research may address this limitation by employing longitudinal designs examining the relation between CSE and psychological safety. A second limitation was that the sample used in the study was composed of undergraduate students. Thus, examining the relation between CSE and psychological safety in other populations to ensure that these findings are generalizable may be worthy of future investigation.

2.3.2 Implications and Future Research Directions

In this study, it was identified that the items in the commonly used measure of psychological safety (Edmondson, 1999) are inconsistent in terms of their target referent. Among the samples included in the meta-analysis conducted by Frazier et al. (2017), the vast majority of

findings were based on Edmondson's (1999) measure. Frazier et al. (2017) examined the relations between psychological safety and variables studied in previous research at both the individual level and team level. Frazier et al.'s (2017) findings showed that the effect sizes of the antecedents and outcomes of psychological safety were comparable across individual and group levels of analysis. Based on these findings, the authors suggested that both individual-level and team-level conceptualizations of psychological safety represent complementary views of the same construct. Specifically, they argued that Kahn's (1990) definition of psychological safety as an *individual perception* is a complementary view to Edmondson's (1999) definition of psychological safety as a shared *group perception*. However, given that the present findings suggest using a mixed-referent version of the measure is problematic, this calls into question the appropriateness of the current operationalization of the construct in previous research. In research where psychological safety is studied at the group level of analysis, statistics that justify its aggregation (e.g., ICCs) are typically included. However, the composition model is rarely specified. Liang et al. (2012) conceptualize psychological safety as an individual perception to justify their use of an individual referent using an adapted measure of psychological safety and analyzing psychological safety at the individual level. In a paper examining psychological safety as a group-level construct, Ortega et al. (2010) refer to Edmondson's original measure as a referent-shift consensus measure which is not accurate given that some of the items include an individual referent.

The results of Study 1 suggest that a referent-shift consensus model in which a team referent approach is used in questionnaire items is appropriate for operationalizing team psychological safety according to Edmondson's (1999) conceptualization as a shared perception of teams whereas an additive model may be more appropriate for operationalizing perceptions of individual psychological safety. An individual referent measure of psychological safety may be more aligned with Kahn's (1990) definition of psychological safety. A next step for future research would be to establish the predictive validity of team psychological safety on important team outcomes using a referent-shift consensus operationalization. The predictive validity of IR and TR psychological safety was examined in Study 2. Study 2 was also conducted to determine whether Study 1 results would be replicated.

Chapter 3

3 Study 2

In Study 2, I examined the role of CSE and core group-evaluations (CGE) on IR and TR psychological safety. I also examined the relations between psychological safety and team-related behaviours (i.e., task performance, citizenship behaviour, and counterproductive behaviour) and performance. Study 2 had five key objectives.

The first objective was to replicate the findings of Study 1 demonstrating the appropriateness of using a referent-shift consensus model for operationalizing TR psychological safety and inappropriateness of using a direct consensus model for operationalizing IR psychological safety. The second objective was to examine the relation between CSE and CGE and IR and TR psychological safety. The third objective was to examine and compare the incremental validity of IR and TR psychological safety in terms of the variance accounted for in team-related behavioural outcomes (in-role behaviour, citizenship behaviour, counterproductive behaviour, perceived team performance) and team effectiveness. The fourth objective was to examine the role of IR psychological safety in mediating the relation between CSE and team-related outcomes and fifth objective was to examine the role of TR psychological safety in mediating the relation between CGE and team-related outcomes. In the following sections, I describe the hypotheses proposed in Study 2 and the rationale that relate to each of the five aforementioned objectives.

3.1 Study 2 Hypotheses and Rationale

3.1.1 Aggregation-based Operationalization of TR Psychological Safety

The findings of Study 1 demonstrated support for the validity of a referent-shift consensus model of TR psychological safety, as indicated by the ICC(1) value [ICC(1) = .19]. In contrast, however, ICC(1) for a direct consensus conceptualization of psychological safety (i.e., IR psychological safety) was quite low [ICC(1) = .04], suggesting that psychological safety in which an individual referent is used may not be considered as a shared construct and may not be characterized by the homogeneity of responses within teams. Thus, I predict the following:

H1a: ICC(1) for TR psychological safety will be sufficiently high to support the validity of a referent-shift consensus operationalization.

H1b: ICC(1) for IR psychological safety will not be sufficiently high to support the validity of a direct consensus operationalization.

3.1.2 Factor Structure of IR and TR Psychological safety

In Study 1, a model involving both IR and TR psychological safety was not evaluated because a subset of the sample filled out only one of two versions (i.e., either the IR or TR version) of the psychological safety measure. In Study 2, all participants filled out both measures, enabling an examination of a model that included both IR and TR psychological safety. Given that the results of Study 1 appeared to suggest that both IR psychological safety and TR psychological safety operate as distinct constructs and are best conceptualized using different composition models (Chan, 1998), the following hypothesis is forwarded:

H2: Scores on the IR psychological safety items and TR psychological safety items will load on two separate factors.

3.1.3 CSE as an Antecedent of IR and TR Psychological Safety

Study 1 showed that CSE related to IR psychological safety (at both the individual and team level) but that CSE related to TR psychological safety *only* at the individual level. This suggests that higher CSE in teams predict psychological safety when it is conceptualized as the perception *individuals* in the team have regarding psychological safety. This conceptualization does not require consensus among team members and follows an additive model (Chan, 1998). CSE may be more likely to influence IR psychological safety regardless of the level of analysis because -- even at the group level -- both CSE and IR psychological safety appeared to reflect a characteristic of individual members within teams. Study 1 also demonstrated that CSE was positively related to TR psychological safety but only at the individual level of analysis. Given the findings of Study 1, I put forth the following hypotheses:

H3: At the team level of analysis, CSE will be a) positively related to IR psychological safety and b) unrelated to TR psychological safety.

H4: At the individual level of analysis, CSE will be positively related to a) IR psychological safety and b) TR psychological safety.

3.1.4 CGEs as Antecedents of IR and TR Psychological Safety

Core group evaluations (CGE) have been defined by Schmidt, Ogunfowora and Bourdage (2012) as representing a "higher order, group-level construct that represents cognitive judgments that a group or unit makes about its own capacity, ability, and overall worth" (p. 929-930). CGE encompass three constructs: collective efficacy, collective esteem, and collective locus of control. Content related to neuroticism was excluded by previous researchers because it was not deemed to be applicable at the group level (Schmidt et al., 2012).

Collective efficacy refers to "a group's shared belief in their conjoint capabilities to organize and execute the courses of action required to produce given levels of attainments" (Bandura, 1997, p. 476). Collective esteem is the overall shared belief held by group members that the "group is valued within the larger system or organization within which it is nested" (Schmidt et al., 2012, p. 930). Collective locus of control refers to "the shared belief that the group is able to control changing environmental factors to achieve desired outcomes" (Schmidt et al., 2012, p. 930). A higher team climate of psychological safety may be more likely to emerge in teams with high CGE. Considering that high CGE involves the perception that the team is capable of effectively working together, high CGE may be likely to engender the perception that the team can also successfully take interpersonal risks. As a result, team members may feel safer engaging in interpersonal risk-taking behaviours. In addition, team members may be more prone to taking interpersonal risks given their belief that the team's actions can control its outcomes. As a result, the team may develop consistent patterns of interpersonal risk-taking behaviour (i.e., group norms) that promote the emergence of high psychological safety.

In the current study, the predictive ability of CGE on both IR psychological safety and TR psychological safety was examined. Given that IR psychological safety appears to capture individual perceptions of each team members' psychological safety rather than perceptions of the entire group, it may be less likely to be influenced by CGE than TR psychological safety.

Thus, my hypotheses regarding the prediction of CGE on IR and TR psychological safety are as follows:

H5: At the team level of analysis, CGE will be a) positively related to TR psychological safety but b) unrelated to IR psychological safety.

H6: At the individual level of analysis, CGE will be a) positively related to TR psychological safety but b) unrelated to IR psychological safety.

3.1.5 IR and TR Psychological Safety and Team Member Behaviours

Task performance

Psychological safety has also been found to relate to in-role behaviours or task performance at the group level. In Frazier et al.'s (2017) meta-analysis, involving 136 samples, a positive relationship between psychological safety and task performance ($\hat{\rho}$ = .29) was found at the group level of analysis. In the present study, task performance was also assessed within the team context and at the team level of analysis. Each team member rated each of the other team members in terms of their task performance. A score for was computed for each individual team member. Each team's task performance score was operationalized as the average of individual team members' scores on task performance within the team.

Given that individuals in a psychologically safe team are less likely to experience the negative consequences of taking interpersonal risks (such as admitting to making mistakes or proposing new ideas), psychological safety is likely to enable team members to focus their attention on performing their tasks (Frazier et al., 2017). This is likely to lead to higher task performance at both the individual, and group, levels of analysis. Moreover, the meta-analytic research conducted by Frazier et al. (2017) also demonstrated that psychological safety generally has a similar impact on task performance at both the individual and group level of analysis. (Frazier et al., 2017).

Thus, I predict the following:

H7: At the team level of analysis, both a) IR psychological safety and b) TR psychological safety will be positively related to aggregated peer ratings of task performance.

H8: At the individual level of analysis, both a) IR psychological safety and b) TR psychological safety will be positively related to peer ratings of task performance.

Citizenship Behaviours

The meta-analysis conducted by Frazier et al. (2017) also demonstrated a positive relationship between psychological safety and citizenship behaviours at the *individual* level. The authors were not able to examine the meta-analytic relationship between psychological safety and citizenship behaviours at the group level due to a limited number of studies. Although the research examining citizenship behaviours as a group-level outcome of psychological safety is limited, I theorize that psychological safety would facilitate citizenship behaviour at the team level as well.

Citizenship behaviours are discretionary but important for aiding in the functioning of groups or organizations (Organ, 1988). Citizenship behaviours may involve taking initiative, challenging the status quo such as speaking up against a popular opinion or challenging a supervisor's decision, and suggesting improvements. While not all citizenship behaviours may involve interpersonal risk, a team member engaging in these behaviours may potentially face rejection or criticism from others who perceive their actions as disruptive, disagree with their ideas or approaches, resist changes due to a fear of change or uncertainty, or feel threatened by being challenged due to a fear or loss of power or control. In a psychologically safe team environment, team members can focus on engaging in behaviours that improve the functioning of the group without fear of reprisal or judgment. Thus, psychological safety may alleviate the risks associated with citizenship behaviours at both the individual and team level, encouraging members to go above and beyond their formal roles to support the team's success.

Moreover, psychological safety may foster a positive experience for team members, which is likely to enhance their motivation to improve the team and its outcomes. As the Social Exchange Theory suggests, individuals are more likely to invest time, energy, and resources in

relationships that provide them with benefits and support (Blau, 1964; Emerson, 1976). When team members feel psychologically safe, they are likely to perceive their interactions as positive and beneficial, leading to increased motivation to contribute to the team's success. Thus, psychological safety not only alleviates the risks that may be associated with citizenship behaviours but also may encourage team members to go above and beyond their formal roles as a form of reciprocity, ultimately driving team performance.

Thus, I propose the following hypotheses:

H9: At the team level of analysis, both a) IR psychological safety and b) TR psychological safety will be positively related to aggregated peer ratings of citizenship behaviour.

H10: At the individual level of analysis, both a) IR psychological safety and b) TR psychological safety will be positively related to peer ratings of citizenship behaviour.

Counterproductive Behaviour

In Study 2, I also examined the relationship between psychological safety and counterproductive behaviour in teams. Applying social exchange theory (Blau, 1964; Emerson, 1976), psychologically safe team climates are likely to engender positive responses from team members to the positive social interactions that are likely to characterize the team. When individual members of a team perceive low psychological safety, they may feel undermined and as a result retaliate with a negative response and engage in counterproductive behaviours such as social loafing and disparaging team members. If team members as a whole perceive the climate as unsafe, this is also likely to drive higher levels of counterproductive behaviour at the team level of analysis. Thus, I predict the following:

H11: At the team level of analysis, both a) IR and b) TR psychological safety will be negatively related to aggregated peer ratings of counterproductive behaviour.

H12: At the individual level of analysis, both a) IR and b) TR psychological safety will be negatively related to peer ratings of counterproductive behaviour.

3.1.6 IR and TR Psychological Safety and Team Performance

Team psychological safety has been shown to be linked empirically to higher team performance (e.g., Edmondson, 1999; Kostopoulos & Bozionelos, 2011; Ortega et al., 2010; Schaubroeck et al., 2011) and there are numerous models and theories that have been applied to explain this link. Edmondson (1999) theorized that psychological safety enhances team members' willingness to engage in behaviours that facilitate learning such as seeking feedback, information sharing, asking for help, bringing up errors, and experimentation. In a psychologically safe climate, team members are less likely to believe that they will suffer negative consequences such as being perceived as incompetent or receiving backlash from engaging in learning behaviour (Edmondson, 1999). Engaging in learning behaviours, in turn, is likely to facilitate higher team performance by enhancing team members' shared understanding of their work requirements and ability to resolve issues and solve problems. In Edmondson's (1999) seminal study using an organizational sample of manufacturing teams, team psychological safety was related to higher team learning behaviour, which in turn related to higher team performance.

Research by Schaubroeck et al. (2011) has also shown a direct relation between psychological safety and team performance. Servant leadership was found to influence team performance through cognition-based trust (i.e., trust in others on performance-related dimensions such as competence and reliability) and psychological safety. Moreover, research suggests that psychological safety can facilitate creative performance in teams by providing a supportive context in which divergent perspectives are handled respectfully and alternative ideas are explored (Hu et al., 2018). Hu et al. (2018) found that team psychological safety predicted higher team creativity among work teams in the information and technology industry.

Social learning theory has also been utilized to explain the positive effects of psychological safety on performance. This theory (Bandura, 1977) suggests that people learn from observing others' behaviours and considering the consequences of those behaviours. A psychologically safe team climate is likely to be characterized by patterns of behavioural interactions that have been reinforced by observation (Newman et al., 2017). Learned behaviours such as information sharing, voice behaviour, help-seeking, and experimentation in turn are likely to facilitate outcomes such as higher performance (Newman et al., 2017).

Social exchange theory (Blau, 1964), which describes social behaviour as an "exchange process", provides another plausible explanation of the link between psychological safety and positive outcomes such as performance. When individuals have a positive social experience, they respond by reciprocating (Blau, 1964). Researchers have suggested that individual team members may reciprocate positive and supportive behaviours that they experience in a psychologically safe climate with their own positive behaviours, leading to favorable outcomes such as higher performance (Newman et al., 2017).

In a systematic review, Newman and colleagues (2017) applied the Conservation of Resources (COR) theory (Hobfoll, 1989) to explain how psychological safety influences performance. According to COR theory (Hobfoll, 1989), humans are motivated by the goal to gain and maintain resources (i.e., what they value). Moreover, individuals will experience stress and burnout when they perceive a *threat* to their resources, even if there is not an actual loss of resources (Halbesleben & Buckley, 2004). Stress and burnout can lead to negative work outcomes such as lower team performance. In a psychologically unsafe team environment, team members may be less likely to take interpersonal risks in fear of losing resources. For example, they may believe that they will be perceived as incompetent if they seek help due to previous experiences of being harshly criticized or reprimanded and worry that this might threaten future opportunities to be rewarded or recognized. On the other hand, if members of a psychologically safe team have become accustomed to receiving social support and valuable information and feedback when engaging in interpersonal risk-taking behaviours such as seeking help, they will be motivated to continue to draw resources from their team members that can help them successfully meet their work demands (Newman et al., 2017).

The vast majority of findings supporting the link between psychological safety and team performance is based on data utilizing Edmondson's (1999) measure which includes a mix of individual, team, and ambiguous referents. It is likely that both IR and TR psychological safety is related to team performance regardless of the level of analysis. Thus, the following hypotheses are proposed:

H13: At the team level of analysis, both a) IR and b) TR psychological safety will be positively related to perceived team performance.

H14: At the individual level of analysis, both a) IR and b) TR psychological safety will be positively related to perceived team performance.

H15: At the team level of analysis[1], both a) IR and b) TR psychological safety will be positively related to graded team effectiveness (i.e., final team project grade).

3.1.7 Incremental Validity

I also aimed to compare the incremental validity of both IR and TR psychological safety at both the individual and team level of analysis. Given that IR psychological safety appears to capture individual perceptions of psychological safety, there is a greater alignment across the level of referent and level of analysis for IR psychological safety at the individual level of analysis. On the other hand, there is a greater alignment between the level of referent and level of analysis for TR psychological safety at the team level of analysis. Thus, I hypothesize stronger predictive ability when there is congruence between the level of analysis and level of referent used. Thus, I put forth the following hypotheses.

H16: At the individual level of analysis, IR psychological safety will predict variance in peer ratings of team members' a) task performance, b) citizenship behaviours, and c) counterproductive behaviours above and beyond TR psychological safety.

H17: At the team level of analysis, TR psychological safety will predict variance in aggregated peer ratings of team members' a) task performance b) citizenship behaviours, c) counterproductive behaviours; d) perceived team performance, and e) graded team effectiveness above and beyond IR psychological safety.

3.1.8 Psychological safety as a mediator

The role of psychological safety as a mediator is also examined. Psychological safety may mediate the relationship between CSE and CGE and team-related behaviours. Past research

[1] Graded team effectiveness was assessed by final project grade which is only meaningful at the team level. Therefore, the relationships between IR and TR psychological safety and graded team effectiveness (final project grade) at the individual level were not examined.

by Haynie (2012) found a significant positive correlation (.32) between mean CSE in teams and team performance. Given the relationship between mean CSE and IR psychological safety found in Study 1 and the established link between psychological safety and team performance (Edmondson, 1999), a potential mediator between mean CSE and team performance may be team psychological safety. In addition, CGE has also been linked to team performance (Schmidt et al., 2012). Thus, I propose the following hypotheses:

> H18: *At the individual level of analysis, IR psychological safety will mediate the relationship between CSE and a) task performance b) citizenship behaviours, and c) counterproductive behaviour.*
>
> *H19: At the team level of analysis, IR psychological safety will mediate the relationship between CSE and a) task performance b) citizenship behaviours, c) counterproductive behaviours d) perceived team performance, and e) graded team effectiveness at the team level.*
>
> *H20: At the individual level of analysis, TR psychological safety will mediate the relationship between CGE and a) task performance b) citizenship behaviours, and c) counterproductive behaviours.*
>
> *H21: At the team level, TR psychological safety will mediate the relationship between CGE and a) task performance b) citizenship behaviours, c) counterproductive behaviours d) perceived team performance, and e) graded team effectiveness.*

The proposed model is shown in Figure 1 below:

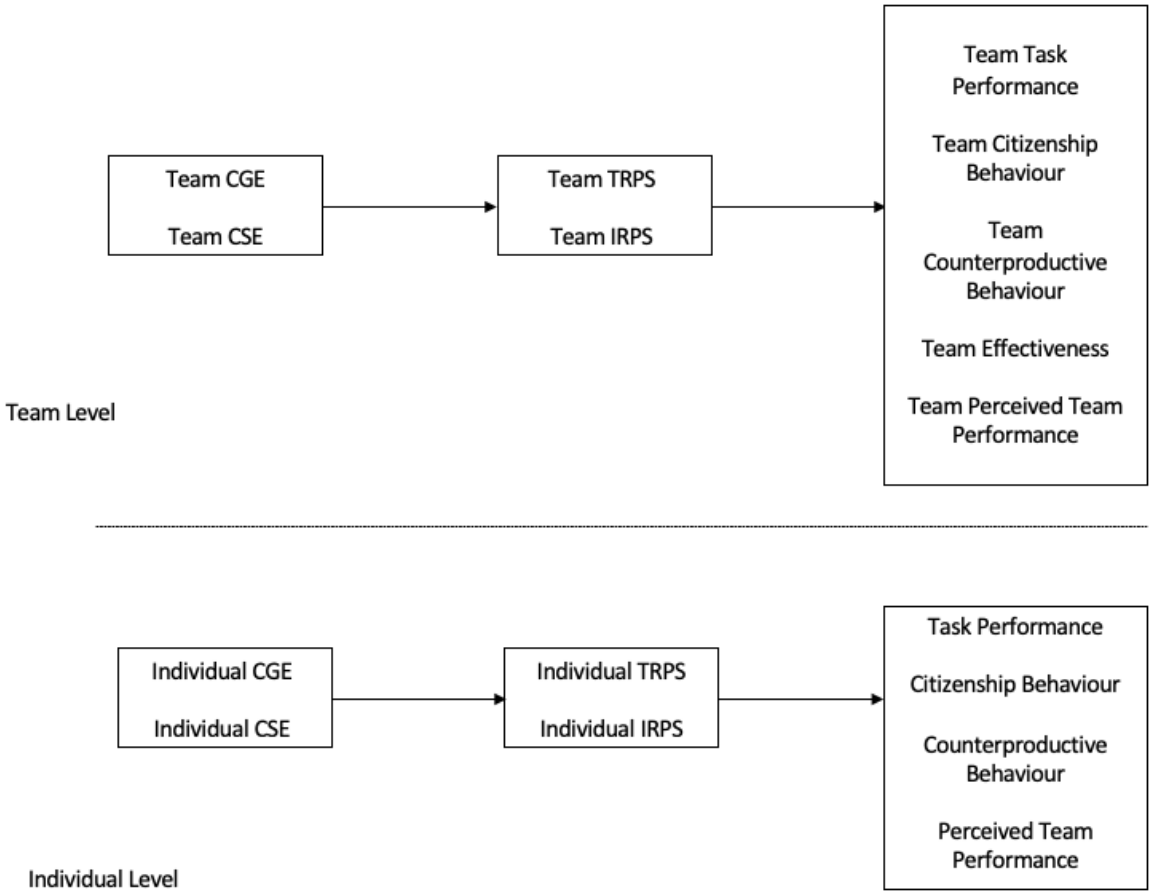

Figure 1. *Study 2 Proposed Model*

3.2 Study 2 Method

3.2.1 Participants and Procedure

Similar to Study 1, the sample consisted of students enrolled in an undergraduate engineering course during the 2013-2014 academic year. During the course, students were required to work in a team to complete two major engineering design projects. Participants provided signed informed consent to acknowledge their willingness to participate in the study.

Teams were formed in the second week of September. Data were collected at three time points as part of a larger data collection. Initially, 432 individuals in 102 teams agreed to participate in the study. Seven individuals who did not complete T2 and T3 questionnaires due to course withdrawal were excluded. Two teams made up of eight individuals with less than three

team members responding at T2 and T3 were also excluded. Five teams of 25 individuals were excluded because there was team membership change between T2 and T3. A flow chart depicting the data exclusion criteria is shown in Figure 1. The final sample consisted of 94 teams composed of 392 individuals. The average group size in the sample was four members.

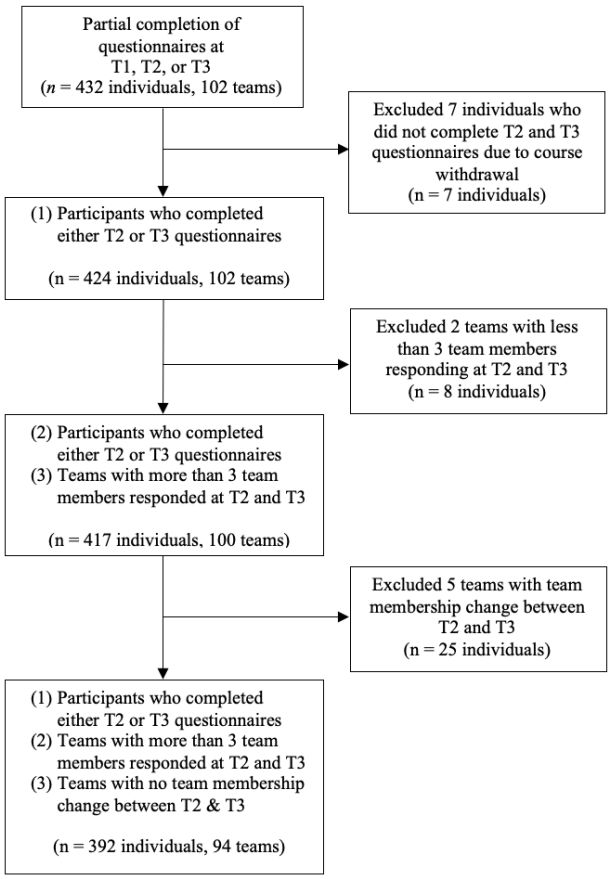

Figure 2. *Study 2 Participant Inclusion and Exclusion*

CSE data was collected in the second week of September when teams were formed. CGE and IR and TR psychological safety assessments were collected after teams had been working together for approximately 10 weeks and had received a grade for their first team project. The first project required that teams submit a written report describing a design solution that would increase the usability of an object. Dependent measures (task performance, citizenship behaviour, and counterproductive behaviour, and perceived team performance) were collected after teams had been working together for approximately 5.5 months. Graded team effectiveness

was the final project grade that was given to each team at the end of the school year which was after the team had been working together approximately 7.5 months. Paper/pencil questionnaires were administered to participants during the weekly laboratory component of the course. Participation in the study was voluntary and students who participated received extra credit towards their final grade in the course.

3.2.2 Study 2 Measures

3.2.2.1 Core Self-evaluations

A 12-item measure developed by Judge et al. (2003) was used to assess CSE. Sample CSE items include, “I am confident I get the success I deserve in life” and “I am filled with doubts about my competence” (reverse-scored). Responses were based on a five-point scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*) scale (α = .82).

3.2.2.2 Core Group Evaluations

Judge et al.’s (2003) measure was adapted to measure CGE. Items were modified to reflect one’s perceptions of the team, rather than the self. For example, items include: “My group completes its assigned tasks successfully” and “When the members of my group try very hard and work together, we generally succeed.” Responses were based on a five-point scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*) scale (α = .79). ICC(1) and ICC(2) were calculated in order to justify aggregating and examining CGE at the team level. The ICC(1) value was .30 and the ICC(2) value was .64.

3.2.2.3 IR and TR Psychological Safety

Similar to Study 1, psychological safety was measured using adapted measures of Edmondson’s (1999) original 7-item scale. Participants completed both the IR and TR measure of psychological safety. One subset of the sample filled out the IR psychological safety measure first and the other subset of the sample filled out the TR psychological safety measure first. Similar to Study 1, each student and their team were assigned to a studio where they attended their laboratory sessions for the engineering course. The order of the psychological safety questionnaires that each team received was randomly assigned to studios.

Table 7 contains the items from the two measures used in the study and the original items from Edmondson (1999). Both the IR and TR psychological safety measures differed slightly between Study 1 and Study 2. Given that Study 1 demonstrated some evidence that individual-referent and team-referent items are not interchangeable and may have distinct differences, it was important to enhance clarity in the items by reducing ambiguity regarding the referents within each measure. The ambiguous items that were used in the IR psychological safety scale (i.e., items 4 and 5) were modified to make the individual referent explicit. After examining the other items carefully for any ambiguity regarding the referent used, item 1 was modified from *"If you make a mistake on this team, it is often held against you"* to *"If I make a mistake on this team, it is often held against me"* to make the referent more explicit to the individual self (rather than a "general" reference to any single team member). In the TR psychological safety measure, item 3 was modified from *"Members of this team sometimes reject others for being different"* to *"Members of this team sometimes reject each other for being different"* to make it clear that the items were referring to team members rather than others external to the team. The Cronbach's alpha for the IR psychological safety scale was .75. The TR psychological safety scale was adapted so that items referred to the team or team members only. Sample items on this measure include: "It is difficult for *team members* to ask other members of the team for help." and "If *team members* make a mistake on this team, it is often held against *them*." The Cronbach's alpha for the TR psychological safety scale was .75. For both the individual and team referent measures, responses were based on a scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*).

Table 7. S*tudy 2 Original Edmondson (1999), Individual-referent and Team-referent Psychological Safety Items*

Original Edmondson (1999) Psychological Safety Scale	Modified Individual-Referent Psychological Safety Scale	Modified Team-Referent Psychological Safety Scale
1. If **you** make a mistake on this team, it is often held against **you**.	**1.** If **I** make a mistake on this team, it is often held against **me.**	1. If **team members** make a mistake on this team, it is often held against **them.**
2. **Members of this team** are able to bring up problems and tough issues.	2. **I** am able to bring up problems and tough issues.	2. **Members of this team** are able to bring up problems and tough issues.
3. People on this team sometimes reject **others** for being different.	3. Members of this team sometimes reject **me** for being different.	3. Members of this team sometimes reject **each other** for being different.
4. It is safe to take a risk on this team.	4. It is safe for **me** to take a risk on this team.	4. It is safe for **team members** to take a risk on this team.
5. It is difficult to ask other members of this team for help.	5. It is difficult for **me** to ask other members of this team for help.	5. It is difficult for **team members** to ask other members of this team for help.
6. No one on this team would deliberately act in a way that undermines **my** efforts.	6. Members of this team would not deliberately act in a way that undermines **my** efforts.	6. Members of this team would not deliberately act in a way that undermines other **team members'** efforts.
7. Working with members of this team, **my** unique skills and talents are valued and utilized.	7. Working with members of this team, **my** unique skills and talents are valued and utilized.	7. Working with **members of this team**, **team members'** unique skills and talents are valued and utilized.

3.2.2.4 Perceived Team Performance Reported by the Team

A 5-item scale, developed by Alper et al (1998), was used to measure perceived team performance. Sample items from the measure include: "The way we manage our work inspires us to better performance" and "All things considered, the team is highly pleased with the way it manages its work." Responses were based on a scale ranging from 1 (*never*) to 7 (*always*) (α = .93). This variable is referred to hereafter as: perceived team performance.

3.2.2.5 Graded Team Effectiveness on a Design Project

The grade on the final group project in the course was used as a measure of team effectiveness. Teams were required to develop an engineering design, construct it, and conduct a presentation describing the design. This variable is referred to hereafter as: graded team effectiveness.

3.2.2.6 Team Members' Task Performance (Peer Ratings)

Peer ratings of each team member's task performance were collected. Team members' task performance was measured using three items adapted from van Dyne and LePine (1998). Sample items on this measure include: "To what extent does (*group member's first name*) successfully perform assigned tasks?" and "To what extent does (*group member's first name*) produce quality work that meets performance expectations?" Responses were based on a seven-point Likert scale ranging from 1 (*never*) to 7 (*always*) (α = .90).

3.2.2.7 Team Members' Citizenship Behaviour (Peer Ratings)

Peer ratings of citizenship behaviour were also collected for each team member. Citizenship behaviour was measured using three items adapted from Lee and Allen's (2002) scale**.** A sample items includes, "To what extent does (*group member's first name*) assist others with their duties?" Responses were based on a seven-point Likert scale ranging from 1 (*never*) to 7 (*always*) (α = .91).

3.2.2.8 Team Members' Counterproductive Behaviour (Peer Ratings)

Four items from Woodley (2017) were used to measure counterproductive behaviour in a team setting. Sample items included, "To what extent does (*group member's first name*) distract team members during team meetings?" and "To what extent does (*group member's first name*) engage in activities that derail the team's progress on the project?" Responses were based on a scale ranging from 1 (*never*) to 7 (*always*) (α = .85).

3.3 Study 2 Results

3.3.1 Study 2 Participant Characteristics

The average age of participants was 19 years, ranging from 16 to 36 years. The majority of the sample was male (79.3%) and Caucasian (59.1%). 16.5% of participants identified as East Asian, 11.5% identified as Arabic or East Indian, 2.5% identified as Southeast Asian, .5% identified as Black, and .5% identified as Native American. The remaining 9.4% identified as Other.

3.3.2 Test of Questionnaire Order Effect

To determine whether questionnaire order affected scores on the IR psychological safety measure and TR psychological safety measure, independent samples *t*-tests were conducted on the scores of both measures with questionnaire version as the grouping variable. There was no significant difference in scores on IR psychological safety and TR psychological safety between the two questionnaire versions, as depicted in Table 8.

Table 8. t-tests Comparing Questionnaire Versions on IR and TR Psychological Safety (Study 2)

Variable	Questionnaire Version 1 (IRPS First)		Questionnaire Version 2 (TRPS First)		*t*(380)	*p*
	M	SD	M	SD		
IR Psychological Safety	3.98	.49	4.03	.57	-.91	.36
TR Psychological Safety	4.02	.54	4.01	.54	.09	.93

I also examined if the correlation between IR and TR psychological safety differed significantly between the subsample that received IR psychological safety was completed first and the subsample that completed TR psychological safety first. In the sample that completed Questionnaire Version 1 (n = 198) (where IR psychological safety was completed first), the correlation between IR and TR psychological safety was .79, $p < .001$. In the sample that completed Questionnaire Version 2 (n = 169) (where TR psychological safety was completed first), the correlation between IR and TR psychological safety was .86, $p < .001$. A Fisher's Z transformation test demonstrated a significant difference the two correlations (p = .035).

I also compared the subsample correlations to the correlation found in the total sample to determine the severity of the order effect. In the total sample (n = 368), the correlation between IR and TR psychological safety was .81, $p < .001$. There was no significant difference between the correlation between IR and TR psychological safety for Questionnaire Version 1 from the correlation observed in the total sample. Moreover, there was no significant difference between the correlation between IR and TR psychological safety for Questionnaire Version 2 from the correlation observed in the total sample.

One possible explanation for the difference in correlations between IR psychological safety and TR psychological safety based on which items were completed first may have to do with a frame-of-reference effect (Schwarz, 1999). This phenomenon suggests that the context in which participants respond influences their responses to subsequent scales (Schwartz, 1999). It may be possible that when participants completed the TR psychological safety measure first, they recalled individual-related information regarding psychological safety (which may be more salient and easier to access), resulting in an assimilation effect. On the other hand, when IR psychological safety items were completed first, it may have been easier for participants to make a distinction between IR and TR psychological safety. Considering this order effect, the results of the study should be interpreted with caution.

3.3.3 Study 2 Descriptive Statistics

The means, standard deviations, Cronbach's alphas, and zero-order correlations for the study variables at the individual level are presented in Table 9. Table 10 contains the means, standard deviations and intercorrelations for the aggregated (team-level) study variables.

Table 9. *Means, Standard Deviations, and Correlations for Individual-Level Variables (Study 2)*

Construct	M	SD	1.	2.	3.	4.	5.	6.
1. CSE	3.67	.52						
2. CGE	3.70	.47	.11*					
3. IR Psychological Safety	4.00	.52	.27**	.50**				
4. TR Psychological Safety	4.02	.54	.20**	.53**	.82**			
5. Task Performance	5.73	1.15	.19*	.13*	.33*	.30**		
6. Citizenship behaviours	4.99	1.33	.16*	.21*	.37**	.33**	.86**	
7. Counterproductive Work Behaviours	1.87	.82	-.02	-.10	-.11*	-.16**	-.47**	-.42**

Note. M = mean, SD = standard deviation, IR = Individual-referent, TR = Team Referent. * $p < .05$. ** $p < .01$. *n* ranged from 356 to 381.

Table 10. *Means, Standard Deviations, and Correlations for Team-Level Variables (Study 2)*

Construct	M	SD	1.	2.	3.	4.	5.	6.	7.	8.
1. Mean CSE	3.67	.27								
2. Mean CGE	3.69	.34	.15							
3. Mean IR Psychological Safety	4.00	.31	.22*	.69**						
4. Mean TR Psychological Safety	4.01	.34	.08	.70**	.88**					
5. Mean Task Performance	5.73	.60	-.04	.53**	.43**	.55**				
6. Mean Citizenship behaviours	5.00	.88	-.02	.53**	.50**	.57**	.83**			
7. Mean Counterproductive Work Behaviours	1.86	.63	-.05	-.28**	-.27**	-.39**	-.38**	-.28**		
8. Mean Perceived Team Performance	5.39	.68	-.02	.49**	.45**	.47**	.64**	.70**	-.35**	
9. Mean Team Effectiveness Grade	25.12	2.44	-.04	.01	.07	.13	.10	.14	.04	.13

Note. M = mean, SD = standard deviation, IR = Individual-referent, TR = Team-referent. * $p < .05$. ** $p < .01$. $n = 94$ teams.

3.3.4 ICCs for IR and TR Psychological Safety

ICCs were calculated to test Hypothesis 1a and 1b. Similar to Study 1, the ICC(1) cutoff value of .12 was used to indicate a sufficient value for justifying aggregating individual scores to the team level. Table 14 contains the computed ICCs for the aggregated measures of psychological safety. H1a was supported as the ICC(1) value yielded for the TR psychological safety measure was sufficiently high, indicating that it is acceptable to aggregate individual responses on the team referent measure to the team level [ICC(1) = .14]. H1b was also supported as the ICC(1) value was below the cutoff [ICC(1) = .11]. Notably, the difference in ICCs between the two measures was not as pronounced as in Study 1. Nevertheless, the findings confirm that the referent-shift model is the appropriate conceptualization for capturing psychological safety as a shared team construct, as the team referent measure yielded a sufficiently high ICC(1) value, while the individual referent measure did not.

Table 11. *Intraclass Correlations for IR and TR Psychological Safety (Study 2)*

	ICC(1)	ICC(2)
1. Mean IR Psychological Safety	0.11	0.33
2. Mean TR Psychological Safety	0.14	0.40

Note. n = 45 for *Individual Referent PS*, n = 48 for *Team Referent PS*; PS = Psychological Safety.

3.3.5 Factor Structure of IR and TR Psychological Safety

Hypothesis 2 proposed that IR and TR psychological safety would load on two separate factors. *Mplus 8.6* was used to investigate the validity of this hypothesized two-factor model, and other possible models. Five measurement models were tested. The results of each model are depicted in Table 12. In the first model, all IR psychological safety items and TR psychological safety items were loaded on one factor. For this model, CFA results indicated that a one-factor model provided a poor fit to the data, $\chi^2(77) = 897.62$, CFI = .60, TLI = .53, RMSEA = .17, and SRMR = .10. An examination of the MIs demonstrated that residuals between matching items were correlated (i.e., item 1 in the IR psychological safety and item 1 in the TR psychological safety measures, item 2 in the IR psychological safety and item 2 in the TR psychological safety

measure etc.). Specifically, MIs indicated that allowing the residuals between the matching items would improve the model fit. Given that each item differs in terms of the referent but is similar in item content, the modifications allowing each pair of matching items were modeled. When the residuals for matching items were allowed to correlate, this improved the model fit, $\chi^2(70) = 248.49$, CFI = .91, TLI = .89, RMSEA = .08, and SRMR = .07.

The second model that was tested examined a two-factor model with all IR psychological safety items loading on one factor and all TR psychological safety items loading on another factor. Factors were constrained to a correlation of 0 (uncorrelated). For this model, CFA results indicated that a two-factor model with uncorrelated factors provided a poor fit to the data, $\chi^2(91) = 1304.95$, TLI = .30. CFI = .41, RMSEA = .21, and SRMR = .24. An examination of the modification indices demonstrated that the highest MI involved correlating the two factors which was tested next.

In this next model, the two factors (i.e., IR psychological safety and TR psychological safety) were allowed to correlate. For this model, CFA results indicated that a correlated two-factor model provided a poor fit to the data, $\chi^2(77) = 884.20$, TLI = .53. CFI = .61, RMSEA = .17, and SRMR = .10. An examination of the modification indices demonstrated that residuals between matching items were correlated. When the residuals of all matching items were allowed to correlate, this improved the model fit, $\chi^2(69) = 217.19$, CFI = .93, TLI = .91, RMSEA = .08, and SRMR = .07. A chi square difference test demonstrated that the model fit for the modified two-factor model was significantly improved in comparison to the modified one-factor model, $\chi^2(1)_{\text{difftest}} = 31.30$, $p < .001$). The two-factor modified model was a superior fit to the one-factor modified model and all other models tested, providing support for Hypothesis 2. I also attempted to run MLCFAs on the individual and team referent items to account for the nested structure of the data. However, the two-factor and one-factor MLCFA models did not converge.

Table 12. *CFA Model Fit Statistics for Measurement Models Tested in Study 2*

Model	χ^2	χ^2df	RMSEA	SRMR	CFI	TLI
Single-factor model	897.62	77	.17	.10	.60	.53
Single-factor model - modified	248.49	70	.08	.07	.91	.89
Two-factor model (orthogonal)	1304.95	77	.21	.24	.41	.30
Two-factor model (correlated)	884.20	76	.17	.10	.61	.53
Two-factor model (correlated) - modified	217.19	69	.08	.07	.93	.91

3.3.6 Examining CSE as an Antecedent of IR and TR Psychological Safety Using Correlational Analysis

H3 proposed that CSE would be positively related to IR psychological safety but unrelated to TR psychological safety at the group level. H3 was supported as CSE scores were positively related to IR psychological safety, *(r = .22, p < .001)* but not TR psychological safety, (*r* = .08, *n.s.*) at the group level of analysis. H4 was also supported as CSE scores were positively related to IR psychological safety *(r = .27, p < .001)* and TR psychological safety, (*r* = .20, *p* < *.001*.) at the individual level of analysis. A Fisher's *z* test found that the relationship between CSE and IR psychological safety and CSE and TR psychological safety did not significantly differ at the individual level (*z* = .98, *p* = .33).

3.3.7 Examining CGE as an Antecedent of IR and TR Psychological Safety Using Correlational Analysis

H5 proposed that CGE would be positively related to TR psychological safety but unrelated to IR psychological safety at the group level. H5 was only partially supported as CGE scores were positively related to both TR psychological safety, *(r = .68, p < .001)* and IR psychological safety, (*r* = .69, *p < .001)* at the group level of analysis. H6, which proposed the same relationships at the individual level was also only partially supported. CGE scores were positively related to TR psychological safety, *(r = .52, p < .001)* and IR psychological safety, (*r* = .53, *p* < *.001*) at the individual level of analysis. A Fisher's z test found that the relationship between CGE and IR psychological safety and CGE and TR psychological safety did not significantly differ at the individual level (*z* = -.13, *p* = .90).

3.3.8 Examining CGE and CSE as Predictors of IR and TR Psychological Safety Using Multilevel Modeling

The predictive ability of CSE and CGE on IR and TR psychological safety was also examined using multilevel modeling. Given that ICC(1) for IR psychological safety was not as low in Study 2 compared to Study 1, and the non-negligible between-group variance that was reflected in an ICC(2) value of 0.33 for IR psychological safety, a multilevel modeling approach is warranted. This approach allowed me to account for the nested structure of the data

(individuals within teams) and explore CSE and CGE as predictors of IR and TR psychological safety at both the individual and team levels.

3.3.8.1 Examining CSE as a Predictor of IR and TR Psychological Safety Using Multilevel Modeling

I first ran a model examining the relationship between CSE and IR and TR psychological safety at the individual and team levels. At the team level, CSE and IR psychological safety was found to be non-significant (b = .25, p = .69). This finding differs in interpretation from the correlational analysis in the previous section and in Study 1 that found a positive relationship between aggregated scores of CSE and IR psychological safety. Moreover, CSE and TR psychological safety was also found to be non-significant (b = -.51, p = .67), partially supporting H3 and replicating what was found in Study 1.

Next, I looked at the relationship between CSE and IR psychological safety and CSE and TR psychological safety at the individual level using MLM. Similar to the correlational analysis in Study 2 and the previous results of Study 1, CSE was positively related to IR psychological safety (b = .27, p < .001) and TR psychological safety at the individual level (b = .24, p < .001). Thus, H4 was supported.

Table 13. *MLM Regressions Examining CSE on IR and TR Psychological Safety*

Predictors	Level	IR Psychological Safety	TR Psychological Safety
CSE	Within	b = .27** [.16, .41]	b =.24** [.13, .34]
CSE	Between	b = .25 [-1.80, 2.63]	b = -.51 [-2.90, 1.85]

Note. b = unstandardized regression coefficient. CSE = Core self-evaluations, CGE = Core group-evaluations, IR = Individual referent, TR = Team referent. Square brackets contain values representing 95% confidence intervals; intervals that contain zero are considered non-significant. * = p < .05, ** = p < .001.

3.3.8.2 Examining CGE as a Predictor of IR and TR Psychological Safety Using Multilevel Modeling

I then ran a model examining the relationship between CGE and TRPS at the individual and team levels using MLM. Similar to the correlational analysis, CGE was found to be positively related to IR psychological safety (b = .47, p < .001) and TR psychological safety (b =

.49, $p < .001$) at the individual level. At the team level, CGE was also found to be positively related to IR psychological safety (b = .71, $p < .001$) and TR psychological safety (b = .97, $p < .001$. These findings were similar to the correlational analysis and partially support H5, which proposed that CGE would not be related to IR psychological safety at the team level.

Table 14. *MLM Regressions Examining CGE on IR and TR Psychological Safety*

Predictors	Level	IR Psychological Safety	TR Psychological Safety
CGE	Within	b = .47** [.34, .60]	b =.49** [.36, .62]
CGE	Between	b = .71** [.50, .91]	b =.97**[.69, 1.15]

Note. b = unstandardized regression coefficient. CSE = Core self-evaluations, CGE = Core group-evaluations, IR = Individual referent, TR = Team referent. Square brackets contain values representing 95% confidence intervals; intervals that contain zero are considered non-significant. * = $p < .05$, ** = $p < .001$.

3.3.8.3 Examining both CGE and CSE as Predictors of IR and TR Psychological Safety Using Multilevel Modeling

I also ran multilevel regressions on IR psychological safety and TR psychological safety including both CGE and CGE as predictors. This allowed me to account for the shared variance between the predictors to provide a more complete picture of the relationships between the variables of interest and reduce the potential for inflated Type 1 error rates.

At the team level, CSE was found to have no significant relationship with IR psychological safety (b = -.15, p = .50). In contrast, CSE was found to be negatively related to TR psychological safety at the team level (b = -.65, p = .03). This finding did not support H3, which proposed that CSE would be positively related to IR psychological safety but unrelated to TR psychological safety at the group level.

At the individual level, CSE was positively related to IR psychological safety (b = .25, $p < .001$) and TR psychological safety (b = .21, $p < .001$). CGE was also positively related to IR psychological safety (b = .45, $p < .001$) and TR psychological safety (b = .48, $p < .001$) at the individual level. Thus, H4 was supported.

At the team level, CGE was positively related to IR psychological safety (b = .47, p < .001) and TR psychological safety (b = .49, p < .001. These findings partially support H5, as H5 proposed that CGE would not be related to IR psychological safety at the team level.

At the individual level, CGE was positively related to IR psychological safety (b = .45, p < .001) and TR psychological safety (b = .48, p < .001). These findings partially support H6, as H6 proposed that CGE would not be related to IR psychological safety at the individual level.

Notably, when I accounted for the effect of CGE on TR psychological safety at the team level and controlled for the interdependent nature of individual data within teams using multilevel modeling, a significant negative relationship between CSE and TR psychological safety emerged at the team level. This finding contrasts with previous analyses, which showed a non-significant correlation between team-level CSE and team-level TR psychological safety. Specifically, the relationship between CSE and TR psychological safety was non-significant in the multilevel model that did not include CGE as a predictor of TR psychological safety. Similarly, when the data was analyzed using correlations between aggregated team-level scores without using MLM, the relationship between CSE and TR psychological safety at the team level appeared to be non-significant. However, the MLM results account for the nested structure of the data and may be more likely to provide a more accurate estimate of the relationships between the variables. Therefore, the significant negative relationship between CSE and TR psychological safety at the team level, as revealed by the MLM analysis, may be likely to be a more reliable finding. The results of the regressions including both CSE and TR psychological safety using MLM are shown below in Table 15.

Table 15. *MLM Regressions Examining CSE and CGE as Predictors of IR and TR Psychological Safety*

Predictors	Level	IR Psychological Safety	TR Psychological Safety
CSE	Within	b = .25** [.16, .34]	b = .21** [.12, .30]
CGE	Within	b = .45** [.33, .57]	b = .48** [.36, .60]
CSE	Between	b = -.15 [-.57, .28]	b = -.65** [-1.24, -.06]

CGE	Between	b = .74* [.52, .95]	b = .92**[.69, 1.15]

Note. b = unstandardized regression coefficient. CSE = Core self-evaluations, CGE = Core group-evaluations, IR = Individual referent, TR = Team referent. Square brackets contain values representing 95% confidence intervals; intervals that contain zero are considered non-significant. * = $p < .05$, ** = $p < .001$.

Wald tests of parameter constraints were conducted to determine whether the strength of the relationships with CSE and CGE as predictors significantly differed between IR and TR psychological safety scales when they were significant and positive (i.e., in the same direction). Specifically, there was no significant difference in the relationships between CSE and IRPS and between CSE and IRPS at the individual level. A Wald test suggested a significant difference in the relationships between CGE and IRPS and between CGE and TRPS at the team level ($\chi^2(1) = 9.778$, $p = .002$) but not at the within level. Thus, this finding suggests that the relation between CGE and TR psychological safety is stronger than the relation between CGE and IR psychological safety at the team level.

3.3.9 IR and TR psychological Safety and Team Member Behaviours

3.3.9.1 Task Performance

H7 proposed that both a) IR psychological safety and b) TR psychological safety would be positively related to peer ratings of task performance at the team level of analysis. H7 was supported, as both IR psychological safety ($r = .43$, $p < .001$*)* and TR psychological safety ($r = .53$, $p < .001$*)* were positively related to peer ratings of task performance at the team level of analysis. H8 proposed that both a) IR psychological safety and b) TR psychological safety would be positively related to peer ratings of task performance at the individual level of analysis. H8 was also supported, as both IR psychological safety ($r = .33$, $p < .001$*)* and TR psychological safety ($r = .30$, $p < .001$*)* were positively related to peer ratings of task performance at the individual level of analysis.

Hierarchical linear modeling using grand mean centering for predictor variables yielded a similar pattern of results. Both a) IR psychological safety (b = .94, $p < .001$) and b) TR psychological safety ($b = .97$, $p < .001$) were positively related to peer ratings of task performance at the team level of analysis. Both IR psychological safety ($b = .66$, $p < .001$*)* and

TR psychological safety (*b* = .59, *p* < .001*)* were also positively related to peer ratings of task performance at the individual level of analysis.

3.3.9.2 Citizenship Behaviours

H9 proposed that both a) IR psychological safety and b) TR psychological safety would be positively related to peer ratings of citizenship behaviour at the team level of analysis. H9 was supported, as both IR psychological safety (*r* = .50, *p* < .001*)* and TR psychological safety (*r* = .57, *p* < .001*)* were positively related to peer ratings of citizenship behaviour at the team level of analysis. H10 proposed that both a) IR psychological safety and b) TR psychological safety would be positively related to peer ratings of citizenship behaviour at the individual level of analysis. H10 was supported, as both IR psychological safety (*r* = .37, *p* < .001*)* and TR psychological safety (*r* = .33, *p* < .001*)* were positively related to peer ratings of citizenship behaviour at the individual level of analysis.

Hierarchical linear modeling using grand mean centering for predictor variables yielded a similar pattern of results. Both a) IR psychological safety (b = 0.53, *p* <.001) and b) TR psychological safety (*b* = 1.51, *p* < .001) were positively related to peer ratings of citizenship behaviour at the team level of analysis. Both IR psychological safety (*b* = .70, *p* < .001*)* and TR psychological safety (*b* = .53, *p* < .001*)* were also positively related to peer ratings of citizenship behaviour at the individual level of analysis.

3.3.9.3 Counterproductive Work Behaviours

H11 proposed that both a) IR psychological safety and b) TR psychological safety would be negatively related to peer ratings of counterproductive behaviour at the team level of analysis. H11 was supported, as both IR psychological safety (*r* = -.27, *p* < .01*)* and TR psychological safety (*r* = -.39, *p* < .001*)* were negatively related to peer ratings of counterproductive behaviour at the team level of analysis. H12 proposed that both a) IR psychological safety and b) TR psychological safety would be negatively related to peer ratings of counterproductive behaviour at the individual level of analysis. H12 was supported, as both IR psychological safety (*r* = -.11, *p* < .01*)* and TR psychological safety (*r* = -.16, *p* < .001*)* were negatively related to peer ratings of counterproductive behaviour at the individual level of analysis.

Hierarchical linear modeling using grand mean centering for predictor variables yielded a similar pattern of results for the team level but not for the individual level. Both a) IR psychological safety (b = -.56, p <.001) and b) TR psychological safety (b = -.72, p < .001) were negatively related to peer ratings of counterproductive work behaviours at the team level of analysis. However, IR psychological safety (b = .00, p = .97*)* and TR psychological safety (b = .02, p = .76*)* were not significantly related to peer ratings of counterproductive behaviours at the individual level of analysis.

3.3.10 IR and TR Psychological Safety, Perceived Team Performance and Graded Team Effectiveness

H13 proposed that both a) IR psychological safety and b) TR psychological safety would be positively related to perceived team performance at the team level of analysis. As shown in Table 10, H13a and b were supported, as both IR psychological safety (r = .45, p < .001*)* and TR psychological safety (r = .47, p < .001*)* were positively related to perceived team performance at the team level of analysis. H14 proposed that both a) IR psychological safety and b) TR psychological safety would be positively related to perceived team performance at the individual level of analysis. Hypothesis 14a and b were supported, as both IR psychological safety (r = .26, < .001*)* and TR psychological safety (r = .23, < .001.*)* were positively related to perceived team performance at the individual level of analysis.

Hierarchical linear modeling using grand mean centering for predictor variables yielded a similar pattern of results. Both a) IR psychological safety (b = 0.34, p <.001) and b) TR psychological safety (b = 1.01, p < .001) were positively related to perceived team performance at the team level of analysis. Both IR psychological safety (b = .34, p = .002*)* and TR psychological safety (b = .38, p < .001*)* were also positively related to perceived team performance at the individual level of analysis.

H15 proposed that both c) IR psychological safety and b) TR psychological safety would be positively related to the graded team effectiveness at the team level of analysis. H15 was not supported as both IR psychological safety *(*r = .07, p = .63*)* and TR psychological safety *(*r = .13, p = .32*)* were unrelated to graded team effectiveness.

3.3.11 Incremental Validity of IR and TR psychological safety

Hierarchical linear modeling (HLM; Raudenbush & Bryk, 2002) was used to assess the incremental validity of IR and TR psychological safety using HLM Version 8.2.3 (Scientific Software International). Hierarchical two-level regression analysis was conducted using maximum likelihood estimation. ΔAIC scores were used to determine the incremental validity in predicting the team outcomes. Traditional ΔR-squared measures were not applicable in the HLM framework, so I relied on AIC changes to assess incremental validity, with decreases in AIC indicating improved model fit and support for the hypotheses as lower AIC scores are considered to reflect a better model fit.

H16 proposed that IR psychological safety would predict variance in individual-level ratings of their team members' a) task performance, b) citizenship behaviours, and c) counterproductive behaviours above and beyond TR psychological safety. As shown in Table 16. , H16a was supported in that the model including both IR and TR psychological safety predicting task performance provided a better model fit beyond TR psychological safety at the individual level (Δ*AIC* = 8.68*)*. Moreover, the relationship between TR psychological safety and task performance was no longer significant when IR psychological safety was included in the model *(b = .18, p = .29). H16b* was supported as the model including both IR and TR psychological safety provided a better model fit beyond TR psychological safety in predicting citizenship behaviours at the individual level (Δ*AIC* = 17.12*)*. Moreover, the relationship between TR psychological safety and citizenship behaviours was no longer significant when IR psychological safety was included in the model *(b* = -.10, *p* = .59*)*. H16c was not supported as the model including IR and TR psychological safety did not predict counterproductive behaviour above and beyond TR psychological safety at the individual level (ΔAIC = -0.69). Neither IR psychological safety nor TR psychological safety significantly predicted counterproductive behaviour when both were included as predictors in the model.

H17 proposed that TR psychological safety would predict team-level ratings of team members' a) task performance b) citizenship behaviours, c) counterproductive behaviours and ratings of d) perceived team performance, and e) graded team effectiveness above and beyond IR psychological safety. As shown in Table 17, Hypothesis 17a was supported in that the model including both IR and TR psychological safety predicting task performance provided a better

model fit beyond IR psychological safety at the team level (Δ*AIC = 10.14)*. Moreover, the relationship between IR psychological safety and task performance was no longer significant when TR psychological safety was included in the model *(b* = -.01, *p* = .96*)*. H17b was supported as the model including both IR and TR psychological safety provided a better model fit beyond IR psychological safety in predicting citizenship behaviours at the team level (Δ*AIC = 7.98)*. Moreover, the relationship between IR psychological safety and citizenship behaviours was no longer significant when TR psychological safety was included in the model *(b* = .48, *p* = .40). H17c was also supported. Including both IR and TR psychological safety predicting counterproductive work behaviours provided a better model fit beyond IR psychological safety at the team level (Δ*AIC* = 11.23*)*. Moreover, the relationship between IR psychological safety and counterproductive work behaviours was no longer significant when TR psychological safety was included in the model *(b* = .58, *p* = .12*)*.

Hypothesis 17d was not supported as a model that included TR psychological safety as a predictor did not provide a better fit for predicting perceived team performance beyond IR psychological safety at the team level. Both the relationship between IR psychological safety and TR psychological safety and perceived team performance became non-significant when both predictors were included in a model predicting perceived team performance. Finally, Hypothesis 17e was not supported as there was no significant relationship between psychological safety and graded team effectiveness, so it was not tested.

Table 16. *Summary of the Hierarchical Linear Model Results for the Incremental Contribution of IR Psychological Safety Beyond TR Psychological Safety at the Individual Level*

Task Performance

Block	Variable	b_{Block1}	b_{Block2}	Deviance	AIC	ΔAIC
1	TR Psychological Safety	.59** (.10)	.18 (.17)	1046.86	1050.86	
2	IR Psychological Safety		.51* (.20)	1038.18	1042.18	8.68

Citizenship Behaviours

Block	Variable	b_{Block1}	b_{Block2}	Deviance	AIC	ΔAIC
1	TR Psychological Safety	.53** (.12)	-.10 (.19)	1133.07	1137.07	
2	IR Psychological Safety		.78** (.19)	1115.95	1119.95	17.12

Counterproductive Work Behaviours

Block	Variable	b_{Block1}	b_{Block2}	Deviance	AIC	
1	TR Psychological Safety	.02 (.07)	.04 (.10)	757.19	761.19	
2	IR Psychological Safety		-.03 (.10)	757.88	761.88	-0.69

Note. *b* = coefficient; standard errors are in parentheses. AIC was calculated as Deviance + 2*k*, where *k* = number of estimated parameters. ΔAIC was calculated as Block 2 AIC – Block 1 AIC. IR = Individual-referent, TR = Team-referent. ** $p < .001$. * $p < .05$.

Table 17. *Summary of the Hierarchical Linear Model Results for the Incremental Contribution of TR Psychological Safety Beyond IR Psychological Safety at the Team Level*

Task Performance

Block	Variable	b_{Block1}	b_{Block2}	Deviance	AIC	ΔAIC
1	IR Psychological Safety	.94** (.16)	-.01 (.32)	1052.23	1056.23	
2	TR Psychological Safety		.98** (.27)	1042.09	1046.09	10.14

Citizenship Behaviours

Block	Variable	b_{Block1}	$b_{Block\ 2}$	Deviance	AIC	ΔAIC
1	IR Psychological Safety	1.57** (.27)	.48 (.57)	1118.63	1122.63	
2	TR Psychological Safety		1.13* (.50)	1110.65	1114.65	7.98

Counterproductive Work Behaviours

Block	Variable	b_{Block1}	b_{Block2}	Deviance	AIC	ΔAIC
1	IR Psychological Safety	-.56* (.21)	.58 (.41)	747.95	751.95	
2	TR Psychological Safety		-1.18* (.37)	736.72	740.72	11.23

Perceived Team Performance

Block	Variable	b_{Block1}	b_{Block2}	Deviance	AIC	ΔAIC
1	IR Psychological Safety	1.06** (.19)	.37 (.43)	950.94	954.94	
2	TR Psychological Safety		.72 (.39)	949.49	953.49	1.45

Note. b = coefficient; standard errors are in parentheses. AIC was calculated as Deviance + 2k, where k = number of estimated parameters. ΔAIC was calculated as Block 2 AIC – Block 1 AIC. IR = Individual-referent, TR = Team-referent. ** $p < .001$. * $p < .05$.

3.3.12 Psychological Safety as a Mediator

Using MPlus 8.6, multilevel modeling was employed to test hypotheses related to the role of psychological safety as a mediator at the individual and team level. The procedure recommended by Kelloway (2015) for calculating and estimating multilevel indirect effects in Mplus was followed. Evidence of mediation is supported if the confidence interval for the estimate of an indirect effect did not include the value 0.

H18 proposed that at the individual level, IR psychological safety would mediate the relationship between CSE and members' a) task performance, b) citizenship behaviours, and c) counterproductive behaviours. Hypothesis 18a was supported. The indirect effect of CSE on task performance through IR psychological safety was b_{indir} = .16 (C.I. = .06, .27). Hypothesis 18b was also supported. The indirect effect of CSE on citizenship behaviours through IR psychological safety was b_{indir} = .15 (C.I. = .05, .26). Hypothesis 18c was not supported. The indirect effect of CSE on counterproductive behaviours through IR psychological safety was b_{indir} = .02 (C.I. = -.03, .07).

H19 proposed that IR psychological safety will mediate the relationship between CSE and members' a) task performance b) citizenship behaviours, c) counterproductive behaviours d) perceived team performance, and e) graded team effectiveness at the team level. Hypothesis 19a was not supported. The indirect effect of CSE on task performance through IR psychological safety was b_{indir} = -.30 (C.I. = -3.55, 2.17). Hypothesis 19b was not supported. The indirect effect of CSE on citizenship behaviours through IR psychological safety was b_{indir} = .56 (C.I. = -10.34, 9.23). Hypothesis 19c was also not supported. The indirect effect of CSE on counterproductive behaviours through IR psychological safety was b_{indir} = -.41 (C.I. = -10.41, 9.60). Hypothesis 19d was not supported. The indirect effect of CSE on perceived team performance through IR psychological safety was b_{indir} = 1.48 (C.I. = -5.88, 8.84). Hypothesis 19e was also not supported. There was no relationship between IR psychological safety and graded team effectiveness, so the mediation effect was not tested.

H20 proposed that TR psychological safety would mediate the relationship between CGE and a) task performance b) citizenship behaviours, and c) counterproductive behaviours at the individual level. Hypothesis 20a was supported. The indirect effect of CGE on task performance

through TR psychological safety was b_{indir} = .31 (C.I. = .16, .46). Hypothesis 20b was supported. The indirect effect of CGE on citizenship behaviours through TR psychological safety was b_{indir} = .26 (C.I. = .11, .40). Hypothesis 20c was not supported. The indirect effect of CGE on counterproductive behaviours through TR psychological safety was b_{indir} = -.08 (C.I. = -.05, .11).

H21 proposed that TR psychological safety would mediate the relationship between CGE and a) task performance b) citizenship behaviours, c) counterproductive behaviours d) perceived team performance, and e) graded team effectiveness at the team level. Hypothesis 21a was not supported. The indirect effect of CGE on task performance through TR psychological safety was b_{indir} = .49 (C.I. = -.23, 1.21). Hypothesis 21b was not supported. The indirect effect of CGE on citizenship behaviours through TR psychological safety was b_{indir} = 1.94 (C.I. = -2.25, 6.12). Hypothesis 21c was not supported. The indirect effect of CGE on counterproductive behaviours through TR psychological safety was b_{indir} = -3.72 (C.I. = -7.38, -.05). Hypothesis 21d was supported. The indirect effect of CGE on perceived team performance through TR psychological safety was b_{indir} = 2.68 (C.I. = .14, 5.21). Hypothesis 21e was not supported. There was no relationship between TR psychological safety and graded team effectiveness, so the mediation effect was not tested.

A summary of hypotheses and findings of Study 2 are presented in Table 18 below.

Table 18. *Summary of Hypotheses and Findings for Study 2*

Hypothesis statement	Supported?	Summary of Finding
Justification for Aggregation-based Conceptualization		
H1a: ICC(1) for TRPS will be sufficiently high to support the validity of a referent-shift consensus operationalization. H1b: ICC(1) for IRPS will not be sufficiently high to support the validity of a direct consensus operationalization.	H1a-b: Supported	• ICCs were lower for IRPS than TRPS • IRPS: ICC1 = .11; ICC2 = .33 • TRPS: ICC1 = .14; ICC2 = .40
Factor Structure of IRPS and TRPS		
H2: Scores on the IRPS items and TRPS items will load on two separate factors.	H2: Supported	• Significant χ^2 difference test [$\chi^2(1)_{\text{difftest}}$ = 31.30, p <.001] for two-factor modified model vs. one-factor modified model ○ One-factor model: $\chi^2(69)$ = 217.19, CFI = .93, TLI = .91, RMSEA = .08, and SRMR = .07 ○ Two-factor model: $\chi^2(70)$ = 248.49, CFI = .91, TLI = .89, RMSEA = .08, and SRMR = .07
CSE as an Antecedent of IRPS and TRPS		
H3: At the team level of analysis, CSE will be a) positively related to IRPS and b) unrelated to TRPS.	H3a: Not Supported H3b: Supported	• H3a: n.s. ○ b = .25 [-1.80, 2.63] • H3b: n.s. ○ b = -.51 [-2.90, 1.85] *Note:* CSE was negatively related to TRPS when accounting the effect of CGE at the team level • b = -.65** [-1.24, -.06]
H4: At the individual level of analysis, CSE will be positively related to a) IRPS and b) TRPS.	H4a: Supported H4b: Supported	• H4a: b = .27** [.16, .41] • H4b: b = .24** [.13, .34]
CGEs as Antecedents of IRPS and TRPS		
H5: At the team level of analysis, CGE will be a) positively related to TRPS but b) unrelated to IRPS.	H5a: Supported H5b: Not Supported	• H5a: b = .97**[.69, 1.15] • H5b: b = .71** [.50, .91]
H6: At the individual level of analysis, CGE will be a) positively related to TRPS but b) unrelated to IRPS.	H6a: Supported H6b: Not Supported	• H6a: b = .47** [.34, .60] • H6b: b = .49** [.36, .62]

IR and TR Psychological Safety and Team Member Behaviours		
H7: At the team level of analysis, both a) IRPS and b) TRPS will be positively related to aggregated peer ratings of task performance.	H7a: Supported H7b: Supported	• H7a: b = .94, p <.001 • H7b: b = .97, p < .001
H8: At the individual level of analysis, both a) IRPS and b) TRPS will be positively related to peer ratings of task performance.	H8a: Supported H8b: Supported	• H8a: b = .66, p < .001 • H8b: b = .59, p < .001
H9: At the team level of analysis, both a) IRPS and b) TRPS will be positively related to aggregated peer ratings of citizenship behaviour.	H9a: Supported H9b: Supported	• H9a: b = 0.53, p <.001 • H9b: b = 1.51, p < .001
H10: At the individual level of analysis, both a) IRPS and b) TRPS will be positively related to peer ratings of citizenship behaviour.	H10a: Supported H10b: Supported	• H10a: b = .70, p < .001 • H10b: b = .53, p < .001
H11: At the team level of analysis, both a) IRPS and b) TRPS will be negatively related to aggregated peer ratings of counterproductive behaviour.	H11a: Supported H11b: Supported	• H11a: b = -.56, p <.001 • H11b: b = -.72, p < .001
H12: At the individual level of analysis, both a) IRPS and b) TRPS will be negatively related to peer ratings of counterproductive behaviour.	H12a: Not Supported H12b: Not Supported	• H12a: b = .00, p = .97 • H12b: b = .02, p = .76
IRPS and TRPS and Team Performance		
H13: At the team level of analysis, both a) IRPS and b) TRPS will be positively related to perceived team performance.	H13a: Supported H13b: Supported	• H13a: b = 0.34, p <.001 • H13b: b = 1.01, p < .001
H14: At the individual level of analysis, both a) IRPS and b) TRPS will be positively related to perceived team performance.	H14a: Supported H14b: Supported	• H14a: b = .34, p = .002 • H14b: b = .38, p < .001
H15: At the team level of analysis, both a) IRPS and b) TRPS will be positively related to graded team effectiveness (i.e., final team project grade).	H15a: Not Supported H15b: Not Supported	• H15a: r = .07, p = .63 • H15b: r = .13, p = .32
Incremental Validity of IRPS and TRPS		
H16: At the individual level of analysis, IRPS will predict variance in peer ratings of team members' a) task performance, b) citizenship behaviours, and c) counterproductive behaviours above and beyond TRPS.	H16a: Supported H16b: Supported H16c: Not Supported	• H16a: ΔAIC = 8.68 • H16b: ΔAIC = 17.12 • H16c: ΔAIC = -0.69
H17: At the team level of analysis, TRPS will predict variance in aggregated peer ratings of team members' a) task performance b) citizenship behaviours, c) counterproductive behaviours; d) perceived team performance, and e) graded team effectiveness above and beyond IRPS.	H17a: Supported H17b: Supported H17c: Supported H17d: Not Supported H17e: Not Supported	• H17a: ΔAIC = 10.14 • H17b: ΔAIC = 7.98 • H17c: ΔAIC = 11.23 • H17d: ΔAIC = 1.45 • H17e: Not tested because H15 was not supported.
Psychological Safety as a Mediator		
H18: At the individual level of analysis, IR psychological safety will mediate the relationship between CSE and a) task performance b) citizenship behaviours, and c) counterproductive behaviour.	H18a: Supported H18b: Supported H18c: Not Supported	• H18a: b_{indir} = .16 (.06, .27) • H18b: b_{indir} = .15 (.05, .26) • H18c: b_{indir} = .02 (-.03, .07)

H19: At the team level of analysis, IR psychological safety will mediate the relationship between CSE and a) task performance b) citizenship behaviours, c) counterproductive behaviours d) perceived team performance, and e) graded team effectiveness at the team level.	H19a: Not Supported H19b: Not Supported H19c: Not Supported H19d: Not Supported H19e: Not Supported	• H19a: b_{indir} = -.30 (-3.55, 2.17) • H19b: b_{indir} = .56 (-10.34, 9.23) • H19c: b_{indir} = -.41 (-10.41, 9.60) • H19d: b_{indir} = 1.48 (-5.88, 8.84) • H19e: n.s.
H20: At the individual level of analysis, TR psychological safety will mediate the relationship between CGE and a) task performance b) citizenship behaviours, and c) counterproductive behaviours.	H20a: Supported H20b: Supported H20c: Not Supported	• H20a: b_{indir} = .31 (.16, .46) • H20b: b_{indir} = .26 (.11, .40) • H20c: b_{indir} = -.08 (-.05, .11)
H21: At the team level, TR psychological safety will mediate the relationship between CGE and a) task performance b) citizenship behaviours, c) counterproductive behaviours d) perceived team performance, and e) graded team effectiveness.	H21a: Not Supported H21b: Not Supported H21c: Not Supported H21d: Supported H21e: Not Supported	• H21a: b_{indir} = .49 (-.23, 1.21) • H21b: b_{indir} = 1.94 (-2.25, 6.12) • H21c: b_{indir} = -3.72 (-7.38, -.05) • H21d: b_{indir} = 2.68 (.14, 5.21) • H21e: Not tested since H15 was not supported.

3.4 Study 2 Discussion

Theoretically, IR psychological safety represents shared individual perceptions of psychological safety given that the IR psychological safety measure in the study references the self, whereas TR psychological safety represents a shared perception of the team's psychological safety. Teams researchers generally recommend that the best practice for conceptualizing team constructs is to specifically reference the appropriate level of the phenomena of interest (Glick, 1985; Klein et al., 1994; Morgeson & Hofmann, 1999; Rousseau, 1985). Given that Edmondson's (1999) definition refers to "a shared belief held by members of the team that the team is safe for interpersonal risk taking" (p. 350) and considers the construct to reside at the group level (Edmondson & Lei, 2014) the appropriate measure would use a team referent.

The first objective of Study 2 was to determine whether the findings of Study 1 regarding the validity of conceptualizing psychological safety as a referent-shift consensus model were replicated. Consistent with Study 1 results, a referent-shift consensus model for operationalizing TR psychological safety was supported by a sufficiently high ICC(1) value to justify aggregation. Thus, it appears that TR psychological safety can be considered a shared group construct. On the other hand, the ICC(1) value for IR psychological safety was not sufficiently high, suggesting that a direct consensus model for operationalizing IR psychological safety may not be viable.

To determine whether IR psychological safety and TR psychological safety operate as distinct constructs, measurement models including both IR psychological safety and TR psychological safety were assessed. The results suggested better model fit for a modified two-factor model in which IR psychological safety and TR psychological safety were correlated compared to a modified one-factor model in which IR psychological safety and TR psychological safety were correlated. Although a high correlation was found between IR psychological safety and TR psychological safety indicating substantial overlap between the constructs (r = .82 at the individual level of analysis and r = .88 at the team level of analysis), the findings of the CFAs that were

conducted suggest that both IR psychological safety and TR psychological safety may be conceptually distinct.

The second objective of Study 2 was to examine CSE and CGE as antecedents of IR and TR psychological safety. Similar to the findings of Study 1, CSE was found to relate positively to IR psychological safety but not to TR psychological safety at the group level. Moreover, CSE was positively related to both IR and TR psychological safety at the individual level similar to Study 1. Interestingly, multilevel modeling resulted in a deviation from correlational results. First, the relationship that was found in the team-level correlational analysis between CSE and IR psychological safety was no longer significant at the team level in the MLM results. When a multilevel model included both CSE and CGE as predictors, an unexpected negative relationship emerged between CSE and TR psychological safety at the team level. This suggests that while individuals' CSE level may not directly influence their scores on how they individually rate the psychological safety of team members as a whole, the overall level of CSE in a team might affect team members' ratings of TR psychological safety at the group level. It is possible that a high level of CSE in a team may influence team dynamics. For example, if there are multiple high CSE team members in a group, they may exude confidence in a way that leads other team members CSE to feel pressured or to conform to a dominant attitude in situations, leaving less room for open discussion and team-decision making. This may dissuade other team members from feeling safe to express their diverse perspectives. An unsafe environment, characterized by suppressed dissent, has been responsible for well-known catastrophic team failures, such as the 1986 NASA Challenger disaster, where a dominant culture of overconfidence and pressure to conform was attributed to a tragic oversight of critical safety concerns.

The results of Study 2 also appear to suggest that TR psychological safety has a different relationship with CSE at the team level from IR psychological safety, but not at the individual level. This finding provides evidence suggesting that IR psychological safety and TR psychological safety may operate as distinct constructs. Moreover, the results suggest that CGE may be a robust predictor of both IR and TR psychological safety.

The third objective of Study 2 was to examine the incremental validity of IR and TR psychological safety in predicting team-related behaviours (task performance, citizenship behaviour, and counterproductive behaviour) and team performance. The incremental validity of IR and TR psychological safety was compared by examining the incremental variance accounted for by IR and TR psychological safety in team member behaviours (task performance, citizenship behaviour, and counterproductive behaviour) and team performance. The results suggested that IR psychological safety predicted team member behaviours (task performance and citizenship behaviour but not counterproductive behaviour) above and beyond TR psychological safety at the individual level of analysis whereas TR psychological safety predicted the team-related behaviours (task performance, citizenship behaviour, and counterproductive behaviour) above and beyond IR psychological safety at the team level of analysis. Interestingly, TR psychological safety did not account for significant incremental variance for perceived team performance above and beyond IR psychological safety at the team level of analysis as predicted. These findings again suggest that IR psychological safety and TR psychological safety may operate as distinct constructs.

The final objective of Study 2 was to examine the mediating role of IR and TR psychological safety. At the individual level, IR psychological safety mediated the relationship between CSE and task performance and citizenship behaviours, but not for counterproductive behaviours. At the team level of analysis, however, IR psychological safety did not mediate the relationship between CSE and team outcomes. TR psychological safety was found to mediate the relationship between CGE and task performance and citizenship behaviours, at the individual level of analysis, but not at the team level of analysis. TR psychological safety was found to mediate the relationship between CGE and counterproductive behaviours and perceived team performance ratings at the team level of analysis.

4 General Discussion

4.1 Contributions of the Current Research

Psychological safety has been shown to be an important factor in determining team-related outcomes in organizational research (Frazier et al., 2017). The current dissertation consisted of two studies that examined the validity of two conceptualizations of psychological safety: IR psychological safety and TR psychological safety. These two conceptualizations of psychological safety varied with respect to the unit of the referent (i.e., individual vs. team).

To my knowledge, this is the first examination that has investigated whether the level of referent use influences relations between psychological safety and both predictor and outcome variables. Overall, the findings of the study suggest that IR psychological safety and TR psychological safety may be constructs that are related, but not the same.

In this dissertation, I identified that the referents used in the most widely used measure of psychological safety developed by Edmondson (1999) are mixed and inconsistent within the scale. General principles of teams research posit that the level of theory and level of referent should be consistent (Kozlowski & Klein, 2000). Given that the unit referenced in Edmondson's (1999) definition is the *team*, it appears that TR psychological safety is the most appropriate conceptualization of the construct that Edmondson (1999) describes as team psychological safety.

Study 1 and 2 findings both demonstrated that TR psychological safety as conceptualized by a referent-shift model appeared to be an appropriate operationalization for psychological safety as a shared team construct. Both studies also demonstrated that IR psychological safety may not be a shared construct among team members as evidenced by a low ICC(1) value. The findings suggest that an appropriate composition model may be an additive composition model for IR psychological safety, in which the meaning a group construct at the group level is a summation of the lower individual-level units regardless of the variance among units (Chan, 1998). However, ICC(1) and ICC(2)

for IR psychological safety were somewhat higher in Study 2 than in Study 1, suggesting that IR psychological safety was more homogenous within teams in Study 2. Thus, future research replicating this finding would strengthen confidence in these conclusions. Moreover, there may be situations where team members show greater within-group agreement on IR psychological safety depending on various factors, such as the degree of interdependence among team members.

Overall, the findings of Study 2 demonstrated that IR psychological safety is a stronger predictor of team-related outcomes in comparison to TR psychological safety at the individual level of analysis whereas TR psychological safety is a stronger predictor of team-related outcomes at the team level of analysis.

Taken together, the findings of Study 1 and Study 2 also contribute to the understanding of the nomological network of psychological safety. CSE and CGE were examined as antecedents of psychological safety. Consistent across both studies, CSE appears to be an antecedent of IR psychological safety at the individual level. At the group level, the findings relating to the relationship between CSE and TR psychological safety were unexpected but interesting. The findings suggested that CSE may not influence how individuals rate the psychological safety of team members at the individual level, but a higher collective level of CSE in a team relates to lower ratings of TR psychological safety at the group level when accounting for the effect of CGE.

Study 2 built on Study 1 by examining IR psychological safety and TR psychological safety as mediators. The findings suggest that at the individual level, IR psychological safety plays a mediating role in the relationship between CSE and both task performance and citizenship behaviours. TR psychological safety plays a mediating role in the relationship between CGE and both task performance and citizenship behaviours at the individual level. Additionally, TR psychological safety mediates the relationship between CGE and perceived team performance ratings at the team level.

4.2 Implications of Findings

In the vast majority of research on psychological safety, researchers do not specify the composition model used to conceptualize psychological safety. Based on the findings of my dissertation, I recommend that researchers explicitly identify the level of construct, level of referent, and level of analysis in their operationalization of psychological safety. If researchers are interested in team psychological safety as a shared group construct, as defined by Edmondson (1999), then TR psychological safety is the appropriate operationalization. If researchers are more interested in examining individual perceptions of psychological safety which is more aligned with Kahn's (1990) definition of psychological safety, then IR psychological safety is likely the appropriate operationalization.

Our results also suggest that using individual referent items when the intention is to examine team psychological safety at the team level is likely to lead to underestimated effect sizes when examining its relation to predictors and outcome variables. Moreover, using team referent items when the intention is to study individual perceptions of psychological safety is likely to result in underestimated effect sizes when examining its relation to predictors and outcome variables.

Given the evidence from this dissertation that IR psychological safety and TR psychological safety may represent distinct constructs, I recommend that researchers measuring psychological safety use a measure that is consistent across items with respect to the referent and matches the construct they are purporting to measure.

4.3 Limitations and Implications for Future Research

One limitation of the dissertation that may affect the generalizability of findings is that students were used to collect the data. Additionally, the majority of the sample for both studies comprised Caucasian males. Future research using more heterogeneous samples regarding demographic characteristics may be a worthy endeavour. A drawback of studying psychological safety using student samples is the lack of real-world experience and long-term professional stakes that students typically have compared to

working professionals. There are likely to be less consequences of taking interpersonal risks (such as sharing ideas or admitting mistakes) for students. As a result, findings from student samples might not fully capture the nuances of psychological safety in real-world organizational settings, limiting the generalizability of the research.

Furthermore, given the changes in the global landscape, including significant world events since the data for the dissertation was collected, there is a compelling need to update the sample. An updated dataset would likely reflect generational shifts, increased demographic diversity, evolving educational environments, and greater societal awareness of mental health, all of which could influence how psychological safety is experienced today. This would ensure that the research findings are more relevant to contemporary experiences of psychological safety.

Despite these limitations related to sample characteristics, the context studied—project teams—may still provide a meaningful basis for understanding psychological safety that is applicable to teams with characteristics of project teams. The sample in the dissertation used teams that are classified as project teams, which consist of individuals with diverse knowledge, skills, and experiences (Chiocchio, 2015). Over the course of the project, which often involves working together over long work cycles, these team members need to gather and combine significant amounts of information (Chiocchio, 2015). This collaboration is essential for defining their goals, developing or modifying approaches, and incrementally or fundamentally creating new concepts, services, products, activities, or driving broader changes (Chiocchio, 2015).

In this regard, a project team context is a suitable setting to study psychological safety because it inherently involves individuals with different backgrounds and expertise working together on complex tasks. The need to share information, clarify objectives, and innovate within a limited timeframe in project teams requires open communication, trust, and a willingness to take risks—key elements related to psychological safety. Since project teams often face uncertainty and must adapt quickly, understanding how psychological safety influences their dynamics can provide valuable insights into how teams manage challenges and succeed in collaborative environments.

Another potential limitation was the need to correlate errors in the confirmatory factor analysis in both the one-factor and two-factor model of psychological safety to achieve an acceptable model fit. The requirement to correlate errors suggests that there is unique variance between certain items that is not fully accounted for by the latent factors alone. Although errors were allowed to correlate based on identifying overlapping content or theoretical relationships between items, it is possible that the unaccounted variance may not be due to the items themselves, but to contextual factors influencing responses. Consequently, this may affect the generalizability and interpretability of the findings.

Additionally, IR psychological safety and TR psychological safety were found to be highly correlated across the two studies in this dissertation. Moreover, there was a statistically significant chi-square difference between the modified one-factor and modified two-factor models, but the overall model fit indices did not show a substantial difference between the two. This mixed evidence suggests that although the two-factor solution may offer some theoretical advantages, the lack of clear distinctiveness of the factors limits the ability to confidently interpret the factors as fully distinct constructs. Consequently, the robustness and theoretical clarity of the study's findings may be affected. However, the evidence of differentiation in relationships with other variables indicates that the factors, while highly correlated, may not be entirely redundant and may be distinct in a meaningful way. Specifically, the relationship between CGE and TRPS appeared to be stronger than the relationship between CGE and IRPS at the team level. Moreover, IR psychological safety generally appears to be a stronger predictor of team members' behaviour at the individual level whereas TR psychological safety generally appears to be a more potent predictor of behaviours at the team level. It was also found that CSE was negatively related to TR psychological safety but not related to IR psychological safety at the team level, when accounting for the effect of CGE. However, this finding should be interpreted with caution as it could be influenced by multicollinearity, which may distort the true relationships between these variables.

It was also found that there may have been a contextual influence on participants' responses to the psychological safety scales as a stronger correlation was found between

IR and TR psychological safety scores for the participants who completed the TR psychological safety first. A plausible explanation may be that there was a greater ease with which respondents could access individual-related information about psychological safety. Thus, completing the IR psychological safety measure first may have helped participants differentiate IR psychological safety from TR psychological safety, whereas completing the TR psychological safety measure first may have led participants to recall individual-related information more readily, thereby blurring the distinction between IR and TR psychological safety. Future research should further investigate the underlying causes of these order effects and explore methodological controls, such as introducing intervening tasks to separate the scales, and consider whether refining the items to enhance the distinctions between the scales could help reduce the overlap or ambiguity that might be contributing to this effect.

Another consideration is the absence of significant relationships between the variables of interest and the final project grade used as a measure for team effectiveness. This may be attributed to the multifaceted nature of team effectiveness and the fact that the team grade may have been influenced by a wide range of factors such as team members' skills and abilities, and external influences beyond the team's control, such as technical issues related to their project innovation (e.g., battery failures or other malfunctions). Given this complexity, the specific constructs examined in this study might not have directly correlated with the final project grade since team effectiveness is typically determined by the interaction of many variables.

Moreover, another limitation of this study was that the IR and TR psychological safety items were not directly compared with the original Edmondson (1999) measure. This was largely due to sample size limitations. A larger sample that enabled me to collect data on the original measure would have allowed for a more robust comparison.

4.3.1 Further Research Directions

An area of future research could examine the mechanisms through which each form of psychological safety (IR and TR psychological safety) influences outcomes. For example, psychological safety has been found to foster team performance through team

learning behaviours (e.g., Edmondson, 1999). Future research could investigate whether these mechanisms operate differently for IR psychological safety compared to TR psychological safety.

Upon close examination of the original Edmondson (1999) items, there appears to be room for improvement in refining them to more closely align with the definition of psychological safety as a shared belief that a team is safe for interpersonal risk-taking. For example, item 3 which states, *"People on this team sometimes reject me for being different"* relates to social acceptance and belonging, which is likely to influence psychological safety. However, it may be slightly misaligned as it focuses more on inclusion than on the specific concept of interpersonal risk-taking. Moreover, item 6 which states, *"No one on this team would deliberately act in a way that undermines my efforts"* is related to trust and support, which is essential for psychological safety. However, it may be more focused on intentional negative actions rather than the broader concept of feeling safe to take interpersonal risks. Item 7, which states, *"Working with members of my team, my unique skills and talents are valued and utilized"* may be problematic as this item focuses more on recognizing and utilizing individuals' contributions, which while important, does not directly relate to the core concept of perceiving a team environment to feel safe for interpersonal risk-taking. Given these observations, these items may benefit from refinement or replacement with items that more directly address the concept of interpersonal risk-taking in a team context. This would help ensure that the scale more accurately captures the essence of psychological safety as defined by Edmondson (1999) and avoid construct contamination.

Another promising direction for future research involves exploring the comparative effectiveness of different consensus models in predicting different categories of team and organizational outcomes. The incremental validity of referent-shift consensus and direct consensus models of other constructs at the group or organizational level have been examined in the literature. For example, a meta-analysis examining organizational climate by Wallace et al. (2016) found that a referent-shift operationalization of organizational climate was a stronger predictor of job performance and customer service performance than direct consensus whereas direct consensus operationalization of

organizational climate was a stronger predictor of job attitudes than referent-shift consensus. It may be interesting to examine the differential predictive ability of IR psychological safety and TR psychological safety for different categories of outcomes in future research.

4.4 Implications for Practice

Understanding how team members perceive psychological safety—both at the individual and team levels—may enhance team management and leadership strategies. The research findings indicate that there tends to be higher within-group agreement on TR psychological safety compared to IR psychological safety. This suggests that while the team as a whole may have a collective perception of being safe, individual members might experience varying levels of psychological safety.

This distinction may be important for leaders to recognize, as it implies that a general sense of safety within the team might mask the discomfort or lack of safety felt by specific individuals. Therefore, a one-size-fits-all team-level intervention aimed at boosting psychological safety might overlook individuals who need additional support. Thus, leaders can consider conducting regular check-ins with individual team members to assess their personal experiences of psychological safety. Identifying individuals who feel less psychologically safe can help leaders provide targeted support and interventions to those specific team members.

The findings also reveal a differentiated impact of Core Self-Evaluations (CSE) on IR and TR psychological safety. CSE was found to relate to higher IR psychological safety at the individual level, whereas CGE was more strongly related to both IR and TR psychological safety regardless of level of analysis, with a more substantial impact on TR psychological safety at the team level.

Interestingly, the research indicated that high CSE within a team might lead to lower TR psychological safety when accounting for CGE. This suggests that while CSE can enhance an individual's perception of safety, it might, under certain conditions, create an environment where team-level psychological safety is compromised.

The findings suggest that it may be beneficial for organizations to help individual employees enhance their CSE through personal development programs, coaching, or confidence-building activities in order strengthen individual's perception of IR psychological safety. At the team level, team-building exercises, workshops, or training programs may be used to enhance CGE, and in turn foster higher levels of IR and TR psychological safety.

Finally, the research suggests the importance of aligning the referent (individual or team) with the level of analysis to maximize predictive validity. By matching survey measures and interventions to the appropriate referent and level of analysis, leaders can develop more targeted and effective strategies that better address the specific needs of their team members and the team as a whole.

References

Alper, S., Tjosvold, D., & Law, K. S. (1998). Interdependence and controversy in group decision making: Antecedents to effective self-managing teams. *Organizational Behavior and Human Decision Processes, 74*(1), 33-52. https://doi.org/10.1006/obhd.1998.2748

Baer, M., & Frese, M. (2003). Innovation is not enough: Climates for initiative and psychological safety, process innovations, and firm performance. *Journal of Organizational Behavior*, *24*(1), 45-68. https://doi.org/10.1002/job.179

Bandura, A. (1977). *Social learning theory.* Prentice Hall.

Bandura, A. (1997). *Self-efficacy: The exercise of control.* Freeman.

Bell, S. T. (2007). Deep-level composition variables as predictors of team performance: A meta-analysis. *Journal of Applied Psychology, 92*(3), 595-615.

Blau, P. (1964). *Exchange and power in social life.* Wiley.

Bliese, P.D. (1998). Group Size, ICC Values, and Group-Level Correlations: A Simulation. *Organizational Research Methods, 1*(4), 355-373.

Bliese, P. D. (2000). Within-group agreement, non-independence, and reliability: Implications for data aggregation and analysis. In K. J. Klein & S. W. J. Kozlowski (Eds.), *Multilevel theory, research, and methods in organizations: Foundations, extensions, and new directions* (pp. 349-381). Jossey-Bass.

Bliese, P. D., & Halverson, R. R. (1998). Group size and measures of group-level properties: An examination of eta-squared and ICC values. *Journal of Management, 24*(2), 157-172. https://doi.org/10.1016/S0149-2063(99)80058-0

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), 151-158. https://doi.org/10.1037/a0024200

Byrne, B. M. (2012). Structural equation modeling with Mplus: Basic concepts, applications, and programming. Taylor Francis/Routledge.

Chan, D. (1998). Functional relations among constructs in the same content domain at different levels of analysis: A typology of composition models. *Journal of Applied Psychology*, *83*(2), 234–246. https://doi.org/10.1037/0021-9010.83.2.234

Chang, C., Ferris, L., Johnson, R. E., & Tan, J. A. (2012). Core self-evaluations: A review and evaluation of the literature. *Journal of Management, 38*(1), 81-128. https://doi.org/10.1177/0149206311419661

Chiocchio, F. (2015). Defining project teams: A review of conceptual underpinnings. In F. Chiocchio, E. K. Kelloway, & B. Hobbs (Eds.), *The psychology and management of project teams* (pp. 40-73). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199861378.003.0003

Collins, C. J., & Smith, K. G. (2006). Knowledge exchange and combination: The role of human resource practices in the performance of high-technology firms. *Academy of Management Journal*, *49*(3), 544-560. https://doi.org/10.5465/amj.2006.21794671

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. (2012). *Teaming: How organizations learn, innovate, and compete in the knowledge economy*. Jossey-Bass.

Edmondson, A. C., & Bransby, D. P. (2023). Psychological safety comes of age: Observed themes in an established literature. *Annual Review of Organizational Psychology and Organizational Behavior*, *10*(1), 55-78. https://doi.org/10.1146/annurev-orgpsych-120920-055217

Edmondson, A. C., & Lei, Z. (2014). Psychological safety: The history, renaissance, and future of an interpersonal construct. *Annual Review of Organizational Psychology and Organizational Behavior*, *1*(1), 23-43. https://doi.org/10.1146/annurev-orgpsych-031413-091305

Edmondson, A. C., & Mogelof, J.P. (2005). Explaining psychological safety in innovation teams. In L. Thompson & H. Choi, (Eds.) *Creativity and Innovation in Organizations* (pp. 109–36). Lawrence Erlbaum Associates.

Elliot, A.J., & Thrash, T. M. (2002). Approach-avoidance motivation in personality: Approach and avoidance temperaments and goals. *Journal of Personality and Social Psychology, 82*(5), 804-818. https://doi.org/10.1037/0022-3514.82.5.804

Emerson, R. M. (1976). Social exchange theory. *Annual Review of Sociology, 2*(1), 335-362. https://www.jstor.org/stable/2946096

Ene, M. C. (2020). Investigating accuracy of model fit indices in multilevel confirmatory factor analysis [Doctoral dissertation, University of South Carolina]. Scholar Commons. https://scholarcommons.sc.edu/cgi/viewcontent.cgi?article=6919&context=etd

Fast, N. J., Burris, E. R., & Bartel, C. A. (2014). Managing to stay in the dark: Managerial self-efficacy, ego defensiveness, and the aversion to employee voice. *Academy of Management Journal, 57*(4), 1013-1034. https://doi.org/10.5465/amj.2012.0393

Ferris, D.L., Rosen, C.C., Johnson, R.E., Brown, D.J., Risavy, S., & Heller, D. (2011). Approach or avoidance (or both?): Integrating core self-evaluations within an approach/avoidance framework. *Personnel Psychology, 64*(1), 137-161. https://doi.org/10.1111/j.1744-6570.2010.01204.x

Frazier, M. L., Fainshmidt, S., Klinger, R. L., Pezeshkan, A., & Vracheva, V. (2017). Psychological safety: A meta-analytic review and extension. *Personnel Psychology, 70*(1), 113-165. https://doi.org/10.1111/peps.12183

Glick, W. H. (1985). Conceptualizing and measuring organizational and psychological climate: Pitfalls in multilevel research. *Academy of Management Review, 10*(3), 601-616. http://doi.org/10.5465/AMR.1985.4279045

Google. (n. d.). Identify dynamics of effective teams. Retrieved from https://rework.withgoogle.com/guides/understanding-team-effectiveness/steps/identify-dynamics-of-effective-teams/

Greenbaum, R. L., Bonner, J. M., Mawritz, M. B., Butts, M. M., & Smith, M. B. (2020). It is all about the bottom line: Group bottom-line mentality, psychological safety, and group creativity. *Journal of Organizational Behavior, 41*(6), 503-517. https://doi.org/10.1002/job.2445

Hackman, J. R. (2012). From causes to conditions in group research. *Journal of Organizational Behavior*, *33*(3), 428-444. https://doi.org/10.1002/job.1774

Halbesleben, J. R., & Buckley, M. R. (2004). Burnout in organizational life. *Journal of Management*, *30*(6), 859-879. https://doi.org/10.1016/j.jm.2004.06.004

Haynie, J. J. (2012). Core-self evaluations and team performance: The role of team-member exchange. *Small Group Research*, *43*(3), 315-329. https://doi.org/10.1177/1046496411428357

Hobfoll, S. E. (1989). Conservation of resources: A new attempt at conceptualizing stress. *American Psychologist, 44*(3), 513-524. https://doi.org/10.1037/0003-066X.44.3.513

Hollenbeck, J. R., DeRue, D. S., & Guzzo, R. (2004). Bridging the gap between I/O research and HR practice: Improving team composition, team training, and team task design. *Human Resource Management*, *43*(4), 353-366. https://doi.org/10.1002/hrm.20029

Hu, J., Erdogan, B., Jiang, K., Bauer, T. N., & Liu, S. (2018). Leader humility and team creativity: The role of team information sharing, psychological safety, and power distance. *Journal of Applied Psychology*, *103*(3), 313-323. https://doi.org/10.1037/apl0000277

IDEO. (2019). *This is the way Google & IDEO foster creativity.* Retrieved from ideou.com/blogs/inspiration/how-google-fosters-creativity-innovation.

James, L. R. (1982). Aggregation bias in estimates of perceptual agreement. *Journal of Applied Psychology*, *67*(2), 219-229. https://doi.org/10.1037/0021-9010.67.2.219

Judge, T.A. (2009). Core self-evaluations and work success. *Current Directions in Psychological Science, 18*(1), 58-62. https://doi.org/10.1111/j.1467-8721.2009.01606.x

Judge, T.A., Erez, A., Bono, J.E., & Thoresen, C.J. (2003). The core self-evaluation scale: Development of a measure. *Personnel Psychology*, *56*(2), 303-331. https://doi.org/10.1111/j.1744-6570.2003.tb00152.x

Judge, T. A., Locke, E. A., & Durham, C. C. (1997). The dispositional causes of job satisfaction: A core evaluations approach. *Research in Organizational Behavior, 19*, 151-188.

Kahn, W. A. (1990). Psychological conditions of personal engagement and disengagement at work. *Academy of Management Journal, 33*(4), 692-724. https://doi.org/10.5465/256287

Kelloway, K. E. (2015). *Using Mplus for structural equation modeling: A researcher's guide*. SAGE Publications, Inc. https://doi.org/10.4135/9781483381664

Kim, S., Lee, H., & Connerton, T. P. (2020). How psychological safety affects team performance: mediating role of efficacy and learning behavior. *Frontiers in Psychology, 11*, 1581.

Klein, K., Dansereau, F., & Hall, R. (1994). Levels issues in theory development, data collection, and analysis. *Academy of Management Review, 19*(2), 195-229. https://doi.org/10.5465/amr.1994.9410210745

Koopmann, J., Lanaj, K., Wang, M., Zhou, L., & Shi, J. (2016). Nonlinear effects of team tenure on team psychological safety climate and climate strength: Implications for average team member performance. *Journal of Applied Psychology, 101*(7), 940-957. https://doi.org/10.1037/apl0000097

Kostopoulos K. C., & Bozionelos, N. (2011). Team exploratory and exploitative learning: Psychological safety, task conflict, and team performance. *Group & Organization Management, 36*(3), 385-415. https://doi.org/10.1177/1059601111405985

Kozlowski, S. W. J., & Klein, K. J. (2000). A multilevel approach to theory and research in organizations: Contextual, temporal, and emergent processes. In K. J. Klein & S. W. J. Kozlowski (Eds.), Multilevel theory, research and methods in organizations: Foundations, extensions, and new directions (pp. 3-90). Jossey-Bass.

Lee, K., & Allen, N. J. (2002). Organizational citizenship behavior and workplace deviance: The role of affect and cognitions. *Journal of Applied Psychology, 87*(6), 1316-1326.

Liang, J., Farh, C. I., & Farh, J. L. (2012). Psychological antecedents of promotive and prohibitive voice: A two-wave examination. *Academy of Management Journal, 55*(1), 71-92.

Maloney, M. M., Johnson, S. G., & Zellmer-Bruhn, M. E. (2010). Assessing group-level constructs under missing data conditions. *Small Group Research, 41*(3), 281–307. https://doi.org/10.1177/1046496410363829

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. Academy of Management Review, 20(3), 709-734.

Morgeson, F. P., & Hofmann, D. A. (1999). The structure and function of collective constructs: Implications for multilevel research and theory development. *Academy of Management Review, 24*(2), 249-265. https://doi.org/10.5465/amr.1999.1893935

Muthén, L. K., & Muthén, B. O. (2012). Mplus user's guide (7th ed.). Los Angeles, CA: Muthén & Muthén.

Muthén, L. K., & Muthén, B. O. (2021). Mplus 8.6. [Computer program]. Los Angeles, CA: Muthén & Muthén.

Newman, A., Donohue, R., & Eva, N. (2017). Psychological safety: A systematic review of the literature. *Human Resource Management Review, 27*(3), 521-535. https://doi.org/10.1016/j.hrmr.2017.01.001

O'Neill, T. A., McLarnon, M. J. W., Hoffart, G. C., Woodley, H. J., & Allen, N. J. (2018). The structure and function of team conflict profiles. *Journal of Management, 44*(2), 811–836. http://doi.org/10.1177/0149206315581662

Organ, D.W. (1988). *OCB: The good soldier syndrome.* Lexington Books.

Ortega, A., Sánchez-Manzanares, M., Gil, F., & Rico, R. (2010). Team learning and effectiveness in virtual project teams: The role of beliefs about interpersonal context. *The Spanish Journal of Psychology, 13*(1), 267-276. https://doi.org/10.1017/S113874160000384X

Raudenbush, S. W., & Bryk, A. S. (2002). Hierarchical linear models: Applications and data analysis methods (2nd ed.). Thousand Oaks.

Rousseau, D. M. (1985). Issues of level in organizational research: Multi-level and cross-level perspectives. *Research in Organizational Behavior, 7*, 1-37.

Rousseau, D., Sitkin, S., Burt, R., & Camerer, C. (1998). Not so different after all: A cross-discipline view of trust. *Academy of Management Review, 23*(3), 393-404. https://doi.org/10.5465/amr.1998.926617

Rozovsky, J. (2015). *The five keys to a successful Google team.* Retrieved from https://rework.withgoogle.com/blog/five-keys-to-a-successful-google-team/

Sims, D. E., Salas, E., & Burke, C. S. (2005). Promoting effective team performance through training. In S. A. Wheelan (Ed.), *The handbook of group research and practice* (pp. 407-425). SAGE Publications.

Schaubroeck J., Lam S. S. K., & Peng, A.C.Y. (2011). Cognition-based and affect-based trust as mediators of leader behavior influences on team performance. *Journal of Applied Psychology, 96*(4), 863–871. https://doi.org/10.1037/a0022625

Schein, E. H., & Bennis, W. G. (1965). *Personal and organizational change through group methods: The laboratory approach.* Wiley.

Schmidt, J. A., Ogunfowora, B., & Bourdage, J. S. (2012). No person is an island: The effects of group characteristics on individual trait expression. *Journal of Organizational Behavior, 33*(7), 925-945. https://doi.org/10.1002/job.781

Seyda, B. A. S., & Tabancali, E. (2020). Correlations between teachers' personality, psychological safety perception and teacher voice. *Eurasian Journal of Educational Research, 20*(85), 185-204. https://orcid.org/0000-0001- 7536-2696

Siemsen, E., Roth, A. V., Balasubramanian, S., & Anand, G. (2009). The influence of psychological safety and confidence in knowledge on employee knowledge sharing. *Manufacturing & Service Operations Management, 11*(3), 429-447. https://doi.org/10.1287/msom.1080.0233

Triplett, S. M., & Loh, J. M. (2018). The moderating role of trust in the relationship between work locus of control and psychological safety in organisational work teams. *Australian Journal of Psychology, 70*(1), 76-84. https://doi.org/10.1111/ajpy.12168

Van Dyne, L., & LePine, J. A. (1998). Helping and voice extra-role behaviors: Evidence of construct and predictive validity. *Academy of Management Journal, 41(1),* 108-119. https://doi.org/10.5465/256902

Wallace, J. C., Edwards, B. D., Paul, J., Burke, M., Christian, M., & Eissa, G. (2016). Change the referent? A meta-analytic investigation of direct and referent-shift consensus models for organizational climate. *Journal of Management, 42*(4), 838-861. https://doi.org/10.1177/0149206313484520

Woodley, H. J. R. (2017). That's not fair! Examining individual differences in perceptions of fairness. *Electronic Thesis and Dissertation Repository. 4470.* https://ir.lib.uwo.ca/etd/4470

Appendices

Appendix A: Core Self Evaluations Measure

Please circle the number that indicates the degree to which you agree or disagree with each statement.

1. I am confident I get the success I deserve in life.
2. Sometimes I feel depressed.
3. When I try, I generally succeed.
4. Sometimes when I fail I feel worthless.
5. I complete tasks successfully.
6. Sometimes, I do not feel in control of my work.
7. Overall, I am satisfied with myself.
8. I am filled with doubts about my competence.
9. I determine what will happen in my life.
10. I do not feel in control of my success in my academic career.
11. I am capable of coping with most of my problems.
12. There are times when things look pretty bleak and hopeless to me.

Appendix B: Core Group Evaluations Measure

*Please **circle** the number that indicates the degree to which you agree or disagree with each statement.*

1. I am confident that my group gets the success it deserves.
2. When the members of my group try very hard and work together, we generally succeed.
3. When the group does not perform well, it is not respected by the rest of the team.
4. My group completes its assigned tasks successfully.
5. Sometimes my group is not in control of achieving its goals or helping the team to succeed.
6. Overall, the rest of the team is satisfied with my group.
7. I am filled with doubts about my group's ability.
8. My group determines the success it receives.
9. The members of my group do not feel that our group is in control of attaining success or achieving group goals.
10. My group is capable of coping with most of its problems.

Appendix C: Team Member Behaviour Measures

*Please rate your group members (but not yourself) on the following items. Fill his or her **first name** in the space below.*

To what extent does _____ (group member's first name) …

Task Performance (Van Dyne & LePine, 1998)

1. …complete work on time?
2. …successfully perform assigned tasks?
3. …produce quality work that meets performance expectations?

Team Citizenship Behaviours (Lee & Allen, 2002)

1. …help others who have been absent?
2. …willingly give time to help others who have project-related problems?
3. …assist others with their duties?

Counterproductive Behaviour (Woodley, 2017)

1. …engage in activities that derail the team's progress on the project?
2. …miss team meetings for unnecessary reasons?
3. …distract team members during team meetings?
4. …treat team members with disrespect?

Appendix D: Perceived Team Performance Measure

Perceived Team Performance (Woodley, 2017)

*Please **circle** the number that best represents how accurately each statement describes your team.*

1. Generally speaking, team members are very satisfied with their work.
2. Team members feel a strong commitment to their work.
3. Team members feel highly committed to the goals of their work.
4. The way we manage our work inspires us to better performance.
5. All things considered, the team is highly pleased with the way it manages its work.

Appendix E: Ethics Approval Letter

Department of Psychology The University of Western Ontario
Room 7418 Social Sciences Centre,
London, ON, Canada N6A 5C1
Telephone: (519) 661-2067Fax: (519) 661-3961

Use of Human Subjects - Ethics Approval Notice

Review Number	**13 07 09**	**Approval Date**	**13 07 24**
Principal Investigator	**Natalie Allen**	**End Date**	**14 04 30**
Protocol Title	**Understanding engineering project teams 2113-14 (Part 1)**		
Sponsor	**n/a**		

This is to notify you that The University of Western Ontario Department of Psychology Research Ethics Board (PREB) has granted expedited ethics approval to the above named research study on the date noted above.

The PREB is a sub-REB of The University of Western Ontario's Research Ethics Board for Non-Medical Research Involving Human Subjects (NMREB) which is organized and operates according to the Tri-Council Policy Statement and the applicable laws and regulations of Ontario. (See Office of Research Ethics web site: http://www.uwo.ca/research/ethics/)

This approval shall remain valid until end date noted above assuming timely and acceptable responses to the University's periodic requests for surveillance and monitoring information.

During the course of the research, no deviations from, or changes to, the protocol or consent form may be initiated without prior written approval from the PREB except when necessary to eliminate immediate hazards to the subject or when the change(s) involve only logistical or administrative aspects of the study (e.g. change of research assistant, telephone number etc). Subjects must receive a copy of the information/consent documentation.

Investigators must promptly also report to the PREB:
a) changes increasing the risk to the participant(s) and/or affecting significantly the conduct of the study;
b) all adverse and unexpected experiences or events that are both serious and unexpected;
c) new information that may adversely affect the safety of the subjects or the conduct of the study.

If these changes/adverse events require a change to the information/consent documentation, and/or recruitment advertisement, the newly revised information/consent documentation, and/or advertisement, must be submitted to the PREB for approval.

Members of the PREB who are named as investigators in research studies, or declare a conflict of interest, do not participate in discussion related to, nor vote on, such studies when they are presented to the PREB.

[redacted]

Clive Seligman Ph.D.

Chair, Psychology Expedited Research Ethics Board (PREB)

The other members of the 2012-2013 PREB are: Mike Atkinson (Introductory Psychology Coordinator), Rick Goffin, Riley Hinson Albert Katz (Department Chair), Steve Lupker, and Adam Piraino (Graduate Student Representative)

CC: UWO Office of Research Ethics

This is an official document. Please retain the original in your files

Curriculum Vitae

EDUCATION

Doctor of Philosophy (PhD), Industrial/Organizational Psychology 2024
Western University

Master of Science (MSc), Industrial/Organizational Psychology 2012
Western University

Honours Bachelor of Science (HBSc), High Distinction 2010
University of Toronto Mississauga

International Student Exchange Program 2008
Seoul National University

Credential of Readiness: The Language of Business 2015
Harvard Business School Online

TEACHING-RELATED EXPERIENCE

Course Instruction Positions
Western University

Psychology at Work	2014, 2016
Statistics for Management and Organizational Studies	2014
Introduction to Industrial/Organizational Psychology	2013

Teaching Assistantships
Western University

Social Psychology	2015
Teams and Workgroups in Organizations	2014 – 2015
Research Design	2013 – 2014
Statistics for Psychology	2011 – 2013
Introduction to Management and Organizational Studies	2010 – 2011

OTHER WORK EXPERIENCE

Social Sciences Research Analyst November 2015 – Present
Human Resources Systems Group

Author March 2022 – December 2023
Penguin Random House

Talent Management Consultant October 2015 – April 2017
Carswell Partners

HONOURS AND AWARDS

Outstanding Research Contributions Award 2018
Western University

Doctoral Fellowship 2013
Social Sciences and Humanities Research Council of Canada

Outstanding Master's Thesis Award of Excellence 2013
Canadian Psychological Association

Graduate Teaching Assistant Award 2012
Society of Graduate Students

Faculty of Social Science Alumni Award 2012
Western University

Joseph-Armand Bombardier Canada Graduate Scholarship 2010
Social Sciences and Humanities Research Council of Canada

ACADEMIC PUBLICATIONS

Tremblay, I., Lee, H., Chiocchio, F., & Meyer, J.P. (2015). Identification and commitment in project teams. In Chiocchio, F., Kelloway, E.K., Hobbs, B. (Eds.), *The Psychology and Management of Project Teams: An Interdisciplinary View*, Oxford University Press.

Girz, L., Polivy, J., Herman, C.P., & Lee, H. (2012). The effects of calorie information on food selection and intake. *International Journal of Obesity, 36,* 1340-1345.

CONFERENCES

Allen, N.J., Stanley, D.J., Cameron, K., McMenamin, J., Ouslis, N., Lee, H., & Woodley, H. (2017). *Group performance: A 10-year bibliometric review of conceptualizations and assessment.* Presented at the annual meeting of the Interdisciplinary Network for Group Research (InGroup), St. Louis, MO.

McMenamin, J., Lee, H., & Daljeet, K. (2017). *Exploring impostor phenomenon personality correlates with the HEXACO.* Poster presented at the Canadian Psychological Association 78th Annual Convention, Toronto, Ontario.

Cameron, K. & Lee, H. (2014). *Examining the relationship between goal orientation and psychological safety.* Poster presented at the Canadian Psychological Association 75th Annual Convention, Vancouver, British Columbia.

Lee, H., Allen, N., Cameron, K., & Woodley, H. (2013). *Rethinking the measurement of psychological safety: referent use validity and the role of core self-evaluations.* Talk presented at the Administrative Sciences Association of Canada 2013 Convention, Calgary, Alberta.

Lee, H. (2013). *Predicting team conflict with the five-factor personality model.* Poster presented at the Canadian Psychological Association 74th Annual Convention, Quebec City, Quebec.

Lee, H. (2010, April). *The effects of calorie information on food selection and consumption.* Talk presented at the 40th Annual Ontario Psychology Undergraduate Thesis Conference, Windsor, Ontario.