Observed Psychological Safety in Agile Teams: A Comparison Across Cultures and Conflict Management Effectiveness

Author: Madelief Veltkamp
University of Twente
P.O. Box 217, 7500AE Enschede
The Netherlands

ABSTRACT

To remain competitive many organisations have shifted from traditional to Agile methodologies. Key characteristics of Agile organisations, such as speed, adaptability and self-management, allow them to thrive in unpredictable and fast-changing environments. However, for Agile teams to function effectively, a strong foundation of psychological safety is essential. Most research on psychological safety relies on self-reported measures, limiting objectivity. Addressing this gap, this study adopts a mixed-methods approach, combining survey data and video-recorded team meetings to examine verbal and non-verbal behavioural indicators of psychological safety in mono- and multicultural Agile teams with varying levels of conflict management effectiveness. Teams with higher conflict management effectiveness displayed more frequent and longer-lasting behaviours indicative of psychological safety, while exhibiting fewer behaviours associated with low psychological safety. This suggests a relationship between the level of psychological safety and conflict management effectiveness. Multicultural teams demonstrated higher levels of silence behaviours, and fewer voice and familiarity behaviours compared to monocultural teams, suggesting that cultural diversity may create communication barriers affecting psychological safety. Although not statistically significant, these patterns provide meaningful insights. This study contributes to the literature by offering an objective, behaviour-based assessment of psychological safety and presents practical implications for Agile coaches and organisations aiming to foster psychological safety, especially in culturally diverse teams. Future research is encouraged to expand the sample size, include diverse industries and geographical contexts, and explore the directionality of the relationship between psychological safety and conflict management effectiveness.

Graduation Committee members:
Dr. L. Carminati
Dr. R. Rajah

Keywords
Agile, Psychological Safety, Conflict Management, Mono- & Multicultural Teams

During the preparation of this work, the author used ChatGPT to improve spelling and grammar, and Litmaps to identify relevant literature. After using this tool/service, the author reviewed and edited the content as needed and takes full responsibility for the content of the work.

1. INTRODUCTION

One of the biggest challenges faced by organisations today is remaining competitive in a business environment characterised by rapid technological evolution, shifting consumer preferences, and increasing global competition (del Pilar Barrera et al., 2025). To succeed, it is essential to adapt quickly and navigate uncertainty effectively (del Pilar Barrera et al., 2025). To overcome these challenges, numerous organisations have transitioned from traditional, predominantly sequential methodologies such as the Waterfall model, where each phase must be completed before moving on to the next (Stanley et al., 2020), to more flexible and responsive approaches like Agile methodologies (Almeida, 2017). Agile is a project management approach characterised by breaking down work into short iterative cycles, incorporating frequent project evaluations, and allowing for continuous adjustment of plans during execution as necessary (Stanley et al., 2020). Indeed, the key characteristics of Agile organisations, like speed and flexibility, allow them to thrive in environments that are highly unpredictable and rapidly changing (Naslund & Kale, 2020). At the team level, this Agility is reflected in self-managing, cross-functional groups that often operate under a shared leadership model (Stanley et al., 2020). As a result, Agile development depends on team members' skills and their ability to work effectively together. In turn, research has shown that autonomy can promote a range of positive outcomes, including greater willingness among team members to experiment and increased openness to exploring solutions, which ultimately leads to higher levels of psychological safety (Buvik & Tkalich, 2022).

To enable Agile team members to thrive, it is essential to foster a team climate where individuals feel safe to contribute and challenge ideas (Buvik & Tkalich, 2022). A crucial role in creating such an environment is played by psychological safety, which refers to a shared belief that the team is safe for interpersonal risk-taking (Edmondson, 1999). In Agile settings, where autonomy, collaboration, and frequent feedback are the norm (Stanley et al., 2020), psychological safety becomes a foundational element for innovation (Andersson et al., 2020), conflict resolution (Edmondson, 1999), team reflexivity and team performance (Buvik & Tkalich, 2022). Moreover, research found that high levels of psychological safety encourage team members to actively participate in Agile meetings, share ideas, and take initiative, making it a key contributor to the success of Agile teams (Hennel & Rosenkranz, 2021).

Although psychological safety has been extensively studied in the literature, research within Agile contexts is more limited. Moreover, psychological safety research has mainly been conducted through the use of self-perceived assessment, meaning that measures of psychological safety typically rely on the use of surveys (Edmondson & Bransby, 2023). Relying only on survey instruments is limited by self-report bias (O'Donovan et al., 2020). Therefore, recent studies have called for implementing more objective assessments, such as video observations, to study individual behaviours in management studies (Zhao et al., 2019; O'Donovan et al., 2020). Such observed measures can complement self-perceived ones by offering alternative insights into psychological safety that are less biased than surveys and capture aspects the teams themselves may not be fully aware of (O'Donovan et al., 2020).

Furthermore, the strong emphasis on collaboration in Agile teams combined with the absence of a formal leader makes them more prone to conflict compared to traditional teams (Crawford et al., 2014; Gren & Lenberg, 2018; Niederman et al., 2018). According to Hennel & Rosenkranz (2021), psychological safety may explain when and why conflict can be beneficial to Agile teams. Whether conflict ultimately benefits or weakens a team largely depends on how effectively it is managed (DeChurch & Marks, 2001). Indeed, higher psychological safety can enhance conflict management (Edmondson, 1999; Faust, 2023). Since psychological safety plays a significant role in team dynamics by creating a favourable environment for speaking up, sharing ideas and asking questions (Edmondson & Bransby, 2023), it supports open and honest communication, which is an essential component of effective conflict management (Tekleab et al., 2009). However, despite substantial work on conflict management, there is a notable gap regarding the interplay with psychological safety (Faust, 2023).

In addition to Agile teams, organisations increasingly rely on multicultural teams, to benefit from diverse perspectives and approaches (Marquard & Horvath, 2001). Although multicultural teams have numerous benefits, they can also be more challenging to manage due to varying approaches to teamwork (Marquard & Horvath, 2001). It is important to have attention for these collaboration challenges, as collaboration is a cornerstone of the Agile methodology (Beck et al., 2001). Psychological safety could potentially play a moderating role in the relationship between diversity and team outcomes (Edmondson and Roloff, 2008) and thus help to overcome diversity related barriers. Additionally psychological safety might be manifested differently between mono- and multicultural teams, as cultural norms shape both attitudes and behaviours by defining what is acceptable, expected, or discouraged (Groysberg et al., 2018). Hence, psychological safety emerged as a vital ingredient in addressing teamwork challenges, including managing conflicts and navigating the complexities introduced by multicultural team dynamics. A study by Thorgren & Caiman (2019) explored psychological safety between Agile teams with different (mono) cultures, however recent research has highlighted the need for a deeper understanding of different cultural influences in an Agile team (Welsch et al., 2024).

Therefore, this thesis contributes to the literature on psychological safety by examining the differences in *observed* psychological safety in relation to conflict management effectiveness and cultural composition within an Agile context.

Hence, this thesis answers the following research question:

"How does team members' observed psychological safety differ between teams with effective versus ineffective conflict management, and between mono- and multicultural teams, within an Agile context?"

To address this research question, this study adopts a mixed-methods design, integrating survey data with video-recorded observations, and offers three key theoretical contributions. First, this study extends current knowledge by employing a novel and more objective

approach to examining psychological safety through the observation and comparison of actual psychological safety related behaviours during retrospective meetings of Agile teams. Second, this thesis provides insight into the relationship between psychological safety and the effectiveness of conflict management within Agile teams. Third, it contributes to the understanding of how psychological safety differs between monocultural and multicultural Agile teams. This thesis also offers practical implications for Agile team development and management. By identifying specific, observable behaviours associated with psychological safety, it equips Agile coaches and organisations with concrete indicators to monitor and support team dynamics more effectively. Moreover, by linking these behaviours to perceived conflict management effectiveness and examining the influence of cultural context, the study provides actionable guidance for enhancing both psychological safety and conflict management practices in diverse team settings.

In the remainder of this thesis, the theoretical framework underpinning the research question will be outlined, followed by a detailed explanation of the methodology employed. Subsequently, the findings will be presented and discussed, highlighting both the strengths and limitations of the research. Finally, the thesis will conclude by addressing the research question and providing recommendations for future research.

2.THEORETICAL FRAMEWORK

This section provides an overview of the fundamental principles of the Agile methodology and its key outcomes, which form the contextual foundation for the teams examined in this study. It is followed by a definition of psychological safety, an exploration of its significance within Agile team dynamics, and a method of assessing it through verbal and non-verbal behaviours. The relationship between psychological safety and conflict management is explored along with the potential differences between mono- and multicultural teams.

2.1 The Agile methodology

Agile gained widespread recognition as a methodology with the publication of the Agile Manifesto in 2001 (Dingsøyr et al., 2012). The Agile methodology emerged as a response to traditional software development approaches, such as the Waterfall model, where each phase must be completed before moving on to the next (Stanley et al., 2020) as they were increasingly viewed as inflexible and inefficient in rapidly changing environments (Abbas et al., 2008). In contrast, Agile methodologies are grounded on four core values that emphasise a more flexible and human-centred approach: (1) individuals and interactions over processes and tools, (2) working software over comprehensive documentation, (3) customer collaboration over contract negotiation, (4) and responding to change over following a plan (Beck et al., 2001).

While these shared values form the foundation of all Agile methods, the term “Agile” remains difficult to define precisely, as it functions as an umbrella concept encompassing a variety of well-defined methods that differ in practical application (Abbas et al., 2008), such as Scrum and Extreme Programming (Dingsøyr et al., 2012). This thesis adopts the definition of Agile as a method that is continually ready “to rapidly or inherently create change, proactively or reactively embrace change, and learn from change, while contributing to perceived customer value (economy, quality and simplicity), through its collective components and relationships with its environment” (Conboy, 2009, p. 340).

2.1.1 Agile Structure

Agile teams, also known as squads, are both self-managing and cross-functional (Stanley et al., 2020), meaning they consist of members with diverse skill sets who collectively take ownership of the planning and execution of their work. Operating under a shared leadership model, all team members are responsible for achieving outcomes (Magpili & Pazos, 2018).

Work in Agile teams is structured around short, iterative development cycles, also referred to as sprints (Stanley et al., 2020). Regular project evaluations enable ongoing adjustments to plans during execution when needed (Stanley et al., 2020). Sprints often include three key meetings: (1) The sprint planning, (2) refinement meeting, and (3) retrospective meeting (Stanley et al., 2020). During the sprint planning, the team defines a clear goal for the sprint. Each day that follows typically includes a brief stand-up meeting (5-15 min), where team members share updates on their progress, what they will do today and communicate any challenges they face (Stanley et al., 2020). The purpose of this is that challenges are identified and handled quickly. As the sprint progresses, a refinement meeting is held followed by the final retrospective meeting, where the team reflects on what went well, what could be improved and how to enhance performance in the next sprint (Andriyani et al., 2017).

2.1.2 Outcomes of Agile

The fundamental characteristics of Agile organisations, allow them to thrive in rapidly evolving and uncertain environments (Naslund & Kale, 2020). Therefore, Agile methodologies have expanded beyond their origins in software development and are now applied across various industries (Conforto et al., 2014). This broader adoption is attributed to several advantages. Research has shown that autonomy can promote a range of positive outcomes, including greater willingness among team members to experiment and increased openness to exploring solutions (Buvik & Tkalich, 2022). Which ultimately results in higher levels of psychological safety, enhanced employee engagement, as well as improvements in team performance (Peeters et al., 2022). The effect of Agile practices on employee engagement and performance was found to be partially mediated by a psychological safe climate (Peeters et al., 2022). Therefore, to support the effective implementation of Agile practices, fostering psychological safety could be highly beneficial. In fact, some practitioners have even claimed that Agile does not work without psychological safety (Alami et al., 2023).

2.2 Psychological safety

Psychological safety was first explored by pioneering organisational scholars in the 1960s (Edmondson & Lei, 2014). Edgar Schein and Warren Bennis introduced the concept in their book *Personal and Organizational Change Through Group Methods: The Laboratory Approach*, describing it as a context that “encourages provisional tries and which tolerates failure” (Shein & Bennis, 1965, p. 45). However, broader recognition emerged much later in 1999, primarily due to Amy Edmondson’s work. In her paper, she described

psychological safety as "a shared belief held by members of a team that the team is safe for interpersonal risk taking" (Edmondson, 1999, p. 2), which is also the definition of psychological safety that this thesis adopts. Her research showed relationships between team psychological safety, team learning, and team performance. Ever since the number and reach of studies on psychological safety have grown exponentially, especially research conducted in the healthcare delivery industry (Edmondson & Bransby, 2023).

2.2.1 Outcomes of Psychological Safety in Agile Teams

It has become clear that psychological safety offers numerous benefits for teams, including those working within Agile frameworks (Hennel & Rosenkranz, 2020). Psychological safety is positively related to team reflexivity, performance (Buvik & Tkalich, 2022) and innovation (Andersson et al., 2020), as it enables essential group processes such as learning behaviour, and conflict resolution by creating an environment that encourages speaking up (Edmondson, 1999). Research by Hennel and Rosenkranz (2020) identifies psychological safety as a critical factor in team members acceptance of Agile practices. Low psychological safety is linked to reduced participation in Agile meetings, as individuals are less likely to speak up, share ideas, or contribute input. In contrast, higher levels of psychological safety promote greater acceptance of Agile methods and more active engagement. When team members feel psychologically safe, they are more likely to help others, propose new ideas and offer valuable input, ultimately enhancing the effectiveness of Agile practices and their outcomes. Psychological safety promotes knowledge sharing (Rivera et al., 2021), especially when it involves high interpersonal risk, since research by Mura et al. (2016) showed that knowledge sharing related to exposing flaws or limitations, e.g., sharing mistakes or seeking feedback, was especially influenced by psychological safety. Psychological safety thus plays a significant role in team dynamics by creating a favourable environment for speaking up, sharing ideas, and asking questions (Edmondson & Bransby, 2023), which is fundamental for effective teamwork, communication, and collaboration. Additionally, psychological safety facilitates improved communication across professional boundaries (O'Leary, 2016), which is especially important in cross-functional Agile teams.

2.2.2 Observing Psychological Safety Through Verbal and Non-Verbal Behaviours

Psychological safety is an inherently interpersonal and behavioural phenomenon that is best understood through observable interactions between team members (O'Donovan et al., 2020. Therefore, examining psychological safety requires going beyond self-reported perceptions as they are limited by self-reported bias and don't capture aspects that participants themselves are not aware of (O'Donovan et al., 2020. Behaviour can be defined as "the internally coordinated responses (actions or inactions) of whole living organisms (individuals or groups) to internal and/or external stimuli, excluding responses more easily understood as developmental changes" (Levitis et al., 2009, p. 108). These internally coordinated responses include both verbal behaviours and physical movements (Bergner, 2010). When these behaviours reflect interpersonal risk-taking, e.g., correcting others or asking questions, this is a sign of psychological safety (Edmondson, 1999; O'Donovan et al, 2020). Conversely, the absence of such verbal behaviours, or overly cautious verbal behaviour, may indicate a lack of psychological safety (Jiang et al., 2019; O'Donovan et al, 2020). Non-verbal behaviours, while often less consciously controlled, can also signal the presence or absence of psychological safety. For example, facial expressions indicating fear or disengagement or closed body language (arms closed, lean backwards) can be a sign of low psychological safety, while active listening (keeping eye contact) can indicate high psychological safety (O'Donovan et al, 2020). Hence, this thesis explores psychological safety through observing verbal and non-verbal behaviours within Agile teams using video observations.

2.3 Conflict management

For Agile teams, both conflict and conflict resolution differ compared to traditional teams (Niederman et al., 2018). The likelihood of conflict in Agile teams is higher compared to traditional teams, because of the intensive collaboration inherent in Agile methodologies (Crawford et al., 2014; Gren & Lenberg, 2018) and the absence of a formal leader (Niederman et al., 2018). Conflict refers to a situation where one party believes that another party is acting against or harming its interests (Wall & Callister, 1995). A distinction can be made between task conflict, i.e., a disagreement between team members about the content of the tasks being performed, including differences in viewpoints, ideas, and opinions, and relationship conflict, i.e., interpersonal incompatibility among members, often characterised by tension, animosity, and annoyance (Jehn, 1995). Conflict can have both positive and negative influences on a team, largely depended on the effectiveness of its management (DeChurch & Marks, 2001). Furthermore, research found that the presence of interpersonal conflict was negatively related to central Agile practices such as iterative development and customer access, making it more difficult for teams to be Agile (Gren, 2017). If a team wants to reach a productive and autonomous stage, it should be able to manage internal conflicts and disagreements efficiently (Gren & Lenberg, 2018). Therefore, effective conflict management is crucial for Agile teams. Previous research has approached conflict management from two main perspectives: as a team-level process involving the degree to which teams actively manage conflict, or by focussing on individual styles of conflict management behaviour displayed by team members (Tekleab, 2009). This thesis follows the approach adopted by Tekleab et al. (2009), placing emphasis on conflict management as a process, specifically on whether teams actively engage in open discussions and are equipped to address conflict when it emerges. Open communication allows teams to successfully resolve their disagreements, which fosters trust and leads to increased team cohesion and overall team effectiveness (Tekleab et al., 2009).

Higher levels of psychological safety can enhance conflict management (Edmondson, 1999; Faust, 2023). By creating a favourable environment for speaking up, sharing ideas and asking questions (Edmondson & Bransby, 2023), psychological safety plays an important role as it supports open communication (Ito et al., 2022), which is an essential component of effective conflict management (Tekleab et al., 2009). Controversially, a lack of

psychological safety is related to knowledge hiding, holding back ideas and observations (Jiang et al., 2019) This makes communication more difficult and, thus in turn negatively influences conflict management as issues are not addressed when they emerge. Research by Sherf et al. (2021) even suggested that psychological safety may be more strongly associated with silence behaviour (negatively) than with voice behaviour (positively) as a lack of psychological safety can serve as an environmental cue that triggers the behavioural inhibition system, leading individuals to avoid potential self-harm through silence.

2.4 Cultural team composition

In today's globalised world, organisations are increasingly relying on multicultural teams to leverage the diverse experiences, perspectives and approaches of individuals from different cultural backgrounds. (Marquardt & Horvath, 2001). Culture refers to a way of thinking, acting, and living that is shared by members of a group and passed down from one generation to the next (Marquardt & Horvath, 2001). This thesis adopts the definition of multicultural teams as proposed by Tirmizi (2008), stating that a multicultural team exists of culturally diverse individuals who are interdependent, share responsibility for outcomes, identify as a cohesive unit within larger systems, and manage relationships across and beyond organisational boundaries. Cultural diversity is thus a variation in shared beliefs, norms and values within a group, that underpin their behaviours and shape how they interpret various aspects of their social world (Hui et al., 2017). Although cultural diversity has advantages, it can also be difficult to manage, as most people simply don't realise it's there (Marquardt & Horvath, 2001). Key cultural differences that affect teamwork include different communication styles, different ways for problem solving, decision making, and dealing with disagreements (Marquardt & Horvath, 2001). Furthermore, these cultural differences can increase the likelihood of friction, misunderstandings, and tension, which if not addressed, may quickly escalate into conflict, making multicultural teams more prone to it (Marquardt & Horvath, 2001). Addressing the greater collaboration challenges in multicultural teams is crucial, as collaboration is central to the Agile methodology (Beck et al., 2001). Edmondson and Roloff (2008) argue that psychological safety can play a moderating role in the relationship between diversity and team outcomes, such as collaboration, learning and performance and thus helps to overcome diversity related barriers. Additionally, since culture shapes the way people behave (Groysberg et al., 2018), psychological safety might differ between mono- and multicultural teams.

3. METHODOLOGY

3.1 Research Design

This thesis adopts a mixed-method research design, combining both quantitative and qualitative approaches to address the research question in a more comprehensive way (Creswell et al., 2003). Each method provides different types of information and has both distinct limitations and strengths (Creswell & Cresswell, 2017). Quantitative methods provide factual, reliable outcomes and generalizable patterns, while qualitative methods offer rich, detailed insights into individual perspectives (Steckler et al., 1992). By integrating both approaches, the study benefits from the complementary nature of the data sources by methodological triangulation, which enhances the reliability and validity of findings as it can neutralise biases and weaknesses of the methods separately (Creswell & Cresswell, 2017). This research utilises pre-existing data, comprising video recordings of real-time Agile team meetings and survey responses from the team members.

The use of video recordings in the qualitative component of this study helps address the limitations of self-reported data and aligns with the growing emphasis in literature on objective behavioural assessment (Zhao et al., 2019; O'Donovan et al., 2020). The video recordings were coded using the psychological safety codebook developed by the OBCC to identify both verbal and non-verbal behaviours associated with high or low levels of psychological safety. The coded videos were used to identify the frequency and duration of these behaviours across mono- and multicultural teams and across teams with high versus low conflict management effectiveness to investigate a potential relationship between cultural composition and psychological safety and between conflict management effectiveness and psychological safety. To categorise teams based on cultural composition, demographic data collected before the first meeting was used, in which participants reported their nationality.

Additionally, the survey captured data on the Agile team members' perceptions of conflict management effectiveness during the meetings, which allowed for a differentiation between effective and ineffective teams. Accordingly, this thesis adopts a mixed-method research design, beginning with a quantitative analysis to categorise the teams, and a qualitative analysis to identify verbal and non-verbal behaviours related to psychological safety, followed by a quantitative approach to investigate differences between mono- and multicultural teams and a potential relationship with conflict management effectiveness.

3.2 Data Collection

The data used in this research originates from a larger research project conducted by the Organisational Behaviour, Change Management and Consultancy Group (OBCC) of the University of Twente between 2018 and 2022. The project was carried out in collaboration with a major Dutch financial services company and focused on multiple Agile teams. These teams were video recorded during three key meetings over the course of their sprint, namely the sprint planning, refinement meeting and retrospective meeting. To preserve anonymity of participants, team members were assigned identification numbers, which were also visible on their bodies during the recordings. While the original dataset includes both virtual and in-person meetings, this thesis exclusively analyses the video recordings of face-to-face meetings, as the observation of psychological safety might differ between the two settings. To capture all the relevant perspectives, three cameras were set up around the meeting table. In addition to the video recordings, a series of surveys were conducted before or after each meeting to gather information on a variety of topics, including demographics and perceived effectiveness of conflict management.

3.3 Sample

The dataset analysed in this thesis consists of observations from four Agile teams operating within a major Dutch financial services company that has long embedded Agile

practices into its organisational structure. Participants in this study were drawn from various squads that function autonomously and represent a range of professional roles and individual backgrounds. The Agile teams selected differ along two key dimensions: cultural composition and conflict management effectiveness. Teams were first grouped by cultural composition as either monocultural or multicultural. Within each of these categories, teams were further differentiated by their conflict management effectiveness, as measured by average scores completed by team members through the surveys. This resulted in a sample of four teams: one monocultural team with high conflict management effectiveness, one monocultural team with low effectiveness, one multicultural team with high effectiveness and one multicultural team with low effectiveness.

Table 1. Perceived conflict management effectiveness scores for each selected team

Team	Conflict management effectiveness	SD	Team	Conflict management effectiveness	SD
Effective monocultural			*Effective multicultural*		
A	5.1	1.5	I	5.7	.63
Ineffective monocultural			*Ineffective multicultural*		
B	5	.9	II	4	1.6

A total of 29 individuals were observed across the four Agile teams. Team sizes ranged from 5 to 9 members, with an average of 7.25 participants per team. Among the 29 participants, 24 identified as male, 4 as female and 1 did not disclose gender information. In terms of nationality, 18 were Dutch and 10 were non-Dutch, including 7 individuals of Indian origin, 1 Belgian, 1 Spanish and 1 Hungarian, and 1 did not disclose information on nationality. The average age of all participants was 37.69, ranging between 22 and 65. The teams represented a diverse mix of professional expertise, with each team including at least two different professional backgrounds. Across all teams, areas such as finance IT, marketing, risk management, and data science were represented. Educational qualifications varied as well, including participants with HBO bachelor's degrees, University bachelor's and master's degrees, and PhDs, as well as alternative qualifications. For three of the teams, the retrospective meeting was observed. During this meeting teams should reflect on collaboration by discussing challenges, feelings, analysing previous action points and identifying reasons for issues (Andriyani et al., 2017), which makes it a critical context for psychological safety. For one team, the retrospective meeting was not available, therefore the second meeting in the sprint cycle was observed instead. In total, four team meetings were included in the analysis.

3.4 Measures

3.4.1 Psychological Safety

Psychological safety was assessed through a behavioural analysis of the video-recorded team meetings. The video recordings were coded using a codebook developed by the OBCC, which was based on the observational framework proposed by O'Donovan et al. (2020). Originally developed to complement survey methods in the context of healthcare teams, this framework enables the identification of behaviours that indicate varying levels of psychological safety. The OBCC codebook defined nine distinct categories of observable behaviours that are expected to either enhance or diminish team members perceived psychological safety within a team. Categories associated with higher perceived psychological safety included behaviours reflecting interpersonal risk-taking, for example: *Voice behaviours* (e.g. disagreeing, providing feedback or correcting others) and *Learning or improvement-oriented behaviours* (e.g. acknowledging as mistake or asking for help) (O'Donovan et al., 2020). Categories associated with lower perceived psychological safety included behaviours reflecting a lack of interpersonal risk-taking, for example: *Silence behaviours (non-verbal)* (e.g. facial expressions indicating fear or disengagement), and *Defensive voice behaviours* (e.g. denying faults and evading confrontation by focussing on the positives) (O'Donovan et al., 2020).

During the coding process, each observed behaviour was assigned to one of the nine categories. Only interactions that clearly reflected behaviour from the predefined behavioural categories were coded. No codes were assigned when team members did not display a single behaviour listed in the codebook. Observations included psychological safety behaviours directed at individual team members, subgroups, as well as those towards the team as a whole.

3.4.2 Mono- & Multi Cultural Teams

In accordance with the definition proposed by Tirmizi (2008), Agile teams were classified as multicultural if they included members form at least two different cultural backgrounds. Cultural background was primarily determined based on team members' self-reported nationality, which was collected through a survey conducted after the first team meeting. In one case where nationality data was not provided, the classification was based on the most fluent language indicated by team members. Given that this team reported a combination of *'Dutch'*, *'English'*, and *'Other'*, it was classified as multicultural. Among the teams that participated in face-to-face meetings, this classification process resulted in the identification of four monocultural and five multicultural teams. One of the multicultural teams included only two nationalities: Dutch and Indian. Although the Dutch member was absent from the meeting analysed, the team was still classified as multicultural. The decision was made because the Indian participants stated three different most fluent languages, suggesting a level of cultural heterogeneity within the group. As Panda and Gupta (2004) demonstrated, cultural diversity can exist within Indian national boundaries, thus, the team was decided to exhibit sufficient cultural diversity to meet the criteria for classification as multicultural.

3.4.3 Conflict Management

Perceived conflict management effectiveness was measured directly after the second meeting, which is also known as the sprint review. This was done by conducting surveys, based on four items adapted from Tekleab et al. (2009): (1) *Conflict is dealt with openly on this squad*, (2) *If conflict arises on this squad, the people involved in the conflict initiate steps to resolve the conflict immediately*, (3) *This squad knows what to do when conflicts between squad members arise*, (4) *This squad is able to avoid the negative aspects of conflict before they occur*. The team members rated their level of agreement with each statement using a seven-point Likert scale, ranging from

'*Strongly Disagree*' (1) to '*Strongly Agree'* (7). Based on these responses, teams were assigned a mean conflict management effectiveness score. For both the mono- and multicultural teams, the teams with the highest and lowest means were selected for analysis.

3.5 Data Analysis

3.5.1 Qualitative Analysis

A deductive thematic analysis was conducted to identify observable instances of psychological safety-related behaviour, based on the framework by Braun and Clarke (2006). Participant verbal and non-verbal behaviours were systematically compared to those specified in the OBCC psychological safety codebook by examining the video recordings and their transcripts. When a participant demonstrated a behaviour that aligned with one of the behaviours in the nine codebook categories, as interpreted by the researcher, it was assigned the corresponding code. For example, when a participant provided information, help, solutions, or corrected another participant, the behaviour was coded as *"Voice behaviour".* When a participant had a facial expression indicating fear or disengagement it was coded as *"Silence behaviour"* (O'Donovan et al., 2020). If a participant did not display a single behaviour listed in the codebook, no code was assigned. When a participant displayed behaviours belonging to two distinct codes, both codes were assigned. Behaviours were coded from their start to conclusion, allowing for the calculation of each behaviour's duration. This process was consistently applied to all participants throughout the meeting. All coding was conducted using Noldus Observer XT to organise and compare the behavioural data systematically.

3.5.2 Quantitative Analysis

After all the psychological safety-related behaviours were identified, both a frequency analysis and a duration analysis were conducted. The frequency analysis quantified the number of observed behaviours within each of the nine categories, while the duration analysis measured the total time these behaviours were displayed. The frequencies and durations of behaviours within each coded category were then added together to determine the total number and total duration of behaviours indicative of higher and lower levels of psychological safety. This was based on the codebook, which categorised four behavioural categories as reflecting higher psychological safety, four as reflecting lower psychological safety, and one as neutral (see Section 3.4.1). These metrics were then used to examine differences in psychological safety across teams with differing levels of conflict management effectiveness and cultural composition.

To enable comparison across teams, both frequency and duration measures were standardised. Given that the analysed meetings varied in team size and length, some teams naturally exhibited more and longer behaviours. Therefore, the frequency and duration of each behavioural category were divided by the total frequency and total duration of all observed behaviours in the entire meeting. This resulted in a percentage-based representation of each behavioural category, enabling consistent comparison across teams. To explore whether the differences in means for frequency and duration were statistically meaningful across cultural composition and conflict management effectiveness of teams, a statistical test was conducted using the Mann-Whitney U test. This non-parametric test was selected because it does not require assumptions of normality and its more appropriate for comparing differences between two independent groups with small sample sizes compared to parametric alternatives (Kim, 2023). In this study, the number of data points per group was too limited to conduct formal normality testing. For example, conducting a Shapiro-Wilk test to examine normality in R Studio requires a minimum of three observations per group. Therefore, to avoid making unsupported assumptions about the data distribution, the Mann-Whitney U test was decided to be the most suitable method.

4. RESULTS

This section outlines the key findings of the study, starting with a descriptive frequency analysis of observed behaviours, followed by a descriptive analysis of the duration of those behaviours, a comparative analysis, and concluding with an exploratory statistical interpretation.

4.1 Frequency Analysis

Table 2 below presents an overview of the standardised frequencies of observed behaviours indicative of higher levels of psychological safety e.g., *voice behaviours* and *collaboration behaviours*, lower levels of psychological safety e.g., *defensive voice behaviours* and *silence behaviours (non-verbal)*, and *neutral behaviours*, i.e. neutral task related behaviours, across the four observed teams. Each team representing a distinct combination of conflict management effectiveness (effective vs. ineffective) and cultural composition (monocultural vs. multicultural). As shown in Table 2, all the teams exhibit more behaviours indicative of higher psychological safety than behaviours indicative of lower psychological safety. Team A displayed the highest percentage of behaviours indicative of higher psychological safety (85.75%), which is interesting since it is the monocultural team with effective conflict management, followed by Team I (67.79%), Team II (63.24%), and Team B (62.24%). In contrast, behaviours indicative of lower psychological safety were most prominent in team II (36.76%) and Team B (34.10%), both of which were characterised by low conflict management effectiveness. For team I this score was slightly lower (34.10%), followed by Team A with the lowest score (11.29%). Neutral behaviours remained minimal across all teams, ranging from 0,00% to 3.66%. However, they were more prevalent in the monocultural teams than in the multicultural teams.

Looking at the specific behavioural categories, collaboration behaviours were most frequently observed across all teams, particularly in Team A (57.38%). Silence behaviours, associated with lower psychological safety, were especially pronounced in Team II (28.06%) and Team I (24.85%), which are both multicultural teams. While both Team A (9.99%) and Team B (13.29%) scored lower in this category. Defensive silence behaviours were notably high in Team B (20.04%) but nearly absent in other teams. Familiarity behaviours were slightly more frequent in both monocultural Team A (6.15%) and Team B (7.13%), compared to multicultural Team I (1.32%) and Team II (3.95%). This is also true for voice behaviours with the highest frequency in Team A (16.64%) and Team B (16.18%), followed by team I (12.65%) and Team II (12.52%). In contrast, unsupportive behaviours were observed more in Team I and Team II (1.62% and 3.29%) compared to Team A and

Table 2. Standardised frequencies of observed behaviours indicative of psychological safety

	Team A (Effective monocultural)	Team B (Ineffective monocultural)	Team I (Effective multicultural)	Team II (Ineffective multicultural)
Voice Behaviours	**16.64%**	**16.18%**	12.65%	12.52%
Defensive Voice Behaviours	0.36%	0.19%	3.82%	0.00%
Silence Behaviours	9.99%	13.29%	**24.85%**	**28.06%**
Defensive Silence Behaviours (Non-verbal)	0.80%	**20.04%**	0.00%	5.40%
Collaboration Behaviours	**57.38%**	29.87%	43.38%	44.27%
Unsupportive Behaviours	0.14%	0.58%	**1.62%**	**3.29%**
Learning or Improvement Oriented Behaviours	5.57%	9.06%	10.44%	2.50%
Familiarity Behaviours	**6.15%**	**7.13%**	1.32%	3.95%
Neutral Behaviours	**2.97%**	**3.66%**	1.91%	0.00%
Total (+) related behaviour	**85.75%**	62.24%	**67.79%**	63.24%
Total (-) related behaviour	11.29%	**34.10%**	30.29%	**36.76%**

4.2 Duration Analysis

Table 3 below presents the same overview as Table 2, listing all behaviours indicative of higher or lower psychological safety, as well as neutral behaviours, but based on the standardised durations. In terms of duration, both monocultural Team A and Team B had similar percentages of time spent on positive psychological safety-related behaviours (74.38% and 74.21%). However, they did differ in time spent on negative psychological safety-related behaviour, with 17.15% for Team A and 23.97% for Team B. In contrast, both multicultural Team I (68.14%) and Team II (61.71%) showed lower durations of positive psychological safety-related behaviours and higher durations of negative psychological safety-related behaviours (28.05% and 38.29%). Neutral behaviours were infrequent in all teams, but more prominent in Team A by duration (8.48%).

Considering the specific behavioural categories, collaboration behaviours occupied the largest share of time spend across all four teams, particularly in Team B (64.16%), Team I (59.10%) and team A (59.05%), while Team II showed a slightly lower percentage (49.90%). Silence behaviours were observed to be the longest in Team II (37.00%) and Team I (25.58%), both multicultural teams. The monocultural Team A and Team B recorded shorter silence durations (15.95% & 14.78%). The percentage of time spent on defensive silence was again the highest in Team B (9.01%), with minimal presence in the other teams. Team I showed no defensive silence by duration, though it did exhibit a more defensive voice compared to the other teams. Learning or improvement-oriented behaviours were most sustained in Team I (3.37%) and Team A (2.08%), while Team B devoted less time (1.26%) to such behaviours, followed by Team II with the least amount of time (0.73%). Notably, Team I and Team A were both categorised as having effective conflict management, while Team B and Team II had ineffective conflict management. In line with the frequency results, Team A followed by Team B exhibited the highest percentage of time spent on familiarity behaviours (2.41% & 1.42%), while Team I and II devoted less time to this (0.71% & 0.92%). Finally, defensive silences were slightly longer in team B and II (9.01% & 1.25%) compared to team A and I (1.14% & 0.00%).

Table 3. Standardised durations of observed behaviours indicative of psychological safety

	Team A (Effective monocultural)	Team B (Ineffective monocultural)	Team I (Effective multicultural)	Team II (Ineffective multicultural)
Voice Behaviours	10.84%	7.37%	4.96%	10.16%
Defensive Voice Behaviours	0.05%	0.01%	2.18%	0.00%
Silence Behaviours	15.95%	14.78%	**25.58%**	**37.00%**
Defensive Silence Behaviours (Non-verbal)	1.14%	**9.01%**	0.00%	**1.25%**
Collaboration Behaviours	59.05%	64.16%	59.10%	**49.90%**
Unsupportive Behaviours	0.01%	0.18%	0.29%	0.04%
Learning or Improvement Oriented Behaviours	**2.08%**	1.26%	**3.37%**	0.73%
Familiarity Behaviours	**2.41%**	**1.42%**	0.71%	0.92%
Neutral Behaviours	**8.48%**	1.82%	3.81%	0.00%
Total (+) related behaviour	**74.38%**	**74.21%**	68.14%	61.71%
Total (-) related behaviour	17.15%	23.97%	**28.05%**	**38.29%**

4.3 Comparative Analysis

4.3.1 Mono- versus Multicultural teams

To investigate psychological safety across different cultural compositions, Team A was compared with Team I, and Team B with Team II, while controlling for conflict management effectiveness to ensure consistent conditions across comparisons. Both multicultural Team I and Team II exhibited higher levels of silence behaviours, both in frequency and duration, compared to the monocultural teams. Specifically, silence behaviours accounted for 24.85% and 28.06% of the behaviours in Team I and Team II, compared to 9.99% and 13.29% in Team A and Team B. This pattern was also visible in the duration data, where silence behaviour made up 25.58% and 37% of the total time spent in multicultural teams, while this was 15.95% and 14.78% in monocultural teams. Furthermore, the percentage of defensive voice behaviours was the highest for multicultural Team I (3.82%), while almost absent in the other teams. Familiarity behaviours were notably more frequent and longer in monocultural teams compared to multicultural teams. Monocultural Team A and B had frequencies of 6.15% and 7.13% with durations of 2.41% and 1.42%, while multicultural Team I and Team II had frequencies of 1.32% and 3.95% with durations of 0.71% and 0.92%. This is interesting because it might suggest that monocultural teams find it easier to establish informal interpersonal connections, potentially due to shared cultural norms and the absence of language barriers. Although voice behaviours were generally more frequent in monocultural teams, the gap was narrower than in the other behavioural categories. Team A and Team B showed the highest frequencies (16.64% and 16.18%), followed by Team I and Team II (12.65% and 12.52%). This may reflect a willingness to speak up across all teams, but with a slightly reduced engagement in multicultural teams.

4.3.2 Teams with effective versus ineffective conflict management

To investigate psychological safety across varying levels of conflict management effectiveness, Team A was compared with Team B, and Team I with team II, while maintaining a consistent cultural composition across each pair. The teams with effective conflict management (Team A and Team I) exhibited higher percentages of behaviours indicative of psychological safety. Positive behaviours comprised 85.75% of observed behaviours in Team A and 67.79% in Team I. In contrast, these figures were 62.24% for Team B and 63.24% for Team II, teams with ineffective conflict management. A similar trend emerged in duration, since Team A and Team I spent more time on positive behaviours (74.38% and 68.14%) compared to Team B and Team II (74.21% and 61.71%). This gap was even more pronounced in the duration of negative behaviours, where Team A and Team I had percentages of 17.15% and 28.05%, while Team B and Team II had percentages of 23.97% and 38.29%. Looking at the individual behavioural categories, the effective teams scored slightly longer durations on learning and improvement-oriented behaviour. In addition, the duration of defensive silence was slightly longer in the ineffective teams. Interestingly to note it that combined with normal silence behaviour, this gap becomes even more pronounced. Thus, silence behaviour could indeed hinder open communication and make conflict management more difficult.

4.4 Exploratory Statistical Analysis

To complement the descriptive and comparative findings, an exploratory statistical analysis was conducted using the Mann-Whitney U test to assess whether the observed differences in psychological safety-related behaviours across mono and multicultural teams and teams with effective versus ineffective conflict management were statistically meaningful. This analysis is considered exploratory, as it does not necessarily aim to confirm hypotheses with inferential certainty, but rather to identify emerging patterns that may guide future research. Although the comparative analysis led to some noticeable trends, such as higher frequencies of silence behaviours in multicultural teams and a lower duration of behaviours negatively related to psychological safety in teams with effective conflict management, none of the results were statistically significant at the standard significance level of 0.05. As shown in Appendices A through D, while differences in behaviour were observable, they did not reach statistical significance within the current sample. Therefore, results should be interpreted with caution. Nonetheless, the behavioural patterns across different team types highlight the potential value for future studies with larger and more diverse samples.

5. DISCUSSION

This thesis explored how psychological safety manifests through observable behaviours in Agile teams, specifically comparing mono- and multicultural team compositions and varying levels of conflict management effectiveness. By incorporating both verbal and non-verbal behavioural analyses, this study contributes to a growing body of literature that advocates for more objective assessments of psychological safety (Zhao et al., 2019; O'Donovan et al., 2020). The findings provide novel insights into the behavioural expressions of psychological safety within Agile teams and offer practical and theoretical implications.

5.1 Theoretical Implications

5.1.1 Psychological Safety and Conflict Management Effectiveness

This research contributes to existing literature by exploring the relationship between the effectiveness of conflict management and psychological safety, an area that is under investigated in literature (Faust, 2023), especially in an Agile context. Research by Faust himself underscored the crucial role of psychological safety as an environmental factor affecting conflict management styles and outcomes among nurse practitioners in acute or critical care settings. He found that psychological safety emerged as a significant predictor of conflict management success. Aligning with his findings, although not statistically significant, this research indicated that the teams categorised as having effective conflict management (Team A and Team I) exhibited a higher percentage of behaviours indicative of higher psychological safety, and a lower percentage of behaviours indicative of lower levels of psychological safety, compared to teams with ineffective conflict management. In addition, this study explored the interplay of both variables through observed behaviour, minimizing the self-reported bias inherent in surveys (O'Donovan et al., 2020).

When examining the individual behavioural categories, the most notable observations were a slightly longer duration of learning and improvement-oriented behaviours in teams with effective conflict management,

and a slightly shorter duration of defensive silence compared to teams with ineffective conflict management. When both silence and defensive silence were considered together, the difference became even more pronounced in both frequency and duration. The multicultural team with ineffective conflict management exhibited by far the highest level of silence behaviours and received the lowest conflict management effectiveness score. Given that silence behaviour is associated with lower psychological safety (O'Donovan et al., 2020), these findings support the idea that psychological safety is essential for constructive conflict resolution (Edmondson, 1999; Faust, 2023), as it enables open communication (Ito et al., 2022; Tekleab et al., 2009).

5.1.2 Psychological Safety in Mono- and Multicultural Teams

This study responds to recent calls for deeper insights into the cultural dynamics of Agile teams as emphasised by Welsch et al. (2024). The comparative findings indicate that multicultural teams displayed higher frequencies and durations of silence behaviours, along with reduced familiarity and voice behaviours, compared to their monocultural counterparts. These patterns suggest that cultural diversity may negatively relate to psychological safety. The increased silence behaviours observed in multicultural teams may reflect discomfort or uncertainty about speaking up in culturally mixed settings as cultural norms around appropriate behaviour vary (Hui et al., 2017; Ng et al., 2019). However, it is also important to interpret silence behaviours from a cultural point of view. Research argues that silence serves different communicative functions in high-context versus low-context cultures (Hayati & Sinha, 2024). In high context cultures, such as those commonly found in India, silence can be seen as a sign of respect or attentiveness. Conversely, in low-context cultures like the Netherlands, silence is more often interpreted as disengagement or a lack of contribution (Hayati & Sinha, 2024). This cultural asymmetry is particularly relevant given that the monocultural teams in this study existed of Dutch participants, while one multicultural team included predominantly Indian team members. This raises important questions about whether psychological safety can be measured uniformly across cultural settings, which is a recommendation for future research (Matsuo et al., 2023; Mahmoud et al., 2022).

The reduction of voice behaviours in multicultural teams, although not statistically significant, partially mirrors findings from Ng et al. (2019). They identified that cultural diversity significantly reduced voice behaviour directed at a supervisor and mediated by (a lack of) cultural intelligence also voice behaviour directed at peers. Another potential explanation for reduced voice behaviour can be found in language barriers, caused by communicating via a corporate language, usually English, resulting in varying proficiency levels (Weinzierl, 2024). When individuals feel pressured to communicate in a foreign language, negative emotions can come up and negatively impact their capacity for action (Weinzierl, 2024) and thus resulting in less voice behaviour.

Similarly, familiarity behaviours, such as informal exchanges or humour, were less frequent and shorter in multicultural teams. This may be attributed to a lack of open communication that is identified as a barrier in multicultural teams, which influences the personal relationships between team members (Welsch et al., 2024). In addition, humour is very culturally dependent both in terms of humour usage and perceptions of humour (Jiang et al., 2019). For example, research showed that Chinese students use humour less frequently compared to Canadian students, especially aggressive humour, indicating potential differences between eastern and western countries (Chen & Martin, 2007). Given that familiarity behaviours are associated with increased psychological safety (O'Donovan et al., 2020), their absence may indicate a more reserved or formal interaction style in multicultural settings. Importantly, while monocultural teams demonstrated a baseline advantage, showing longer durations of positive psychological safety behaviours and shorter durations of negative ones, the multicultural team with higher psychological safety exhibited more effective conflict management than its counterpart. This highlights that, although monocultural teams may more easily sustain psychological safety, fostering psychological safety in multicultural teams can help overcome cultural barriers and enable more constructive conflict resolution. Taken together, these findings contribute to the growing body of literature that highlights both the complexity and potential of multicultural teams.

5.2 Practical Implications

The findings of this study offer several practical implications for organisations and Agile coaches aiming to foster psychological safety within teams. Given the connection between psychological safety and conflict management effectiveness, both fostering psychological safety as equipping teams with tools and structured approaches for handling disagreements is essential. This could include offering conflict resolution training, promoting open dialogue during meetings, and establishing norms for constructive feedback.

For multicultural teams in particular, language barriers (Marquardt & Horvath, 2001) and differing communication norms (Hayati & Sinha, 2024) may hinder interpersonal risk-taking, given the observed tendency towards more silence behaviour and a lower frequency and duration of voice behaviours. Even though Agile teams don't have a formal leader, it is recommended that the individual leading a meeting adopts inclusive facilitation practices. They should spot signs of discomfort/confusion and check in with quitter members or they could implement structured speaking opportunities, where every member gets the chance to speak while reducing the pressure to do this spontaneously, as that can be difficult in a second language (Tenzer & Pudelko, 2025). Overcoming other cultural differences involves learning about one another's cultural backgrounds, including differing communication styles, and the ways in which a member's culture influences team dynamics (Marquardt & Horvath, 2001).

Additionally, the multicultural teams showed slightly less engagement in familiarity behaviours. Since familiarity behaviours are an important part of developing psychological safety within a team, it is recommended to promote and enable leisure-time gatherings and team-building activities to build trust, strengthen informal bonds and ultimately enhance team performance (Welsch et al., 2024).

Last, organisations could consider using video segments from their team meetings to highlight both positive and negative behavioural patterns. This can help teams identify problems in communication and self-diagnose issues related to psychological safety, leading to improvements over time. Moreover, increasing employee

awareness and understanding of psychological safety and its impact can be beneficial across all team types. Providing training on what psychological safety looks like in practice, and how it contributes to team performance, could help team members to become more mindful of their own behaviours during meetings. It can also empower them to gently address or correct behaviours in others that may unintentionally undermined a safe team climate.

6. LIMITATIONS AND FUTURE RESEARCH

While this study provides valuable insights into Agile team dynamics, several limitations must be acknowledged. First, all Agile teams in the sample were drawn from a single organisation in the Netherlands. The specific organisational culture, structure, and implementation of Agile practices may differ from other contexts, industries or geographical regions, limiting the generalisability of the findings. However, given the exploratory nature of this work and the novel, top-notch methodological approach implemented to study teams within this organisation, the focus on one company was deemed appropriate. Still, future research could conduct a similar study in other organisations, industries or cultural contexts.

A second limitation concerns the relatively small sample size, consisting of only four Agile teams (two monocultural and two multicultural, of which two with high and two with low conflict management effectiveness). A small sample size limits the generalisability of the findings (Etz & Arroyo, 2015; Tipton et al., 2016), as the observed behavioural patterns might be influenced by the individual dynamics within the team or the behaviour of a single team member could have disproportionately affected the results. Nonetheless, the small sample allowed for an in-depth, context-rich analysis of psychological safety with a novel methodology, that therefore is still highly valuable (Boddy, 2016). Yet future research may benefit from including a larger sample of Agile teams to improve generalisability.

Another limitation lies in the categorisation of Agile teams into groups with “effective” and “ineffective” conflict management. Due to the relatively small sample size, the differences in mean scores between teams on conflict management effectiveness were not clearly different. For example, the two monocultural teams had mean scores of 5.0 (SD = 0.9) and 5.1 (SD =1.2). Despite the marginal difference, one was classified as ineffective and the other as effective. This small difference, combined with overlapping standard deviations, suggests that the categorisation lacks robustness. Nonetheless, using this binary categorisation was a practical starting point for exploratory research, enabling insights into potential behavioural patterns related to conflict management.

Moreover, teams were selected on a voluntary basis, which may have caused self-selection bias. Teams with more positive dynamics, such as conflict management effectiveness, may have been more willing to participate, potentially affecting the overall results. At the same time, voluntary participation likely increased authenticity as they were being comfortable with being observed. However, future research would benefit from a more diverse and randomised sample that allows for a clearer differentiation between levels of conflict management effectiveness and reduces the risk of selection bias.

Furthermore, conflict management effectiveness was assessed solely through self-reported survey data, which may be influenced by biases such as inaccurate self-assessment. Even though surveys are a widely accepted method in research, to gain a more objective and detailed understanding, future studies could benefit from incorporating observational methods. Conflict management effectiveness could be assessed by analysing team interactions and conflict episodes.

A final limitation concerns the coding of the video recordings. Although the original plan was to have two independent coders per video to improve reliability and reduce observer bias, time constraints caused each video to be coded by only one person. Additionally, all four meetings were analysed by four different coders. This may have introduced systematic differences in interpreting behaviours and led to coding variations across videos. These factors increase the risk of subjective interpretation and reduce the overall reliability. Even so, the use of a detailed and structured codebook helped to maintain a consistency across coders. To strengthen future research, it is recommended that at least two independent coders analyse each video and work towards consensus to enhance coding consistency and data robustness.

Last, while this study indicated a relationship between psychological safety and conflict management effectiveness, it did not investigate the directionality. It is recommended that future research explores whether psychological safety leads to more effective conflict management, vice versa, or whether they are mutually reinforcing. This could potentially be done by a longitudinal study, also to investigate how psychological safety evolves over time.

7. CONCLUSION

This thesis explored how psychological safety is manifested in Agile teams by comparing observed behaviours across teams with varying levels of conflict management effectiveness and cultural diversity. Using a mixed-method approach, integrating video recordings with survey data, this study provided objective insights into team dynamics beyond self-reported perceptions. Although the findings were not significant, the results showed notable patterns. Teams with effective conflict management displayed a higher frequency and longer duration of behaviours indicative of higher psychological safety, while behaviours indicative of lower psychological safety were less frequent, compared to teams with ineffective conflict management. This supports the idea that psychological safety and conflict management effectiveness are interconnected. Furthermore, multicultural teams exhibited more silence behaviours and slightly lower voice, and familiarity behaviours compared to monocultural teams, possibly reflecting the influence of varying cultural norms and language barriers on psychological safety. These findings highlight the importance of increasing awareness of how cultural composition influences Agile teams and of actively fostering a climate of open communication and psychological safety, both of which are essential for effective conflict management.

8. REFERENCES

Abbas, N., Gravell, A. M., & Wills, G. B. (2008). Historical roots of Agile methods: Where did "Agile thinking" come from? In P. Abrahamsson, R. Baskerville, K. Conboy, B. Fitzgerald, L. Morgan, & X. Wang (Eds.), *Agile Processes in Software Engineering and Extreme Programming* (pp. 94–103). Springer. https://doi.org/10.1007/978-3-540-68255-4_10

Alami, A., Zahedi, M., & Krancher, O. (2023). Antecedents of psychological safety in agile software development teams. *Information and Software Technology, 162*, 107267. https://doi.org/10.1016/j.infsof.2023.107267

Almeida, F. (2017). Challenges in migration from waterfall to agile environments. *World Journal of Computer Application and Technology*, *5*(3), 39–49. https://doi.org/10.13189/wjcat.2017.050302

Andersson, M., Moen, O., & Brett, P. O. (2020). The organizational climate for psychological safety: Associations with SMEs' innovation capabilities and innovation performance. *Journal of Engineering and Technology Management*, *55*, 101554. https://doi.org/10.1016/j.jengtecman.2020.101554

Andriyani, Y., Hoda, R., & Amor, R. (2017). Reflection in agile retrospectives. In H. Baumeister, H. Lichter, & M. Riebisch (Eds.), *Agile processes in software engineering and extreme programming: 18th International Conference,* (pp. 3–19). Springer. https://doi.org/10.1007/978-3-319-57633-6_1

Beck, K., Beedle, M., van Bennekum, A., Cockburn, A., Cunningham, W., Fowler, M., Grenning, J., Highsmith, J., Hunt, A., Jeffries, R., Kern, J., Marick, B., Martin, R. C., Mellor, S., Schwaber, K., Sutherland, J., & Thomas, D. (2001). *The Agile Manifesto*. https://agilemanifesto.org/

Bergner, R. M. (2011). What is behavior? And so what? *New Ideas in Psychology, 29*(2), 147–155. https://doi.org/10.1016/j.newideapsych.2010.08.001

Boddy, C. R. (2016). Sample size for qualitative research. *Qualitative Market Research: An International Journal, 19*(4), 426–432. https://doi.org/10.1108/QMR-06-2016-0053

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77-101. https://doi.org/10.1191/1478088706qp063oa

Buvik, M. P., & Tkalich, A. (2022). Psychological safety in agile software development teams: Work design antecedents and performance consequences. *Proceedings of the 55th Hawaii International Conference on System Sciences*. https://doi.org/10.48550/arXiv.2109.15034

Chen, G. H., & Martin, R. A. (2007). A comparison of humor styles, coping humor, and mental health between Chinese and Canadian university students. *Humor: International Journal of Humor Research, 20*(3), 215–234. https://doi.org/10.1515/HUMOR.2007.011

Conboy, K. (2009). Agility From First Principles: Reconstructing the Concept of Agility in Information Systems Development. *Information Systems Research*, *20*(3), 317-480. https://doi.org/10.1287/isre.1090.0236

Conforto, E. C., Salum, F., Amaral, D. C., da Silva, S. L., & Magnanini de Almeida, L. F. (2014). Can agile project management be adopted by industries other than software development? *Project Management Journal*, *45*(3), 21–34. https://doi.org/10.1002/pmj.21410

Crawford, B., Soto, R., de la Barra, C. L., Crawford, K., & Olguín, E. (2014). Agile software teams can use conflict to create a better product. In C. Stephanidis (Ed.), *HCI International 2014 – Posters' extended abstracts*, *434,* (pp. 24 -29) Springer. https://doi.org/10.1007/978-3-319-07857-1_5

Creswell, J. W., Clark, V. L. P., Gutmann, M. L., & Hanson, W. E. (2003). Advances in mixed methods in research designs. In A. Tashakkori, & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 209-240). SAGE Publications.

Creswell, J. W., & Creswell, J. D. (2017). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). SAGE Publications.

DeChurch, L. A., & Marks, M. A. (2001). Maximizing the benefits of task conflict: The role of conflict management. *International Journal of Conflict Management*, *12*(1), 4–22. https://doi.org/10.1108/eb022847

Del Pilar Barrera, L., & de la Fuente, J. (2025). Dynamic capabilities to drive innovation and competitiveness in a changing business world. In M. D. Lytras, A. C. Șerban, P. Ordóñez de Pablos, & A. Alkhaldi (Eds.), *Open innovation strategies for effective competitive advantage* (pp. 1–15). IGI Global. https://www.igi-global.com/chapter/dynamic-capabilities-to-drive-innovation-and-competitivenss-in-a-changing-business-world/359424

Dingsøyr, T., Nerur, S., Balijepally, V., & Moe, N. B. (2012). A decade of agile methodologies: Towards explaining agile software development. *Journal of Systems and Software, 85*(6), 1213–1221. https://doi.org/10.1016/j.jss.2012.02.033

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., & Bransby, D. P. (2023). Psychological safety comes of age: Observed themes in an established literature. *Annual Review of Organizational Psychology and Organizational*

Behavior, 10, 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*, 23–43. https://doi.org/10.1146/annurev-orgpsych-031413-091305

Edmondson, A. C., & Roloff, K. S. (2008). Overcoming barriers to collaboration: Psychological safety and learning in diverse teams. In E. Sales, G. F. Goodwin, & C. S. Burke (Eds.), *Team effectiveness in complex organizations: Cross-disciplinary perspectives and approaches* (pp. 183–208). Routledge. https://doi.org/10.4324/9780203889312

Etz, K. E., & Arroyo, J. A. (2015). Small sample research: Considerations beyond statistical power. Prevention Science, *16*(7), 1033–1036. https://doi.org/10.1007/s11121-015-0585-4

Faust, S. T. (2023). *Psychological safety and conflict management among nurse practitioners in interprofessional teams* (Doctoral dissertation, Abilene Christian University). Digital Commons @ ACU. https://digitalcommons.acu.edu/etd/701/

Gren, L. (2017). The links between Agile practices, interpersonal conflict, and perceived productivity. *In Proceedings of the 21st International Conference on Evaluation and Assessment in Software Engineering (EASE '17)* (pp. 292–297). ACM. https://doi.org/10.1145/3084226.3084269

Gren, L., & Lenberg, P. (2018). The importance of conflict resolution techniques in autonomous agile teams*. In Proceedings of the 19th International Conference on Agile Software Development: Companion (XP '18)* (pp. 1–4). ACM. https://doi.org/10.1145/3234152.3234185

Groysberg, B., Lee, J., Price, J., & Cheng, J. Y. J. (2018). The leader's guide to corporate culture. *Harvard Business Review*, https://egn.com/dk/wp-content/uploads/sites/3/2020/01/HBR-The-Leaders-guide-to-Corporate-Culture.pdf

Hayati, D., & Sinha, S. (2024). *Decoding silence in digital cross-cultural communication: Overcoming misunderstandings in global teams. Language, Technology, and Social Media, 2*(2), 128–144. https://doi.org/10.70211/ltsm.v2i2.60

Hennel, P., & Rosenkranz, C. (2021). Investigating the "socio" in socio-technical development: The case for psychological safety in agile information systems development. *Project Management Journal*, *52*(1), 11–30. https://doi.org/10.1177/8756972820933057

Hui, K. M., Bednall, T. C., Wang, L., & Sanders, K. (2017). Effects of cultural diversity and emotional regulation strategies on team member performance. *Australasian Journal of Organisational Psychology*, *10*, 1–14. https://doi.org/10.1017/orp.2017.4

Ito, A., Sato, K., Yumoto, Y., Sasaki, M., & Ogata, Y. (2022). A concept analysis of psychological safety: Further understanding for application to health care. *Nursing Open, 9*(1), 467–489. https://doi.org/10.1002/nop2.1086

Jehn, K. A. (1995). A multimethod examination of the benefits and detriments of intragroup conflict. *Administrative Science Quarterly*, *40*(2), 256–282. https://doi.org/10.2307/2393638

Jiang, T., Li, H., & Hou, Y. (2019). Cultural differences in humor perception, usage, and implications. *Frontiers in Psychology, 10*, 123. https://doi.org/10.3389/fpsyg.2019.00123

Jiang, Z., Hu, X., Wang, Z., & Jiang, X. (2019). Knowledge hiding as a barrier to thriving: The mediating role of psychological safety and the moderating role of organizational cynicism. *Journal of Organizational Behavior, 40*(7), 800–818. https://doi.org/10.1002/job.2358

Kim, J. (2023). Non-parametric tests for beginners (Part 1: Rank and sign tests). Towards Data Science. https://towardsdatascience.com/non-parametric-tests-for-beginners-part-1-rank-and-sign-tests-629704f27f2f-2/

Levitis, D. A., Lidicker, W. Z., & Freund, G. (2009). Behavioural biologists do not agree on what constitutes behaviour. *Animal Behaviour, 78*(1), 103–110. https://doi.org/10.1016/j.anbehav.2009.03.018

Magpili, N. C., & Pazos, P. (2018). Self-Managing Team Performance: A Systematic Review of Multilevel Input Factors. *Small Group Research*, *49*(1)*,* 3-33. https://doi.org/10.1177/1046496417710500

Mahmoud, M. A., bin Ahmad, S., & Poespowidjojo, D. A. L. (2022). Validation of the psychological safety, psychological empowerment, intrapreneurial behaviour and individual performance measurements. *RAUSP Management Journal, 57*(3), 219–234. https://doi.org/10.1108/RAUSP-11-2020-0252

Marquardt, M. J., & Horvath, L. (2001). *Global teams: How top multinationals span boundaries and cultures with high-speed teamwork*. Palo Alto, Davies-Black Publishing.

Matsuo, A., Tsujita, M., Kita, K., Ayaya, S., & Kumagaya, S. (2023). Developing and validating Japanese versions of psychological safety scale, knowledge sharing scale and expressed humility scale. *Management and Labour Studies, 49*(3), 375–388. https://doi.org/10.1177/0258042X231191871

Mura, M., Lettieri, E., Radaelli, G., & Spiller, N. (2016). Behavioural operations in healthcare: A knowledge sharing perspective. *International Journal of Operations & Production Management*, 36(10), 1222–1246. https://doi.org/10.1108/IJOPM-04-2015-0234

Naslund, D., & Kale, R. (2020). Is agile the latest management fad? A review of success factors of agile transformations. *International Journal of Quality and Service Sciences, 12*(4)*,* 489–504. https://doi.org/10.1108/IJQSS-12-2019-0142

Ng, K., Van Dyne, L., & Ang, S. (2019). Speaking out and speaking up in multicultural settings: A two-study examination of cultural intelligence and voice behavior. *Organizational Behavior and Human Decision Processes, 151*, 150–159. https://doi.org/10.1016/j.obhdp.2018.10.005

Niederman, F., Lechler, T., & Petit, Y. (2018). A research agenda for extending agile practices in software development and additional task domains. *Project Management Journal, 49*(6), 3–17. https://doi.org/10.1177/8756972818802713

O'Donovan, R., Van Dun, D., & McAuliffe, E. (2020). Measuring psychological safety in healthcare teams: Developing an observational measure to complement survey methods. *BMC Medical Research Methodology, 20*(1), 203. https://doi.org/10.1186/s12874-020-01066-z

O'Leary, D. F. (2016). Exploring the importance of team psychological safety in the development of two interprofessional teams. *Journal of Interprofessional Care, 30*(1), 29–34. https://doi.org/10.3109/13561820.2015.1072142

Panda, A., & Gupta, R. K. (2004). Mapping cultural diversity within India: A meta-analysis of some recent studies. *Global Business Review, 5*(1)*,* 27–49. https://doi.org/10.1177/097215090400500103

Peeters, T., Van De Voorde, K., & Paauwe, J. (2022). The effects of working agile on team performance and engagement, *Team Performance Management*, *28*(1/2), 61-78. https://doi.org/10.1108/TPM-07-2021-0049

Rivera, A. E., Rodríguez-Aceves, L., & Mojarro-Duran, B. I. (2021). Enabling knowledge sharing through psychological safety in inter-organisational arrangements. *Journal of Knowledge Management*, *25*(5), 1170–1193. https://doi.org/10.1108/JKM-04-2020-0241

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

Sherf, E. N., Parke, M. R., & Isaakyan, S. (2021). Distinguishing voice and silence at work: Unique relationships with perceived impact, psychological safety, and burnout. *Academy of Management Journal*, *64*(1), 114–148. https://doi.org/10.5465/amj.2018.1428

Stanley, J. C., & Gross, E. D. (2020). *Project management handbook: Simplified Agile, Scrum and DevOps for beginners.* Prosper Consulting Inc.

Steckler, A., McLeroy, K. R., Goodman, R. M., Bird, S. T., & McCormick, L. (1992). Toward integrating qualitative and quantitative methods: An introduction. *Health Education Quarterly, 19*(1), 1–8. https://doi.org/10.1177/109019819201900101

Tekleab, A. G., Quigley, N. R., & Tesluk, P. E. (2009). A Longitudinal Study of Team Conflict, Conflict Management, Cohesion, and Team Effectiveness. *Group & Organization Management*, *34*(2), 170-205. https://doi.org/10.1177/1059601108331218

Tenzer, H., & Pudelko, M. (2015). Leading across language barriers: Managing language-induced emotions in multinational teams. *The Leadership Quarterly, 26*(4), 606–625. https://doi.org/10.1016/j.leaqua.2015.05.006

Thorgren, S., & Caiman, E. (2019). The role of psychological safety in implementing agile methods across cultures. *Research-Technology Management*, *62*(2), 31–39. https://doi.org/10.1080/08956308.2019.1563436

Tipton, E., Hallberg, K., Hedges, L. V., & Chan, W. (2016). Implications of small samples for generalization: Adjustments and rules of thumb. *Evaluation Review, 41*(5), 1–34. https://doi.org/10.1177/0193841X16655665

Tirmizi, S. A. (2008). Towards understanding multicultural teams. In C. B. Halverson & S. A. Tirmizi (Eds.), *Effective multicultural teams: Theory and practice* (pp. 1–20). Springer.

Wall, J. A., & Callister, R. R. (1995). Conflict and its management. *Journal of Management*, *21*(3), 515–558. https://doi.org/10.1177/014920639502100306

Welsch, D., Burk, L., Mötefindt, D., & Neumann, M. (2024). Navigating cultural diversity: Barriers and benefits in multicultural agile software development teams. *In Proceedings of the 39th ACM/SIGAPP Symposium on Applied Computing (SAC '24)* (pp. 818–825). Association for Computing Machinery. https://doi.org/10.1145/3605098.3635988

Zhao, E., Thatcher, S. M. B., & Jehn, K. (2019). Instigating, Engaging in, and Managing Group Conflict: A Review of Literature Addressing the Critical Role of the Leader in Group Conflict. *Academy of Management Annals*, *13*(1). https://doi.org/10.5465/annals.2016.0153

10. APPENDIX

Appendix A, B, C and D present an overview of the outcome of the Mann Whitney U tests, including the *p*-value and an interpretation on significance.

Appendix A. Mann-Whitney U Test, Standardised Frequencies between Mono- and Multicultural teams

Behavioural Category	Test Statistic	*p* – value	Significant? (Critical value α = .05)
Voice Behaviours	4	0.3333	No
Defensive Voice Behaviours	2	1.0000	No
Silence Behaviours	0	0.3333	No
Defensive Silence Behaviours (Non-verbal)	3	0.6667	No
Collaboration Behaviours	2	1.0000	No
Unsupportive Behaviours	0	0.3333	No
Learning or Improvement Oriented Behaviours	2	1.0000	No
Familiarity Behaviours	4	0.3333	No
Neutral Behaviours	4	0.3333	No

Appendix B. Mann-Whitney U Test, Standardised Frequencies between Teams with Effective versus Ineffective Conflict Management, critical value a = .05

Behavioural Category	Test Statistic	*p* – value	Significant? (Critical value α = .05)
Voice Behaviours	3	0.6667	No
Defensive Voice Behaviours	4	0.3333	No
Silence Behaviours	1	0.6667	No
Defensive Silence Behaviours (Non-verbal)	0	0.3333	No
Collaboration Behaviours	3	0.6667	No
Unsupportive Behaviours	1	0.6667	No
Learning or Improvement Oriented Behaviours	3	0.6667	No
Familiarity Behaviours	1	0.6667	No
Neutral Behaviours	2	1.0000	No

Appendix C. Mann-Whitney U Test, Standardised Durations between Mono- and Multicultural teams

Behavioural Category	Test Statistic	*p* – value	Significant? (Critical value α = .05)
Voice Behaviours	3	0.6667	No
Defensive Voice Behaviours	2	1.0000	No
Silence Behaviours	0	0.3333	No
Defensive Silence Behaviours (Non-verbal)	3	0.6667	No
Collaboration Behaviours	3	0.6667	No
Unsupportive Behaviours	1	0.6667	No
Learning or Improvement Oriented Behaviours	2	1.0000	No
Familiarity Behaviours	4	0.3333	No
Neutral Behaviours	3	0.6666	No

Appendix D. Mann-Whitney U Test, Standardised Durations between Teams with Effective versus Ineffective Conflict Management

Behavioural Category	Test Statistic	*p* – value	Significant? (Critical value α = .05)
Voice Behaviours	2	1.0000	No
Defensive Voice Behaviours	4	0.3333	No
Silence Behaviours	2	1.0000	No
Defensive Silence Behaviours (Non-verbal)	0	0.3333	No
Collaboration Behaviours	2	1.0000	No
Unsupportive Behaviours	2	1.0000	No
Learning or Improvement Oriented Behaviours	4	0.3333	No
Familiarity Behaviours	2	1.0000	No
Neutral Behaviours	4	0.3333	No