

**How is Observed Psychological Safety related to Job
Satisfaction in Self-Managing Agile Team Members?**
**‘A Mixed Methods Study Using Video Observations and
Surveys’**

Master Thesis

Harm Jan Meijer

MSc Business Administration

International Management & Consultancy

University of Twente

Graduation Committee Members:

1st: Dr. L. Carminati

2nd: Dr. D. H. van Dun

Date: 31-01-2023

Word Count: 14.753 words

Abstract

Purpose – Working in self-managing agile teams (SMATs) has become increasingly popular in organizations during the last decade. Even though research on SMATs has shown positive outcomes of SMAT implementation, how observed psychological safety and job satisfaction of SMAT members are related is still unclear.

Methodology – To explore the relationship between observed psychological safety and job satisfaction of SMAT members, this research used a mixed methods type of design by combining novel video observations with survey data. Five hypotheses were formulated in which relationships about perceived and observed psychological safety and job satisfaction were suggested. The data was collected from a prominent Dutch service organization working agile. A frequency count and statistical analysis were conducted in the quantitative analysis of this research. Episode analysis was conducted in the qualitative analysis of this research.

Findings – Although the results from the quantitative analysis were not significant, the in-depth explorations of the video observations through the episode analysis seemed to reveal that there is a positive relationship between observed psychological safety in SMATs and the level of job satisfaction of SMAT members. By improving the working environment and working atmosphere in organizations, the level of psychological safety of SMAT members is likely to be positively influenced. As a consequence, the level of job satisfaction of SMAT members may increase.

Originality – This study uses an innovative mixed methods type of design by combining novel video observations with survey data. It answers the recent calls by authors to study the concept of psychological safety with the help of observational techniques and in an agile setting.

Practical Implication – Organizations shifting to an agile setting with self-managing teams may face issues regarding the level of psychological safety of their SMAT members. This study reveals several insights on how to improve the level of psychological safety of SMAT members and what effects this may have, in particular on their level of job satisfaction.

Keywords: Agile, Job Satisfaction, Observed Psychological Safety, Self-Managing Teams, Surveys, Video Observations

Table of Contents:

1. Introduction	p. 4-6
2. Theoretical Background	p. 6-13
2.1 Self-Managing Agile Teams	p. 6-8
2.2 Psychological Safety	p. 8-11
2.3 Psychological Safety and Job Satisfaction	p. 11
2.4 Sub-Dimensions of Psychological Safety and Job Satisfaction	p. 11-13
3. Methodology	p. 13-20
3.1 Research Design	p. 13-14
3.2 Data Collection & Research Instruments	p. 14-16
3.3 Quantitative Measurements	p. 16-17
3.3.1 <i>Observed Psychological Safety</i>	p. 16-17
3.3.2 <i>Perceived Psychological Safety</i>	p. 17
3.3.3 <i>Job Satisfaction</i>	p. 17
3.4 Quantitative Data Analysis	p. 17-20
3.5 Qualitative Data Analysis	p. 20
4. Results	p. 20-35
4.1 Sample Characteristics and Frequency Count	p. 21-22
4.2 Relationship between Perceived and Observed Psychological Safety	p. 22-23
4.3 Psychological Safety and Job Satisfaction	p. 23
4.4 Sub-Dimensions of Psychological Safety and Job Satisfaction	p. 23-25
4.4.1 <i>Collaboration Behaviours</i>	p. 23-24
4.4.2 <i>Voice Behaviours</i>	p. 24-25
4.4.3 <i>Silence Behaviours</i>	p. 25
4.5 Episode Analysis	p. 25-35
4.5.1 <i>Episode Analysis - First Type of Episode</i>	p. 26-30
4.5.2 <i>Episode Analysis - Second Type of Episode</i>	p. 30-34
4.5.3 <i>Summary of the Episode Analysis</i>	p. 34-35
5. Discussion	p. 35-41
5.1 Theoretical Contribution	p. 36-38
5.2 Practical Contribution	p. 39-41
6. Limitations & Future Research	p. 41-42
7. Conclusion	p. 42-44
8. Acknowledgements	p. 44
9. References	p. 44-46
10. Appendix	p. 46-51

1. Introduction

Working in teams has become increasingly popular during the last decade. Previously, traditional organizations were mostly hierarchically structured with a top-down approach and a clear division of power and influence. Recently, many organizations shifted from this hierarchical structure to a flatter organizational structure in which power and influence is more equally divided among employees (Meyer, 2015). To illustrate this, Meyer (2015) argued that the introduction of self-managing teams (SMTs) in organizations exemplifies the shift in the way businesses are organized. SMTs are characterized by having high levels of autonomy, and team members having collective responsibility for the process outcomes (Humphrey et al., 2007). This enhances decision-making in SMTs and improves the quality of ongoing business processes. In addition to the rise of SMTs, over the last decade the concept of 'agile' also became increasingly popular for organizations to implement (Fernandez et al., 2008). The concept originated from the world of software development and has a connection to SMTs as 'agile' is often implemented in such teams (Dikert et al., 2016). Self-managing agile teams (SMATs) are independent teams in organizations where team members closely collaborate on small projects and therefore are highly dependent on each other (De Jong et al., 2004).

In order to boost such collaboration, among other factors such as trust and taking time together to get used to the new working situation (Bagshaw et al., 2007), psychological safety is a very important concept. SMATs seem to work particularly well when there is a psychological safe working environment for team members in which they feel free to speak up and are able to openly discuss all kinds of issues (De Jong et al., 2004). Edmondson et al. (2007, p. 283) defined psychological safety as 'a shared belief amongst individuals as to whether it is safe to engage in interpersonal risk-taking in the workplace.' By enhancing communication and collaboration between team members, psychological safety can ultimately positively affect the level of job satisfaction of team members (Spector, 1997). Brief et al. (2002, p. 280) defined job satisfaction as 'an employee's cognitive and emotional evaluation of their job, which has either a positive or negative form.' They further suggested that an increased level of job satisfaction likely leads to a higher level of involvement of team members to the ongoing business processes.

However, psychological safety has usually been studied as a self-perceived variable which means that individuals were rating themselves through survey questions. Donaldson et al. (2002) noted that only relying on survey measures is a key shortcoming in organizational behaviour research, since there is a high risk of creating bias as a result of formulating survey questions into a certain direction unintentionally. This is in line with findings of more recent papers, such as O'Donovan et al. (2020), Sherf et al. (2021) and Hoogeboom et al. (2021), in which the authors suggested that further research on measuring psychological safety in teams should be done in a more innovative way by, for instance, using video observations. Indeed, 'observational techniques may offer insights into team psychological safety that the team themselves are not fully aware of and that complement findings from survey measures' (O'Donovan et al., 2020, p. 2). In this type of research both verbal and nonverbal behaviours of team members are analyzed which provides new insights on psychological safety in SMATs. We are then speaking about *observed* psychological safety and this concept has not received much scientific attention, yet.

Therefore, the goal of this thesis is to explore how *observed* psychological safety can influence job satisfaction in SMAT members. By using video observations and surveys this research aims to present an innovative way of conducting research. Hence, the research question of this thesis is:

'How is observed psychological safety related to job satisfaction in self-managing agile team members?'

In terms of theoretical contribution, this research aims to extend the current literature on psychological safety by focusing on an innovative way of measuring the concept (Sherf et al., 2021), i.e., via video observations, and link it to job satisfaction within the peculiarity of the agile setting. In terms of practical contribution, this research aims to increase awareness of managers on the importance of both psychological safety and job satisfaction in organizations by providing new insights, while focusing on an agile setting. Existing literature such as Brief et al. (2002) suggested that higher levels of job satisfaction of team members lead to positive effects such as increased personal well-being and more job involvement. Managers could use the findings of this research to improve the work environment for employees in their organizations. Consultants could use the findings of this research in their advice to managers

to inform them about the importance of both psychological safety and job satisfaction and their potential positive effects to employees.

The remaining part of this thesis is structured as follows. The theoretical background builds on the introduction by providing an in-depth analysis on the main concepts of this paper which are SMATs, psychological safety and job satisfaction. This in-depth analysis is used as a basis for the formulation of hypotheses. After the theoretical background the methodology used in this research is presented. Then, the outcomes of the research are presented to the reader in the results chapter. Afterwards, the meaning and interpretation of the results is addressed in the discussion chapter. The following chapter discusses the limitations of this research and suggestions for future research. Thereafter, the conclusion of this research is presented. Lastly, the acknowledgements, reference list and appendices are shown.

2. Theoretical Background

This chapter presents the theoretical background of this thesis. Subchapter 2.1 covers the concept of SMATs. The first part of the subchapter provides insights on SMTs, and the second part focuses on SMTs in an agile setting (SMATs) and recent developments in agile methodologies. Subchapter 2.2 provides insights on the concept of psychological safety and its relationship to SMATs. Subchapter 2.3 addresses the concept of job satisfaction and its relationship to psychological safety in SMAT members. Subchapter 2.4 builds on the preceding subchapters and delves into the subdimensions of psychological safety and their relationships to job satisfaction. In the subchapters several hypotheses are formulated which are expected to help in answering the research question of this thesis.

2.1 Self-Managing Agile Teams

Self-managing teams (SMTs) were defined as ‘a group of individuals with diverse skills and knowledge with the collective autonomy and responsibility to plan, manage and execute tasks interdependently to attain a common goal’ (De Jong et al., 2004, p. 19). This means that SMTs have the autonomy to make decisions regarding their projects for themselves without having to discuss and justify these external. Humphrey et al. (2007) noted that a SMT can independently decide how to deal with issues like project and conflict management, formulation of a strategy, problem solving, evaluation of performance and development of

personal skills. According to Hackman (2002), power in SMTs is divided among team members through a shared leadership model where all members hold collective responsibility for the project outcome. Members of SMTs are able to collaboratively divide the tasks among each other based on their unique competences. Therefore, it can be argued that decision making in SMTs is enhanced and performance increases through harnessing the specialized knowledge and skills from the team members. This is in line with the findings of Cooney (2004), who found that the ability of SMTs to make decisions independently positively affects team performance. Furthermore, members of SMTs are highly dependent on each other and forced to collaborate a lot during projects (De Jong et al., 2004). Millikin et al. (2010) showed that SMTs often consist of self-regulating individuals who are highly skilled to fulfill their tasks and are willing to back up on other team members when needed. All these findings suggested a link between more intense collaboration and better performance of SMTs through the collection of actions of team members.

Due to their proven potential the development and use of SMTs in larger projects and bigger companies became more attractive recently (Dikert et al., 2016). At the same time, the concept of 'agile' emerged and quickly became very popular. As a consequence, a new agile methodology called 'scrum' originated. Agile scrum is based on iterative and incremental processes and is designed to deliver value throughout the development of a project (Khalid et al., 2020). Agile scrum projects are characterized by being fast, flexible, adaptable and effective (Srivastava et al., 2017). In so-called 'sprints' SMT-members work intensively together on projects focusing on incremental development (Khalid et al., 2020). Normally, a sprint lasts between two and four weeks. A sprint typically contains several meetings intended to reach alignment in SMTs (Khalid et al., 2020). During a 'planning' meeting the planning for the upcoming project(s) is made, 'refinement' meetings are held to ensure that SMT-members are on the right track and at the end of a sprint a 'retrospective' meeting is held to reflect on the sprint as a whole.

As one of the most prominent scholars on agile way of working, Cockburn (2005) identified seven agile principles in his so called 'Crystal' methodology which are related to SMTs. The 'Crystal' methodology focused on people, community, talents, communications, skills and interaction and described underlying values and principles which led to successful practices while using agile methodologies. The seven agile principles focused on interactions within

teams to enhance unity and smooth collaboration in SMTs. The principles relate to agile and therefore focus shifts to self-managing agile teams (SMATs). The first and main principle identified is *osmotic communication*, which means that ‘questions and answers flow naturally and with surprisingly little disturbance among the team.’ (Cockburn, 2005, p. 55). This principle favors open communication, as bad communication between team members may cause unforeseen costs, delays and irritations. *Reflective improvement* is the second principle. During reflection is discussed how things went and how to improve on certain aspects in the future. This is likely to stimulate unity in SMATs (Cockburn, 2005, p. 55). It requires the team to critically assess the performance over a beforehand set period of time. The third principle is *frequent delivery*, which is very common in agile methodology. It provides SMATs the opportunity to get continuous feedback from their stakeholders which can be used to optimize and streamline the ongoing processes. *Focus* is the fourth principle, which emphasizes the importance of the ability of SMAT members to do their jobs unimpeded. Cockburn (2005) expressed that team members should not be asked to perform multiple tasks at once as this comes at the expense of the effectiveness of their work. The fifth principle, *easy access to expert users*, is related to the third principle (*frequent delivery*), as both aim to ‘put the customer at the center’ (Cockburn, 2005, p. 51) in order to gain continuous feedback on the ongoing process. The sixth principle is *working in a technological environment*, as this facilitates SMAT members the opportunity to work with modern and innovative tools which should increase the quality of the ongoing process. The last principle is *personal safety*, which implies that ‘team members should feel free to speak up without fear for negative consequences.’ (Cockburn, 2005, p. 56). Given its focus on individuals’ freedom to voice their opinions without fearing adverse repercussions, this last principle is associated with the concept of psychological safety. Since SMAT members have to collaborate intensively during business processes psychological safety has a huge impact on communication, performance and feelings of SMAT members (Edmondson, 1999). For instance, Hennel et al. (2021) found that psychological safety, together with transparency and trust, is a critical success factor for agile teams. It is thus important to have a psychological safe working environment to stimulate the performance of SMATs.

2.2 Psychological Safety

The concept of psychological safety was originally developed by Schein et al. (1965, p. 23) who described the concept as 'the extent to which individuals feel secure and confident in their ability to manage change.' In a more recent and commonly-used definition, psychological safety was defined as 'a shared belief amongst individuals as to whether it is safe to engage in interpersonal risk-taking in the workplace' (Edmondson et al., 2007, p. 283). Edmondson (1999, p. 364) already described a psychological safe work environment as a working environment in which 'employees feel safe to voice ideas, are willing to seek feedback, provide honest feedback to others, collaborate, take risks and experiment in order to stimulate both individual and organizational learning.' It is a working environment in which team members feel support from their colleagues and have respect for each other's competences. It is safe to experiment and to take risks as these activities are regarded as positive intentions in order to learn and intent to stimulate personal development.

Studies established positive links between psychological safety and learning behaviours in SMATs (Edmondson, 1999; Liu et al., 2014). This means that when there is the feeling of psychological safety among team members, the level of openness increases which positively affects the amount of knowledge-sharing. The more knowledge is shared in SMATs the more likely it is that personal development can take place. Furthermore, the findings of Edmonson (1999) were also supported by Carmeli et al. (2007), whose work showed that psychological safety supports individuals to learn from failure. This is because there is space for, and understanding of, taking risks while experimenting new concepts when there is psychological safety. Leroy et al. (2012) found that psychological safety can improve communication outcomes such as more face-to-face conversations and knowledge sharing by team members. Furthermore, psychological safety is often linked and/or confounded with trust. Newman et al. (2017, p. 524) clarified this as was shown that psychological safety is 'conceptually different as it focuses on how group members perceive a group norm, whilst trust focuses on how one person views another.' The concept of psychological safety is considered at different levels, namely the individual, team and organizational level. This research focused on the individual level because psychological safety is perceived differently by each team member. All teams consist of different individuals and therefore it is important to start the analysis of psychological safety on the individual level instead of directly focusing on the team level. The

insights this research provides could be used in future analysis of psychological safety on the team and organizational level.

However, the numerous studies around psychological safety have tended to rely on survey-based measurements, which means that psychological safety is studied mostly as a self-perceived variable. Recent studies such as O'Donovan et al. (2020) and Hoogeboom et al. (2021) suggested that video observations may provide new insights on psychological safety in SMAT members. Consequently, O'Donovan et al. (2020) created a codebook for video observations in which different subdimensions of psychological safety in teams were identified. Some of the subdimensions nicely linked to the descriptions of psychological safety by Edmondson (1999 & 2007). As example, in subdimension 'voice behaviours' providing positive feedback and providing help or solutions were identified as types of behaviours present in a team with a high level of psychological safety. In subdimension 'collaboration behaviours' these were active listening and sharing knowledge, whereas in subdimension 'learning or improvement-oriented behaviours' these were asking for ideas and speaking up with ideas for improvement. On the other hand, in subdimension 'defensive voice behaviours' evading confrontation and denying faults or blame others were identified as types of behaviours present in a team with a low(er) level of psychological safety. In subdimension 'silence behaviours' these were facial expression or body language indicating fear or indifference. In subdimension 'defensive silence behaviours' these were reacting cold and ignoring a joke.

Following this, Sherf et al. (2021) showed that observed psychological safety has a strong effect on voice-usage of team members during a business meeting and that 'a lack of psychological safety should trigger silence to avoid threats and punishments and move away from negative stimuli' (Sherf et al, 2021, p. 117-118). These findings suggested that what we see (observe) during a business meeting would match with what the team members actually feel (perceive) during the business meeting. Hence, perceived and observed psychological safety in an agile setting (SMAT members) are likely to match. Therefore, the following hypothesis is suggested:

Hypothesis 1:

‘There is a positive relationship between perceived and observed psychological safety of SMAT members.’

2.3 Psychological Safety and Job Satisfaction

The concept of job satisfaction was grounded by Hoppock (1935, p. 47), who defined it as ‘any combination of psychological, physiological and environmental circumstances that cause a person truthfully to say I am satisfied with my job.’ The definition focused on internal personal feelings of employees regarding their jobs. In a more commonly-used definition Spector (1997) argued that job satisfaction relates to both the way people feel about their jobs and the various aspects the jobs bring along. This implies that every employee has his or her own unique perception about a certain job and its requirements which can be both positive and negative. This is in line with the findings of Brief et al. (2002, p. 280) who defined job satisfaction as ‘an employee’s cognitive and emotional evaluation of their job, which has either a positive or negative form.’

According to both Hoppock (1935) and Spector (1997), psychological safety is regarded as one of the circumstances to cause higher levels of job satisfaction. Psychological safety enables team members to feel safe to speak up, not to fear to ask for help and to be comfortable with challenging themselves. This leads to a positive perception about the job and its requirements and causes the level of job satisfaction of team members to increase. Furthermore, Brief et al. (2002) noted that increased levels of job satisfaction relate to smooth collaboration between team members. Kim et al. (2020) found that increased levels of job satisfaction relate to more enthusiasm on the job and positive attitudes shown by team members. Given these findings, perceived psychological safety seems to have a reasonable effect on the level of job satisfaction of SMAT members. Hence, in line with the goal of this thesis that is focus on *observed* psychological safety, the following hypothesis is formulated:

Hypothesis 2:

‘Observed psychological safety of SMAT members is positively associated with their job satisfaction.’

2.4 Sub-Dimensions of Psychological Safety and Job Satisfaction

O'Donovan et al. (2020) created a codebook on psychological safety which distinguished observable behaviours into different behavioural categories. 'Collaboration behaviours' is one of the categories. Types of behaviours identified in this category are for example active listening, agreeing/responding positively or enthusiastically to input and acknowledging achievements/congratulating. Already was noted that psychological safety plays an important role in the collaboration between SMAT members (Spector, 1997). Also, a link between smooth collaboration between SMAT members and positive attitudes shown by SMAT members and their level of job satisfaction was made (Brief et al., 2002; Kim et al., 2020). Therefore, the following hypothesis is advanced:

Hypothesis 3:

'Observed collaboration behaviours of SMAT members are positively associated with their job satisfaction.'

'Voice behaviours' is another category identified in the codebook of O'Donovan et al. (2020). Types of behaviours identified in this category are for example disagreeing, providing positive feedback, asking a question and correcting others. Sherf et al. (2021) found that voice behaviours of SMAT members are affected by the level of psychological safety in their team. The research noted that there is a positive relationship between the perception of psychological safety in a team and the voice-usage of SMAT members. Hence SMAT members speak up more frequently when they feel comfortable in the team leading to 'low levels of silence or in-frequent silence' (Sherf et al, 2021, p. 119). That is, voice behaviours are likely to be more frequently displayed in a team with a high level of psychological safety. However, the opposite, i.e., silence behaviours, is also possible. Sherf et al. (2021, p. 118-119) argued that 'a lack of psychological safety should trigger silence to avoid threats and punishments and move away from negative stimuli.' This means that voice behaviours are less frequently displayed in a team with a low(er) level of psychological safety. Feeling comfortable in the team would ultimately positively affect the level of job satisfaction of SMAT members. Hence, the following hypotheses are put forward:

Hypothesis 4a:

'Observed voice behaviours of SMAT members are positively associated with their job satisfaction.'

Hypothesis 4b:

‘Observed voice behaviours are more often displayed by SMAT members in SMATs scoring high on perceived psychological safety than in SMATs scoring low(er) on perceived psychological safety.’

Hypothesis 4c:

‘Observed silence behaviours are more often displayed by SMAT members in SMATs scoring low(er) on perceived psychological safety than in SMATs scoring high on perceived psychological safety.’

Beside the aforementioned types of behaviours, the codebook consisted of other behavioural categories, namely: ‘defensive voice behaviours’, ‘defensive silence behaviours’, ‘unsupportive behaviours’, ‘learning or improvement-oriented behaviours’, ‘familiarity behaviours’ and ‘neutral behaviours’. Concerning these other behavioural categories hypotheses were not formulated, but they were examined in a more exploratory way. This is because for these behavioural categories both positive and negative relationships with job satisfaction were expected, which may be of value for this and future research.

3. Methodology

This chapter presents the methodology applied in this thesis. The research design is addressed in the first subchapter. Then, a subchapter about the data collection & research instruments follows. The remaining subchapters present the quantitative measurements, quantitative data analysis and qualitative data analysis.

3.1 Research Design

This thesis used an innovative mixed methods type of design. Terrell (2011) found that mixed methods research provides a researcher many design options to choose from which ultimately lead to new and more insights than when only relying on one specific type of research. In this research, video observations in combination with survey research were conducted in order to explore the relationship between psychological safety in SMATs and the level of job satisfaction of SMAT members. This mixed methods type of research was chosen because of recent calls in literature that suggested that doing research on organizations working agile

should be done in a more innovative way (O'Donovan et al., 2020; Sherf et al., 2021; Hoogeboom et al., 2021). Methodological innovation is often credited for generating new theoretical insights and is needed to make significant theoretical contributions (Lê et al., 2022). Furthermore, Fetters et al. (2013) noted that integrating qualitative and quantitative data can bring several advantages, such that 'qualitative data can be used to assess the validity of quantitative findings' and 'quantitative data can also be used to help to explain findings from the qualitative data' (Fetters et al., 2013, p. 2135). Consequently, episode analysis was applied to better understand the quantitative findings, since it can provide more in-depth insights into the hypothesis testing. Moreover was expected that the observed behaviours of SMAT members during meetings would provide more objective data than the outcomes of corresponding surveys which were held after those meetings. Since video observations is a quite new and time-consuming research technique, a relatively small sample size was analyzed in this research which to some extent weakens the strength of the findings of this thesis.

3.2 Data Collection & Research Instruments

Data used in this research was obtained by the Organizational Behaviour, Change Management & Consultancy (OBCC) research group at the University of Twente during a large-scale project at a prominent Dutch service organization working agile. In this particular research video and survey data were gathered. The collaborating organization and all participating members agreed upon taking part in the research by providing written consent to OBCC. Moreover, all people working with the data gathered in the research by OBCC signed a non-disclosure agreement at the University of Twente to make sure that the privacy regulations are respected. The data is treated confidentially and is not shared with third parties.

In the project data on a total of 13 different SMATs was gathered. Per SMAT, three different team meetings during a sprint were filmed, for a total of 39 meetings. In so-called 'planning' meetings, the planning for the upcoming sprints were made and discussed. 'Refinement' meetings were held to make sure that every SMAT member was on the right track and to discuss potential opportunities or issues. In 'retrospective' meetings was reflected on how things went during the sprints. After all meetings corresponding surveys were held among

SMAT members to capture more specific data on their perception about the meetings on the individual and team level. The meetings of SMATs 1-8 were physical meetings and took place before the COVID-19 pandemic. The meetings of SMATs 9-12 were online meetings as a consequence of the COVID-19 pandemic. The meeting of SMAT 13 was a hybrid meeting, also as a consequence of the COVID-19 pandemic.

This thesis focused on the retrospective meetings. During retrospective meetings SMAT members were triggered to critically assess their own performance as well as the team performance during the last sprint. This led to situations in which SMAT members ended up into a vulnerable position as they had to judge themselves and others in front of the whole SMAT. Especially then, the verbal and nonverbal behaviours of SMAT members gave a clear indication whether there was a high level of psychological safety in the SMAT or not. During planning and refinement meetings there was much less discussion or judgement between SMAT members. This gave a less obvious indication about the level of psychological safety in the SMATs and therefore those types of meetings were considered less important for this research.

This thesis focused on the retrospective meetings of the first eight SMATs which took place physically before the COVID-19 pandemic. This because the concept of psychological safety could be better assessed during physical meetings than in virtual meetings. During physical meetings, SMAT members were closely together in the same room and were able to look to each other directly. During virtual meetings, however, SMAT members were mostly not in the same room causing a distance between them. As a consequence, the interrelationships between the SMAT members, which gave an indication about the level of psychological safety in the SMAT, could be better observed during the retrospective meetings held physically. Hence, eight retrospective meetings of eight different SMATs were observed and coded with the use of a further finetuned codebook based on by O'Donovan et al. (2020). Furthermore, individual perceptions of SMAT members during the retrospective meetings, captured through corresponding surveys held after these retrospective meetings, were analyzed in order to explore if the video observations and the surveys measured the data consistently.

The eight SMATs consisted of a total of 51 SMAT members and served as the units of analysis of this thesis. Of the 51 SMAT members 43 were male (84.3%) and eight were female (15.7%). 34 SMAT members had a Dutch nationality (66.7%) whereas 17 had an international

background (33.3%). The ages of the SMAT members ranged from 25 years to 61 years with an average age of 41 years. The eight retrospective meetings lasted for an average duration of 57 minutes. The shortest meeting lasted 34 minutes, whereas the longest meeting lasted 100 minutes.

3.3 Quantitative Measurements

Different types of quantitative measurement were used in this thesis, from video observations related to observed psychological safety to survey-based data for both perceived psychological safety and job satisfaction. Below a detailed explanation of these three key variables.

3.3.1 Observed Psychological Safety

Observed psychological safety was measured through the use of a further finetuned codebook based on O'Donovan et al. (2020) for coding. The codebook distinguished observable behaviours that either point to a high or a low(er) level of perceived psychological safety in a team. It consisted of subdimensions 'voice behaviours', 'defensive voice behaviours', 'silence behaviours', 'defensive silence behaviours', 'collaboration behaviours', 'unsupportive behaviours', 'learning or improvement-oriented behaviours', 'familiarity behaviours' and 'neutral behaviours'. Behaviours of SMAT members were observed and coded. 'Voice', 'silence' and 'collaboration' behaviours were especially of importance in this thesis as relationships with both psychological safety and job satisfaction were hypothesized. During the coding process, there was discussion about how to exactly code the behavioural category 'collaboration behaviours'. This because there was a different interpretation about this behavioural category, in particular about subcategory 'active listening', between the coders.

Two different coders were involved in this research. They used Observer XT (version 16) as software. Observer XT is an application used to observe and to code video meetings. The coders independently coded one of the retrospective meetings. Afterwards, four meetings were planned in which the two different observations were compared and discussed about. This was done in an attempt to ensure objectivity and to increase reliability in the coding process. Also, the Cohen's Kappa statistic (inter-coder agreement coefficient) for this retrospective meeting was calculated during these meetings. Eventually, the Cohen's Kappa

statistic turned out to be 0.28, which was rather low, and indicated that there was considerable disagreement between the coders. As mentioned, the main difference between the two coding styles lay in the interpretation of subcategory 'active listening' of the behavioural category 'collaboration behaviours'. After discussion, it was decided that the difference between the two coding styles was too considerable. As a result, the other seven retrospective meetings were coded by only one coder. The reader should take this into account while interpreting the outcomes presented in this thesis.

3.3.2 Perceived Psychological Safety

Measurement of psychological safety was based on the 3-item psychological safety scale by Detert et al. (2007). After a meeting was held, SMAT members were asked about their perception of psychological safety during that meeting. The questions were asked using a 7-point scale anchored to 1 = "*strongly disagree*" and 7 = "*strongly agree*". An example item is: 'It felt safe for me to speak up' (Appendix 1). The Cronbach's Alpha coefficient was 0.961 (Appendix 2).

3.3.3 Job Satisfaction

Measurement of job satisfaction was based on the 4-item job satisfaction scale by Thompson et al. (2012). The scale was used because 'it contains data that affirm the usefulness of using a very short job satisfaction scale to analyze this variable' (Thompson et al., 2012, p. 294). SMAT members were asked about their perception of job satisfaction. The questions were asked using a 7-point scale anchored to 1 = "*strongly disagree*" and 7 = "*strongly agree*". An example item is: 'I find real enjoyment in my job' (Appendix 3). The Cronbach's Alpha coefficient was 0.862 (Appendix 4).

3.4 Quantitative Data Analysis

In this research a relatively small sample size was used for analysis, and the observational data was mainly based on the work of one coder. Therefore, the data was explored with a frequency count and basic statistical tests. As a result, the outcomes need to be treated cautiously.

Before the quantitative analysis started, a dataset was created in Microsoft Excel which contained data resulting from the video observations, retrieved from Observer XT, and the surveys. This dataset was later converted to SPSS Statistics for further statistical analysis.

The data of the video observations was firstly converted into a common format to make it suitable for analysis. This was done as follows: all frequencies of each displayed type of behaviour by a SMAT member during the retrospective meetings were count per individual SMAT member. After this was done, the frequencies were standardized by dividing the number of one particular type of behaviour displayed by a SMAT member during a retrospective meeting by the total number of all behaviours displayed by the SMAT members during that meeting. This standardization of data was done in an attempt to account for the varying duration times of the retrospective meetings. Furthermore, it was done to get an insight on how many times a single SMAT member displayed a particular type of behaviour during a retrospective meeting. Thereafter, in order to determine the scores for the variable observed psychological safety, the values of each type of displayed behaviour per SMAT member were added up and divided by the total number of types of behaviours, which in this case was nine.

The data of the surveys was used to calculate the individual scores per SMAT member for variables perceived psychological safety and job satisfaction. This was done by adding up the scores on the questions on concepts 'meeting psychological safety' and 'job satisfaction' per SMAT member and dividing this number by the total number of questions on these concepts. There were three questions on the concept 'meeting psychological safety', whereas there were four questions on the concept 'job satisfaction'.

Then, a frequency count was applied to get an insight on how the SMAT members behaved (differently) in the different SMATs during the retrospective meetings, according to the video observations. A frequency count shows the number of times a specific circumstance occurs, in this case a specific type of behaviour displayed by a SMAT member during a retrospective meeting. Based on the survey outcomes, the SMATs and SMAT members were divided in either scoring high or scoring low(er) on perceived psychological safety (Table 1). The frequency count provided an insight on how often a certain type of behaviour, for example a 'voice' behaviour, was displayed by SMAT members in SMATs either scoring high or low(er) on perceived psychological safety. With these insights, similarities and differences in displayed

behaviours by SMAT members in the different SMATs could be explored. Data of the video observations and surveys were used in the frequency count in the way as explained above.

To continue, SPSS Statistics software (version 28) was used to run correlation analysis to find evidence to support or reject the hypotheses formulated in the theoretical background (see Chapter 2). Correlation analysis is usually conducted to see whether there is a relationship between two (or more) selected variables (Field, 2018). Furthermore, it provides insights on the strength and the direction of the relationship (Field, 2018). More specifically, correlation coefficients range between a value of -1 and 1. A score of '0' indicates that there is no relationship between the selected variables. A score of -1 indicates a perfect negative relationship and a score of 1 indicates a perfect positive relationship. The P-value gives an insight on the significance level of the relationship.

Before running correlation analysis, it was checked whether the selected variables were normally distributed or not (Appendix 5). This is because, dependent on normality, different types of correlation tests had to be conducted. For example, when a variable is normally distributed the Pearson's R test can be used, whereas when a variable is not normally distributed the Spearman's Rho test can be used. A variable is normally distributed when the P-value of the Shapiro-Wilk test is greater than 0.05 (>5%). In the correlation analysis both normally and not-normally distributed variables were included, therefore the Pearson's R test and the Spearman's Rho test were applied.

Furthermore, comparative tests were run in this thesis. Comparative tests are conducted to compare the mean scores of different groups to see if there are significant differences between the groups (Field, 2018). Before running the comparative tests was checked whether the selected variables were normally distributed or not (Appendix 5). This is because, similarly to correlation analysis, dependent on normality, different types of comparative tests had to be conducted. For example, when a variable is normally distributed t-test can be used, whilst when a variable is not normally distributed the Mann-Whitney U test can be used. In the comparative tests non-normally distributed variables were included, therefore the nonparametric Mann-Whitney U test was applied to investigate the difference in displayed behaviours by SMAT-members in SMATs either scoring high or scoring low(er) on perceived psychological safety. The Mann-Whitney U test examines whether two independently sampled groups differ on a continuous variable (McKnight & Nejab, 2010). As mentioned, the

SMAT members were divided in two groups based on the scores of their SMATs on perceived psychological safety according to the survey outcomes.

3.5 Qualitative Data Analysis

As mentioned, the outcomes of the frequency count and basic statistical tests need to be treated cautiously. Therefore, episode analysis was applied to see if a qualitative analysis could corroborate the quantitative exploration of data in an attempt to strengthen the contents of this thesis.

Episode analysis was conducted to identify differences and/or similarities in the display of certain types of behaviours by SMAT members, as well as in the reactions on these displayed behaviours by other SMAT members in different SMATs. An episode is defined as ‘a sequence of events in terms of a beginning and an ending’ (Jarrett & Liu, 2016, p. 370). For this research, four episodes were analyzed. The shortest episode lasted for about 20 seconds, whilst the longest episode lasted for about one and a half minute. It was decided to analyze two episodes taken from SMATs scoring highest on perceived psychological safety and two episodes taken from SMATs scoring lowest on perceived psychological safety. This in accordance with the formulation of hypotheses 4b and 4c, which focused on differences in displayed types of behaviours by SMAT members, while dividing the SMAT members in two groups based on the survey score of their SMAT on perceived psychological safety. Furthermore, it was decided that two similar types of episodes for both SMATs scoring highest on perceived psychological safety and SMATs scoring lowest on perceived psychological safety were analyzed. This is because it was expected to generate new insights, as it would be very interesting to see how SMAT members behaved in different types of SMATs during similar situations.

4. Results

This chapter presents the results of this thesis. Firstly, subchapter 4.1 shows the sample characteristics and the frequency count with a further interpretation of the outcomes. Then, subchapter 4.2 focuses on the concept of psychological safety as the relationship between perceived and observed psychological safety of SMAT members is investigated. Subchapter 4.3 explores the relationship between both perceived and observed psychological safety and the level of job satisfaction of SMAT members. Thereafter, in the first instance subchapter 4.4

examines the relationship between the display of certain types of behaviours identified in the codebook by O'Donovan et al. (2020) and the level of job satisfaction of SMAT members. Later on, it focuses on the display of certain types of behaviours identified in the codebook by O'Donovan et al. (2020) in different types of SMATs, after dividing the SMAT members in two groups based on the score of their SMAT perceived psychological safety. Lastly, subchapter 4.5 presents the episode analysis which shows how SMAT members behave in different SMATs during similar types of episodes.

4.1 Sample Characteristics and Frequency Count

This subchapter shows the frequency count which was applied in this research. The sample characteristics are shown in Table 1. The table shows which SMAT members and SMATs were scoring high on perceived psychological safety and which SMAT members and SMATs were scoring low(er) on perceived psychological safety according to the survey outcomes.

Table 1. Sample Characteristics

	SMAT 1	SMAT 2	SMAT 3	SMAT 4	SMAT 5	SMAT 6	SMAT 7	SMAT 8
<i>Participants</i>	1-8	9-16	17-24	25-29	30-35	36-41	42-46	47-51
<i>PPS*</i>	6.6	6.3	6.4	6.3	6.2	4.8	6.3	5.6
<i>High/Low(er)</i>	High	High	High	High	High	Low(er)	High	Low(er)

*Perceived Psychological Safety: average of individual PPS scores of SMAT members per SMAT.

The frequency count on displayed types of behaviours by SMAT members during the retrospective meetings, as a result of the video observations, is shown in Table 2. As explained in Chapter 3, 'voice', 'silence' and 'collaboration' behaviours were considered to be important in this research. Therefore, these types of behaviours are highlighted in Table 2.

Table 2. Frequency Count of Displayed Behaviours by SMAT Members (Standardized / %)

	SMAT 1	SMAT 2	SMAT 3	SMAT 4	SMAT 5	<u>SMAT 6</u>	SMAT 7	<u>SMAT 8</u>
<i>Voice Behaviours</i>	4	26	23	2	7	<u>11</u>	10	<u>17</u>
<i>Def. Voice Behaviours</i>	0	0	0	0	0	<u>67</u>	17	<u>17</u>
<i>Silence Behaviours</i>	7	11	30	3	14	<u>10</u>	1	<u>24</u>
<i>Def. Silence Behaviours</i>	0	0	0	0	0	<u>100</u>	0	<u>0</u>
<i>Collaboration Behaviours</i>	7	23	21	6	10	<u>10</u>	9	<u>13</u>
<i>Unsupportive Behaviours</i>	47	13	0	0	20	<u>13</u>	0	<u>7</u>
<i>Learning or Improvement-Oriented Behaviours</i>	9	20	17	10	5	<u>13</u>	0	<u>7</u>
<i>Familiarity Behaviours</i>	17	19	21	19	5	<u>6</u>	11	<u>2</u>
<i>Neutral Behaviours</i>	12	28	27	8	5	<u>8</u>	5	<u>7</u>
Meeting Duration (min.)	43	80	80	36	34	<u>55</u>	48	<u>60</u>

*The scores presented in this table are standardized. This was done in an attempt to account for the varying duration times of the retrospective meetings. Furthermore, the scores are presented as percentages.

**The scores which are underlined in this table belong to SMAT members in SMATs scoring low(er) on perceived psychological safety, as was also shown in Table 1.

The frequency count, surprisingly, showed that there was no striking difference in the display of ‘voice’, ‘silence’ and ‘collaboration’ behaviours by SMAT members in SMATs scoring high and scoring low(er) on perceived psychological safety. This was interesting, because in this research it was hypothesized that both ‘voice’ and ‘collaboration’ behaviours were likely to be displayed more by SMAT members in SMATs scoring high on perceived psychological safety, whereas it was hypothesized that ‘silence’ behaviours were likely to be displayed more by SMAT members in SMATs scoring low(er) on perceived psychological safety. However, both ‘defensive voice’ and ‘defensive silence’ behaviours were mostly displayed by SMAT members in SMATs scoring low(er) on perceived psychological safety. This was as expected in advance of this research. On the contrary, ‘familiarity’ behaviours were mostly displayed by SMAT members in SMATs scoring high on perceived psychological safety. This was also as expected in advance of this research. Lastly, it was striking that ‘unsupportive’ behaviours were obviously displayed more by SMAT members in SMATs scoring high on perceived psychological safety. This was in advance of this research expected to happen more often by SMAT members in SMATs scoring low(er) on perceived psychological safety.

4.2 Relationship between Perceived and Observed Psychological Safety

The first hypothesis was formulated to investigate whether there was a positive relationship between perceived and observed psychological safety of SMAT members. In order to test the first hypothesis a correlation was therefore made between the variables perceived psychological safety and observed psychological safety. It appeared that the variable perceived psychological safety was not normally distributed ($W(51) = 0.761, p < .001$), whereas the variable observed psychological safety was normally distributed ($W(51) = 0.961, p = .088$). Therefore, the Spearman's Rho correlation test was applied. According to the test there was a non-significant weak negative correlation between the two variables, $r(49) = -.00, p = .985$ (Appendix 6). Hence, the first hypothesis was not supported.

4.3 Psychological Safety and Job Satisfaction

The second hypothesis was formulated to investigate whether there was a positive relationship between observed psychological safety and the level of job satisfaction of SMAT members. In order to test the second hypothesis a correlation was therefore made between the variables observed psychological safety and job satisfaction. The variable observed psychological safety was already checked on normality, whereas the variable job satisfaction appeared to be normally distributed ($W(51) = 0.957, p = .065$). Therefore, the Pearson's R correlation test was applied. According to the test there was a non-significant weak positive correlation between the two variables, $r(49) = .10, p = .486$ (Appendix 7). Hence, the second hypothesis was not supported.

For better insight, also a correlation was made between the variables perceived psychological safety and job satisfaction. Both the variables were already checked on normality. The Spearman's Rho correlation test was applied for the variables perceived psychological safety and job satisfaction. According to the test there was a non-significant weak positive correlation between the two variables, $r(49) = .21, p = .139$ (Appendix 8).

4.4 Sub-Dimensions of Psychological Safety and Job Satisfaction

4.4.1 Collaboration Behaviours

The third hypothesis was formulated to investigate whether there was a positive relationship between the display of collaboration behaviours by SMAT members and their level of job satisfaction. In order to test the third hypothesis a correlation was therefore run between the

variables collaboration behaviours and job satisfaction. It appeared that the variable collaboration behaviours was normally distributed ($W(51) = 0.957$, $p = .063$), whereas the variable job satisfaction was checked on normality already. Therefore, the Pearson's R correlation test was applied. According to the test there was a non-significant weak positive correlation between the two variables, $r(49) = .06$, $p = .671$ (Appendix 9). Hence, the third hypothesis was not supported.

4.4.2 Voice Behaviours

Hypothesis 4a was formulated to investigate whether there was a positive relationship between the display of voice behaviours by SMAT members and their level of job satisfaction. In order to test the fourth hypothesis a correlation was therefore made between the variables voice behaviours and job satisfaction. It appeared that the variable voice behaviours was not normally distributed ($W(51) = 0.919$, $p = .002$), whereas the variable job satisfaction was checked on normality already. Therefore, the Spearman's Rho correlation test was applied. According to the test there was a non-significant weak negative correlation between the two variables, $r(49) = -.04$, $p = .796$ (Appendix 10). Hence, the fourth hypothesis was not supported.

Hypothesis 4b was formulated to investigate whether SMAT members in SMATs scoring high on perceived psychological safety displayed voice behaviours more often in comparison to SMAT members in SMATs scoring low(er) on perceived psychological safety. Before starting the analysis, the SMAT members were divided in two groups based on the scores of their SMAT on perceived psychological safety according to the survey outcomes (Table 3).

Table 3. Distribution of SMAT members

SMAT	Participants	PPS	High/Low(er)
1	1-8	6.6	High
2	9-16	6.3	High
3	17-24	6.4	High
4	25-29	6.3	High
5	30-35	6.2	High
6	36-41	4.8	Low(er)
7	42-46	6.3	High
8	47-51	5.6	Low(er)

*The scores which are in bold in this table belong to SMAT members in SMATs scoring high on perceived psychological safety.

As it was already checked, the variable voice behaviours appeared to be non-normally distributed. Therefore, the Mann-Whitney U test was conducted. The display of voice behaviours of SMAT members in SMATs scoring high on perceived psychological safety (Mdn = .03*) was lower than the display of voice behaviours of SMAT members in SMATs scoring low(er) on perceived psychological safety (Mdn = .04*). The Mann-Whitney U test indicated that this difference was statistically non-significant, $U(N_{high} = 40, N_{low(er)} = 11,) = 142.00, z = -1.79, p = .074$ (Appendix 11). Hence, the fifth hypothesis was not supported.

*Standardized Values

4.4.3 Silence Behaviours

Hypothesis 4c was formulated to investigate whether SMAT members in SMATs scoring low(er) on perceived psychological safety displayed silence behaviours more often in comparison to SMAT members in SMATs scoring high on perceived psychological safety. The distribution of SMAT members is shown in Table 3.

It appeared that the variable silence behaviours was not normally distributed ($W(51) = 0.797, p < .001$). Therefore, the Mann-Whitney U test was conducted. The display of silence behaviours of SMAT members in SMATs scoring low(er) on perceived psychological safety (Mdn = .02*) was higher than the display of silence behaviours of SMAT members in SMATs scoring high on perceived psychological safety (Mdn = .00*). The Mann-Whitney U test indicated that this difference was statistically non-significant, $U(N_{high} = 40, N_{low(er)} = 11,) = 143.00, z = -1.79, p = .075$ (Appendix 12). Hence, the sixth hypothesis was not supported.

4.5 Episode Analysis

This subchapter presents the episode analysis which was applied in this research. The episode analysis was conducted with the intention to identify differences and/or similarities in the display of certain types of behaviours by SMAT members, as well as in the reactions on these displayed behaviours by other SMAT members, in different SMATs. A total of four episodes were analyzed. Firstly, it was decided to analyze two episodes taken from SMATs scoring highest on perceived psychological safety, and two episodes taken from SMATs scoring lowest on perceived psychological safety. This was expected to generate insights on how SMAT

members behaved (differently) in different types of SMATs. Furthermore, it was decided to analyze two similar types of episodes for both SMATs scoring highest on perceived psychological safety and SMATs scoring lowest on perceived psychological safety. This was expected to generate insights on how SMAT members behaved (differently) in similar situations, but in different types of SMATs.

To illustrate, in the first type of episode a particular SMAT member, who seemed to be quite frustrated, constructively criticized the way of working of another SMAT, while showing a voice behaviour. This type of episode was analyzed for SMAT 1 and for SMAT 8. This type of episode was chosen because it gave a clear indication about the cohesion of the SMAT members in the different SMATs, which had an effect on the level of psychological safety and job satisfaction in those SMATs. In the second type of episode, a particular SMAT member was constructively criticized by the product owner of the SMAT, who displayed a voice behaviour. This type of episode was analyzed for SMAT 2 and for SMAT 6. This type of episode was chosen because it gave a clear indication about the relationships between the SMAT members and the atmosphere during the retrospective meetings, which had an effect on the level of psychological safety and job satisfaction in the SMATs.

4.5.1 Episode Analysis - First Type of Episode

The first episode came from the retrospective meeting of SMAT 1 which had 8 participants. This episode was chosen for the episode analysis as it shows the cohesion between SMAT members in a SMAT scoring high on perceived psychological safety.

Table 4. Transcript of the First Episode Analysis

The episode starts at 18:32 and ends at 18:52.

Transcript	Follower(s) - Type of Behaviour(s)
'But I always say to them, you should test against acceptance and not against test'	F1 Voice Behaviour
(Crosstalk) 'Yes, because than it is testing what it is developed for'	F4 Collaboration Behaviour
'Yeah, than we are busy with acceptance and they complain about the test'	F5 Voice Behaviour
'Yeah'	F2 Collaboration Behaviour
'Yes, indeed'	F8 Collaboration Behaviour
'You always have this issue, right?'	F7 Learning or Improvement-Oriented Behaviour
'Unfortunately, yeah'	F1 Voice Behaviour
'But acceptance is usually quite stable, so it should work better'	F4 Voice Behaviour
'It is quite stable indeed'	F7 Collaboration Behaviour

It was an active meeting in which all participants, except for participant 6, who seemed disengaged, were very involved in the discussion. All participants were able to see the screen with the sticks on it and some of them were standing. The episode starts about 20 minutes into the meeting and lasts for about 20 seconds. In the episode, participant 1 showed a voice behaviour as he, seemingly being quite frustrated, expressed to have corrected another SMAT regarding their way of working on a certain test multiple times by then. The criticism was brought in a constructive way. In the following crosstalk, most of the other participants showed collaboration behaviours as they enthusiastically backed participant 1 on his frustration and agreed with him. To illustrate, participant 4 expressed that she agreed with the suggestions participant 1 gave to the other SMAT. Afterwards, the discussion continued on how the other SMAT should have adjusted their way of working regarding the test.

Noticeable in this episode was the enthusiastic involvement of almost all SMAT members while they displayed supportive voice and collaboration behaviours after participant 1 expressed his frustration. Together they tried to come to a solution which visibly did

participant 1 well. This gave the indication that there was a strong cohesion between the SMAT members. The atmosphere throughout this retrospective meeting was very pleasant. Especially, a lot of supportive voice and collaboration behaviours were displayed by the SMAT members. This seemed to have a positive effect on the level of job satisfaction of the SMAT members. Considering all of this, there seemed to be a high level of psychological safety while observing this SMAT. The survey outcomes indicated that SMAT members of this SMAT had a high level of perceived psychological safety. Therefore, it can be pointed out that this episode supports hypothesis 1, in which a positive relationship between perceived and observed psychological safety of SMAT members was expected. Furthermore, it can be pointed out that this episode supports hypothesis 2, in which a positive relationship between observed psychological safety and the level of job satisfaction of SMAT members was expected. Next, it can be pointed out that this episode supports hypothesis 3, in which a positive relationship between the display of collaboration behaviours and the level of job satisfaction of SMAT members was expected. Also, it can be pointed out that this episode supports hypothesis 4b, which suggested that voice behaviours would be more often displayed by SMAT members in SMATs scoring high on perceived psychological safety than by SMAT members in SMATs scoring low(er) on perceived psychological safety.

The second episode came from the retrospective meeting of SMAT 8 which had 5 participants. This episode was chosen for the episode analysis as it shows that there was no cohesion between the SMAT members, which resulted in a depressing atmosphere during the retrospective meeting in a SMAT scoring low(er) on perceived psychological safety.

Table 5. Transcript of the Second Episode Analysis

The episode starts at 51:46 and ends at 52:38.

Transcript	Follower(s) - Type of Behaviour(s)	
'If they have an issue, they are first putting a blame on people'	F3	Voice Behaviour
'No, the main problem is that they are not supposed to test in the test environment'	F9	Voice Behaviour
<i>(Starts looking bored at laptop)</i>	F5	Silence Behaviour
<i>(Starts looking bored in the distance)</i>	F6	Silence Behaviour
'It wasn't like that. They have started the acceptance now'	F7	Voice Behaviour
<i>(Crosstalk)</i>		
'They don't have'	F9	Voice Behaviour
'That is not correct'	F3	Voice Behaviour
'Listen, I was not involved in the topic. I am just explaining the diagram'	F7	Defensive Voice Behaviour

It was an active meeting, in which especially participant 3 spoke very often, in a loud tone, while making a lot of hand gestures. This sometimes visibly annoyed the other participants. The participants sat around a table, and all were able to see the screen. Furthermore, they all had a laptop in front of them. The episode starts about 51 minutes into the meeting and lasts for about a minute. Participant 3 displayed a voice behaviour in which he, seemingly being quite frustrated, constructively criticized another SMAT on how they dealt with an issue, in a very loud tone. Participant 9 disagreed with participant 3 as he expressed that the problem of the other SMAT they were talking about at that moment lay somewhere else. He also made hand gestures which made clear that participant 3 should calm down. Participant 9 was clearly annoyed because of participant 3 speaking in a very loud tone constantly. Immediately, participants 5 and 6 started displaying silence behaviours as they looked bored to either their laptop or in the distance. It seemed that this issue was discussed multiple times already. In a response to participant 9, participant 7 expressed his point of view on the problem of the other SMAT, which was completely different from the one than what was discussed up until that moment. Directly, participants 3 and 9 quite roughly expressed to disagree with participant 7, who seemed to feel attacked by this as he looked indignant. He reacted by displaying a defensive voice behaviour as he wanted to evade confrontation. He mentioned,

in a fairly loud tone, that he was not involved in this topic, but just tried to explain a diagram. While doing so, he raised his hands desperately into the air.

Noticeable in this episode was that the participants could not agree on what the problem another SMAT faced actually was. Instead of backing participant 3 on his frustration, the other participants only gave their own opinions and did not really listen to the input of the others. The SMAT members seemed to have a good relationship with each other, however, there was no cohesion between them during this meeting. Furthermore, it was remarkable that participants 5 and 6 directly started displaying silence behaviours after the discussion started, as if this discussion had occurred multiple times before. This clearly indicated that they had enough of this returning discussion. At the end, participant 7 was cut off quite roughly by participants 3 and 9, which visibly annoyed him and led him to displaying a defensive voice behaviour. The atmosphere throughout this retrospective meeting was quite tense and sometimes even aggressive. Remarkable was the display of silence and even defensive voice behaviours by some of the SMAT members during the discussion. Although these behaviours were not always disrespectful, it created a lot of irritations by the SMAT members. Considering all of this, there seemed to be a low level of psychological safety while observing this SMAT. The survey outcomes indicated that SMAT members of this SMAT had a low(er) level of perceived psychological safety. Therefore, it can be pointed out that this episode supports hypothesis 1. Furthermore, it can be pointed out that this episode supports hypothesis 4c, which suggested that silence behaviours would be more often displayed by SMAT members in SMATs scoring low(er) on perceived psychological safety than by SMAT members in SMATs scoring high on perceived psychological safety.

4.5.2 Episode Analysis - Second Type of Episode

The third episode came from the retrospective meeting of SMAT 2 which had 8 participants. This episode was chosen for the episode analysis as it shows the strong relationship and goodwill between the SMAT members in a SMAT scoring high on perceived psychological safety.

Table 6. Transcript of the Third Episode Analysis

The episode starts at 37:50 and ends at 39:15.

Transcript	Follower(s) - Type of Behaviour(s)	
'When this is the result, then I actually think it is 'excellent'. And then I am modest'	F4	Voice Behaviour
'Does everybody agree?'	F9	Collaboration Behaviour
(Crosstalk) 'Yes'	F3, F6, F7	Collaboration Behaviours
'I think it is a bit funny, you start with 'good' and now you move to 'excellent''	F8	Voice Behaviour
'Yeah, but..'	F4	Voice Behaviour
(Interrupts) 'While I think you are in the highest scale already. Then we can also expect more from you'	F8	Unsupportive Behaviour / Voice Behaviour
'Well, that is because I know what else I have done'	F4	Voice Behaviour
'I agree with you, because you really dealt with a lot of screeners for us'	F5	Collaboration Behaviour
(Crosstalk) 'Exactly'	F3	Collaboration Behaviour
'Yes'	F9, F10	Collaboration Behaviour
'Okay, in that case we can leave it at 'excellent''	F8	Voice Behaviour

It was an active, but very structured meeting, in which the SMAT members critically reflected on the achievement of targets during the last sprint. The participants sat around a large table and were all able to see the screen. Some had a laptop in front of them. The episode starts about 38 minutes into the meeting and lasts for about one and a half minute. In the episode, participant 4 displayed a voice behaviour as he proudly reflected on his own performance in a very positive way. Thereafter, the agile coach asked the other SMAT members if they agreed with participant 4 on his positive reflection. In the following crosstalk, almost all participants enthusiastically expressed to agree with participant 4. Then, quite suddenly, participant 8, who was the product owner of the SMAT, displayed a voice behaviour in which he expressed that he did not agree with participant 4 on his positive reflection. Participant 4 tried to react

on this, but was interrupted by participant 8 who continued with his constructive criticism in a fairly neutral tone. Hereafter, participant 4, who sounded and looked quite surprised, gave an argument on why he still thought he had done well during the last sprint. Participant 5 backed participant 4 as he mentioned a certain action participant 4 had done very well during the sprint. In the following crosstalk, a lot of other SMAT members, again enthusiastically, expressed to agree with both participants 4 and 5. As a result, participant 8 eventually agreed to rate participant 4 with an 'excellent'.

Noticeable in this episode was that almost all SMAT members enthusiastically backed participant 4, after he was constructively criticized by the product owner regarding his positive reflection on his own performance during the last sprint. This led to a better rating for participant 4 and this visibly did him well. Also, the other SMAT members seemed satisfied. Striking was that the atmosphere in this retrospective meeting remained quite relaxed and pleasant, despite the fact that the SMAT members were critically reflecting on the achievement of targets during the last sprint. Especially, a lot of (positive) voice and collaboration behaviours were displayed by the SMAT members as they were judging each other. This seemed to have a positive effect on their level of job satisfaction. Also, it gave a clear indication that the SMAT members had very close relationships with each other. Considering all of this, there seemed to be a high level of psychological safety while observing this SMAT. The survey outcomes indicated that SMAT members of this SMAT had a high level of perceived psychological safety. Therefore, it can be pointed out that this episode supports hypothesis 1, in which a positive relationship between perceived and observed psychological safety of SMAT members was expected. Furthermore, it can be pointed out that this episode supports hypothesis 2, in which a positive relationship between observed psychological safety and the level of job satisfaction of SMAT members was expected. Next, it can be pointed out that this episode supports hypothesis 3, in which a positive relationship between the display of collaboration behaviours and the level of job satisfaction of SMAT members was expected. Also, it can be pointed out that this episode supports hypothesis 4b, which suggested that voice behaviours would be more often displayed by SMAT members in SMATs scoring high on perceived psychological safety than by SMAT members in SMATs scoring low(er) on perceived psychological safety.

The fourth and last episode came from the retrospective meeting of SMAT 6 which had 6 participants. This episode was chosen for the episode analysis as it shows a heated discussion between two of the SMAT members, which obviously affected the atmosphere during the retrospective meeting in a negative way in a SMAT scoring low(er) on perceived psychological safety.

Table 7. Transcript of the Fourth Episode Analysis

The episode starts at 08:35 and ends at 09:30.

Transcript	Follower(s) - Type of Behaviour(s)	
'*Name*, just in general right? Okay, it is blocked but also people can help you getting it unblocked'	F1	Voice Behaviour
'Yeah, true'	F4	Collaboration Behaviour
<i>(Start looking at laptop to evade confrontation)</i>	F3, F6	Silence Behaviours
'So we can also say to *name*: “*name*, can you just sit together now with *name* and just really explain what the issue is?”'	F1	Voice Behaviour
'But what is than the role of the scrum master? I don't remember what the scrum master should do?'	F5	Learning or Improvement-Oriented Behaviour
'No, for the impediment it is the responsibility for the whole team, not for the scrum master. If you need to address it, then you need to address it in your stand-up'	F7	Voice Behaviour
'No, that is a whole different story'	F1	Defensive Voice Behaviour

It was quite an active meeting with a lot of discussions, especially between two of the participants. The participants sat around an enormous table, which led them to be sit apart quite far from each other. Yet, all participants were able see the screen and some had a laptop in front of them. The episode starts about eight minutes into the meeting and lasts for about a minute. In the episode, participant 1, the product owner of the SMAT, showed a voice behaviour as she constructively criticized participant 5 on a certain action in a fairly neutral tone. Directly, two other participants started looking at their laptops as they clearly wanted to evade the confrontation. Participant 1 continued by giving, albeit a little frustrated and in a louder tone, a suggestion to participant 5 on how to improve on the action in the future.

Participant 5 desperately raised his hands into the air, as he expressed that he did not understand the role of the scrum master in the situation. Hereafter, participant 7 intervened when she tried to explain the situation to participant 5. In reaction to this, participant 1, who looked annoyed, displayed a defensive voice behaviour as she corrected participant 7 quite roughly on her explanation and therewith cut off the discussion.

Noticeable in this episode was that the participants were not sitting close to each other around the table during the meeting. There was a lot of space between them. This clearly indicated that there was no cohesion between the SMAT members at all. Furthermore, it was remarkable that two of the participants clearly wanted to evade the confrontation, as they started displaying silence behaviours after the first voice behaviour of participant 1 was displayed. They immediately started looking at their laptops in front of them. Also, it was remarkable that, except for participant 7, none of the other SMAT members seemed inclined and willing to help participant 5 with his misunderstanding, which visibly annoyed him. Striking was the display of silence and even defensive voice behaviours by some of SMAT members during the heated discussion. These behaviours were sometimes quite disrespectful, which again indicated that the SMAT members did not have a good relationship with each other. There was a tense atmosphere throughout this retrospective meeting. Considering all of this, there seemed to be a low level of psychological safety while observing this SMAT. The survey outcomes indicated that SMAT members of this SMAT had a low(er) level of perceived psychological safety. Therefore, it can be pointed out that this episode supports hypothesis 1. Furthermore, it can be pointed out that this episode supports hypothesis 4c, which suggested that silence behaviours would be more often displayed by SMAT members in SMATs scoring low(er) on perceived psychological safety than by SMAT members in SMATs scoring high on perceived psychological safety.

4.5.3 Summary of the Episode Analysis

The episode analysis showed that in the chosen episodes SMAT members reacted with supportive voice and collaboration behaviours to a voice behaviour displayed by one of the SMAT members in a SMAT scoring high on perceived psychological safety. This was in line with the formulation of hypotheses 1, 2, 3 and 4b. The episode analysis further showed that in the chosen episodes SMAT members reacted with (non-supportive) voice, silence or even

defensive voice behaviours to a voice behaviour displayed by one of the SMAT members in a SMAT scoring low(er) on perceived psychological safety. This was in line with the formulation of hypothesis 4c. It became clear that both the atmosphere during the retrospective meetings and the underlying relationships between the SMAT members obviously had an effect on the behaviours which were displayed, as well as on the level of psychological safety and job satisfaction in the SMATs. In SMATs 1 and 2 there was a pleasant overall atmosphere and the participants seemed to be close to, and respectful with, each other. A strong cohesion between the SMAT members and a seemingly high level of psychological safety was observed in those SMATs. This led to the display of supportive voice and collaboration behaviours by SMAT members, which seemingly had a positive effect on their level of job satisfaction. On the contrary, in SMATs 6 and 8 there was a tense atmosphere and the participants seemed not to be close with each other. This was for example illustrated by the fact that the SMAT members were sitting quite far apart from each other. There was no cohesion between the SMAT members resulting in a seemingly low(er) level of psychological safety. This led to the display of (non-supportive) voice, silence and defensive voice behaviours by SMAT members. Sometimes, these behaviours were quite disrespectful.

Considering all this, a striking difference in the displayed behaviours by SMAT members in the different types of SMATs may be identified. SMAT members in SMATs scoring high on perceived psychological safety displayed a lot of voice and collaboration behaviours, whereas SMAT members in SMATs scoring low(er) on perceived psychological safety displayed a considerable amount of silence behaviours.

The statistical analyses in this research, through which the data was explored, pointed to non-significant results and did not support the hypotheses which were formulated in Chapter 2. This was among others the result of the use of a relatively small sample size. As a consequence, the outcomes of the statistical analyses need to be treated cautiously. The episode analysis, however, pointed to other, very interesting, outcomes, as support was found for hypotheses 1, 2, 3, 4b and 4c. Hence, the outcomes of the episode analysis indicated the opposite for most of the hypotheses in comparison to the outcomes of the statistical analyses. This may have strengthened the contents of this research.

5. Discussion

This chapter presents the discussion of this thesis. Firstly, the theoretical contribution is presented which is followed by the practical contribution.

5.1 Theoretical Contribution

In terms of theoretical contribution, this thesis aimed to enrich literature on psychological safety, while measuring the concept in an innovative way by combining video observations and surveys. This was expected to provide new insights on psychological safety (O'Donovan et al., 2020; Hoozeboom et al., 2021), and in particular on psychological safety in SMAT members. Furthermore, this thesis aimed to enrich literature on psychological safety in an agile setting. Hennel et al. (2021) argued that psychological safety is a critical success factor for agile teams, which is among others likely to stimulate their team cohesion. This research aimed to explore whether this finding could also be applied to SMATs. To continue, this thesis also aimed to enrich literature on psychological safety and its relationship to job satisfaction in an agile setting. Spector (1997) argued that psychological safety is regarded as one of the circumstances to cause higher levels of job satisfaction. This research aimed to explore whether this finding could also be applied to SMAT members.

Firstly, the positive relationship between perceived psychological safety and observed psychological safety was explored. The quantitative analysis did not provide evidence to support this relationship. However, the qualitative analysis, via the episode analysis, seemed to support this positive relationship. It clearly showed that SMAT members who expressed to feel a high level of psychological safety during a retrospective meeting also were observed to act accordingly, for example by displaying (supportive) voice and collaboration behaviours. The episode analysis furthermore showed that SMAT members who expressed to feel a low(er) level of psychological safety during a retrospective meeting were observed to act accordingly, for example by displaying (non-supportive) voice, silence and even defensive voice behaviours.

This was contrary to the findings of Hennel et al. (2021) and Sherf et al. (2021). The studies noted that the perception of SMAT members during a retrospective meeting would match with the observed behaviours of the SMAT members. This was not the case in this research, as the surveys and video observations did appear not to measure the data consistently. This may be a consequence of the use of a relatively small sample size in this research. As a result,

the outcomes of the statistical analyses need to be treated cautiously, because they are likely to have given a distorted picture. This may explain the discrepancy between perceived and observed psychological safety in this research.

Then, observed psychological safety and the display of collaboration behaviours were positively linked to job satisfaction. The quantitative analysis did not provide evidence to support these relationships. However, the qualitative analysis, via the episode analysis, seemed to support these positive relationships. The episode analysis showed that, while displaying a lot of supportive types of behaviours, there was a smooth collaboration between SMAT members in SMATs with a seemingly high level of psychological safety. This seemed to have a positive effect on the level of job satisfaction of the SMAT members. This is in line with Kim et al. (2020), who suggested that enthusiasm on the job may cause high(er) levels of job satisfaction in SMAT members. In contrast, in SMATs with a seemingly low(er) level of psychological safety smooth collaboration was not observed that much. Also, SMAT members in those SMATs barely showed enthusiasm in their displayed behaviours. This was also in line with the findings of Kim et al. (2020). Therefore, based on the episode analysis, it can be argued that psychological safety, indeed, is regarded as one of the circumstances to cause higher levels of job satisfaction in SMAT members, as Spector (1997) proposed.

To continue, observed voice behaviours were expected to have a positive link to job satisfaction in hypothesis 4a. Both the quantitative and qualitative analyses did not support this expectation. To illustrate, the episode analysis did not give the indication that SMAT members were acting as if their level of job satisfaction had increased, after the display of a voice behaviour in the retrospective meetings. They, for example, did not necessarily show more enthusiasm on the job after the display of a voice behaviour, as was proposed by Kim et al (2020).

Furthermore, it was expected that SMAT members with a higher perception of psychological safety would speak up more frequently in a retrospective meeting than SMAT members with a low(er) perception of psychological safety. To illustrate, Sherf et al. (2021) noted that there was a positive relationship between the perception of psychological safety in a SMAT and the voice-usage of SMAT members. Both the quantitative and qualitative analyses did not support this expectation. To illustrate, the episode analysis showed that there was hardly any

difference in the amount SMAT members used their voice in SMATs scoring high and scoring low(er) on perceived psychological safety.

However, striking was that the episode analysis showed that there actually was a clear difference in the way SMAT members used their voice in the different SMATs. To illustrate, in SMATs scoring high on perceived psychological safety SMAT members used their voice mostly in a normal and friendly tone. They let each other finish their sentences and any criticism was brought gently, constructively and with respect. In SMATs scoring low(er) on perceived psychological safety, however, SMAT members used their voice in a much louder and sometimes even unfriendly tone. They were interrupting each other quite often, and SMAT members even yelled at each other at some moments. Criticism was brought very directly and sometimes even an aggressive tone and disrespectful words were used.

The difference in the way SMAT members used their voice in the different types of SMATs was interesting to observe. The episode analysis clearly showed that SMAT members in SMATs with a seemingly high level of team cohesion used their voice in a much more positive way than SMAT members in SMATs with a seemingly low(er) level of team cohesion. This may well have been the result of a higher level of psychological safety in those SMATs. Therefore, based on the episode analysis, it can be argued that psychological safety, indeed, is a critical success factor for SMATs, and is likely stimulate the level of team cohesion in SMATs, as Hennel et al. (2021) proposed.

Lastly, it was expected that SMAT members with a low(er) perception of psychological safety would display silence behaviours more often than SMAT members with a high perception of psychological safety. To illustrate, Sherf et al. (2021, p. 118-119) noted that 'a lack of psychological safety should trigger silence to avoid threats and punishments and move away from negative stimuli.' The quantitative analysis did not provide evidence to support this relationship. However, the qualitative analysis, via the episode analysis, did find support for this relationship. It clearly showed that SMAT members in SMATs scoring low(er) on perceived psychological safety displayed silence behaviours more often than SMAT members in SMATs scoring high on perceived psychological safety. To illustrate, they were more often looking bored, looking at their laptops or acting indifferent. Also, they were more often not involved in a discussion, or were staring in the distance.

5.2 Practical Contribution

In terms of practical contribution, this thesis aimed to raise managers and consultants' awareness concerning the importance of both psychological safety and job satisfaction in organizations working agile. This research showed that a working environment with a high level of psychological safety in SMATs seems to have positive effects for the SMAT members. The video observations and episode analysis showed that in SMATs with high levels of psychological safety there was more cohesion between the SMAT members than in SMATs with low(er) levels of psychological safety. This led to the display of supportive types of behaviours, a pleasant working atmosphere and a smooth collaboration between the SMAT members. Furthermore, it appeared that SMAT members in SMATs with high levels of psychological safety used their voice in a much more supportive and respectful way than SMAT members in SMATs with low(er) levels of psychological safety. The effects of a high(er) level of psychological safety in SMATs seemed to have a positive effect on the level of job satisfaction of the SMAT members.

Therefore, managers could use the findings of this thesis to become more familiar with the potential positive effects of psychological safety and its sub-dimensions in their organizations. Moreover, they could learn from the insights on the dynamics in self-managing (agile) teams, which are extensively analyzed in this research. To illustrate, the episode analysis showed how SMAT members behaved differently in different types of SMATs. Once managers know more about the benefits of psychological safety, and about the dynamics in SMATs, they should be better able to organize the organizational structure and culture of their organizations, so that the working environment, working atmosphere and the wellbeing of their employees may improve.

To illustrate, the episode analysis showed that in retrospective meetings often is focused on the achievement of targets or the team performance. These meetings could, for example, be arranged differently, so that there is also a moment for SMAT members to discuss things like personal well-being, individual problems or to shortly celebrate the achievement of targets. In this way, some more positive and non-business related aspects are added to the retrospective meetings. Through this, managers could let SMAT members feel more valued,

which is likely to stimulate their positivity and the level of psychological safety during the retrospective meeting (Lehmann-Willenbrock et al., 2017).

Furthermore, the video observations indicated that SMAT members were more involved in retrospective meetings with a shorter duration than in ones with a longer duration. To illustrate, SMAT members in retrospective meetings with a longer duration were observed to display more negative types of behaviours, like looking bored or indifferent. Also, the longer a retrospective meeting lasted, the more any criticism was brought in a non-constructive way. This is in line with the findings of Tilahun et al. (2017). The display of these types of behaviours negatively impacted the atmosphere during the retrospective meetings and seemingly also the level of psychological safety in the SMATs. Therefore, managers are advised to organize retrospective meetings with a maximum duration of 45 minutes. This is to prevent the display of negative types of behaviours by SMAT members out of boredom during the retrospective meetings. Furthermore, this is likely to stimulate the level of involvement of SMAT members during the retrospective meetings. As a result, the level of psychological safety of SMAT members during retrospective meetings may increase.

Consultants who read this thesis could use its findings in their advice to managers and organizations to stimulate them to think and/or act proactively regarding the creation of a working environment with a high level of psychological safety. To illustrate, they can prescribe managers to implement training programs in their businesses aimed at the creation of a working place with a high(er) level of psychological safety. These training programs could, for example, focus on how SMAT members can professionalize their behaviours during teamwork or retrospective meetings. This sounds obvious, however, the episode analysis showed that a considerable number of SMAT members did not always behave in a supportive or respectful way during the retrospective meetings. Sometimes, there were a lot of interruptions, voices were raised and even aggressive language was used. Education and training on how to behave more professionally during teamwork or retrospective meetings could definitely have a positive impact on the behaviours of SMAT members. As a result, the level of psychological safety of SMAT members during teamwork or retrospective meetings may increase.

Lastly, consultants who read this thesis could advice to implement video observations as a method of self-reflection for SMAT members. This would enable them to reflect on their own behaviours displayed during retrospective meetings. In a recent paper, O'Donovan et al.

(2020, p. 2), was suggested that ‘the use of observational techniques may offer insights into team psychological safety that the teams themselves are not fully aware of.’ This means that, while using video observations as a method of self-reflection, SMAT members may become better aware of the behaviours which they display during retrospective meetings, and the effects of it, for example on the level of psychological safety in the SMAT. In this way, SMAT members are confronted with their own behaviours and from this they can learn. This is likely to stimulate the display of supportive types of behaviours by SMAT members, which as a result may have a positive effect on the level of psychological safety during the retrospective meetings.

6. Limitations & Future Research

As in all research, there are some limitations in this thesis. A quite new mixed methods type of research was conducted by combining video observations with surveys. This type of research was time-consuming and therefore a relatively small sample size was used for analysis. If a larger sample size had been used, different insights and more convincing results could have been found.

Furthermore, this thesis explored the concept of psychological safety on the individual level, as it studied the relationship between psychological safety and job satisfaction while focusing on SMAT members. Future research on the concept of psychological safety could focus on the team and/or organizational level. Teams consist of several individuals, and therefore the insights this research provides can be used as starting point for future analysis of psychological safety in teams or even in organizations.

Moreover, future research could explore the relationships between psychological safety and other concepts on the individual level, like job performance, job meaningfulness, product owner performance and turnover intention. On the team level, future research could explore the relationships between psychological safety and, for example, team learning behaviour, team resilience, team adaptability and conflict management.

To continue, the data used in this research was obtained by OBCC and only focused on one prominent Dutch service organization working agile. More interesting findings could have been found when data on several (international) organizations acting in multiple branches and working agile was used. Also, the surveys taken by OBCC may not have captured all relevant

aspects on the of concepts psychological safety and job satisfaction. The questions in the surveys were to some extent framed into a certain direction and this may have prevented that all relevant insights on the concepts were found.

Then, 'only' two coders were involved in the coding process of this research. The involvement of more coders (earlier on in the coding process) could have improved the quality of coding. This is because comparing and discussing about different observations provided a lot of new insights from which both coders learned and afterwards were able to improve their coding capabilities. In this research, the coders were not able to reach a sufficient score (>0.8) on the Cohen's Kappa statistic for the eight observed retrospective meetings. Furthermore, during this research some improvements were made in the codebook of O'Donovan et al. (2020). Therefore, the outcomes of this thesis may not be fully reliable as different versions of the codebook were used in the coding process.

Lastly, a suggestion for further improvement of the codebook of O'Donovan et al. (2020) is to also include the subcategory 'evading confrontation' in the behavioural category 'silence behaviours'. In the latest version of the codebook, this subcategory is only included in the behavioural category 'defensive voice behaviours'. The episode analysis showed a situation in which two SMAT members suddenly started looking at their laptops in order to evade a confrontation. They pretended to be distracted by something on their laptops, but it was clear that they faked this in order to prevent getting involved in the confrontation. During this situation, they did not use their voice, neither were they looking bored, indifferent or scared. Therefore, it was difficult to decide which behavioural category had to be coded in this situation. The inclusion of subcategory 'evading confrontation' to behavioural category 'silence behaviours' would have given a solution to this issue. As example could then be added 'suddenly starts looking at laptop, in order to evade confrontation'.

The reader is advised to take all the limitations into account while interpreting the outcomes and conclusion of this research.

7. Conclusion

This thesis aimed to find an answer on the question how observed psychological safety is related to job satisfaction in SMAT members. In order to find out, several hypotheses were formulated in the theoretical background. A mixed methods type of research was conducted,

with a combination of video observations and surveys, in order to explore the concepts psychological safety and job satisfaction in an agile setting. The analysis started with an exploration of the data while conducting a frequency count and a statistical analysis. The statistical analysis did not provide support for the hypotheses which were formulated, which was among others the result of the use of a relatively small sample size. Therefore, a qualitative episode analysis was conducted in an attempt to strengthen the contents of this research.

The episode analysis revealed several interesting findings. Firstly, it seemed that SMAT members who expressed to feel a high level of psychological safety during a retrospective meeting were also observed to act accordingly. Similarly, it seemed that SMAT members who expressed to feel a low(er) level of psychological safety during a retrospective meeting were also observed to act accordingly. Hence, there seemed to be a link between perceived and observed psychological safety. Secondly, it seemed that while displaying a lot of supportive types of behaviours, there was a smooth collaboration between SMAT members in SMATs with a high level of psychological safety. This seemed to have a positive effect on the level of job satisfaction of the SMAT members. Hence, there seemed to be a link between observed psychological safety in SMATs and the level of job satisfaction of SMAT members. Thirdly, striking was that the episode analysis seemed to show that there was no difference in the amount SMAT members used their voice in different SMATs, but that there actually was a clear difference in the way SMAT members used their voice in different SMATs. Hence, there seemed to be a link between observed psychological safety in SMATs and a supportive use-of-voice by SMAT members. Lastly, it seemed that SMAT members in SMATs scoring low(er) on perceived psychological safety displayed silence behaviours more often than SMAT members in SMATs scoring high on perceived psychological safety. Hence, there seemed to be a link between a low(er) level of perceived psychological safety in SMATs and the display of silence behaviours by SMAT members.

Considering in particular the outcomes of the qualitative episode analysis, this research has found evidence for a positive relationship between observed psychological safety in SMATs and the level of job satisfaction of SMAT members. Managers and consultants can use the insights and findings of this thesis to improve the working environment, working atmosphere

and the level of psychological safety of SMAT members in organizations working in an agile setting.

8. Acknowledgements

I would like to thank my first supervisor Dr. Lara Carminati who provided guidance and support during this research. Her flexibility, positivity, knowledge and constructive feedback helped a lot while writing this thesis. Next, I would also like to thank my second supervisor Dr. Desirée van Dun for giving the opportunity to contribute to a larger research conducted by the Organizational Behaviour, Change Management & Consultancy research group at the University of Twente. Her involvement, positivity and supportive feedback definitely strengthened the content of this thesis. Finally, I would like to thank my 2nd coder for her time, help, support and valuable insights during the coding process of this thesis.

9. References

- Bagshaw, D., Lepp, M., & Zorn, C. R. (2007). International Research Collaboration: Building Teams and Managing Conflicts. *Conflict Resolution Quarterly*, 24(4), p. 433-447.
- Brief, A.P., & Weiss, H.M. (2002). Organizational Behaviour: Affect in the Workplace. *Annual Review of Psychology*, 53, p. 279-307.
- Carmeli, A. (2007). Social Capital, Psychological Safety and Learning Behaviours from Failure in Organizations. *Long Range Planning*, 40, p. 30-44.
- Cockburn, A. (2005). *Crystal Clear - A Human-Powered Methodology for Small Teams*. Addison-Wesley.
- Cooney, R. (2004). Empowered Self-Management and the Design of Work Teams. *Personnel Review*, 33, p. 677-692.
- De Jong, A., De Ruyter, K., & Lemmink, J. (2004). Antecedents and Consequences of the Service Climate in Boundary-Spanning Self-Managing Service Teams. *Journal of Marketing*, 68, p. 18-35.
- Dikert, K., Paasivaara, M., & Lassenius, C. (2016). Challenges and Success Factors for Large-Scale Agile Transformations: a Systematic Literature Review. *Journal of Systems and Software*, 119, p. 87-108.
- Detert, J. R., & Burris, E. R. (2007). Leadership Behaviour and Employee Voice: Is the Door Really Open? *Academy of Management Journal*, 50(4), p. 869-884.
- Donaldson S. I., & Grant-Vallone E. J. (2002). Understanding Self-Report Bias in Organizational Behaviour Research. *Junior Business Psychologist*, 17, p. 245-60.
- Edmondson, A. C. (1999). Psychological Safety and Learning Behaviour in Work Teams. *Administrative Science Quarterly*, 44, p. 350-383.

- Edmondson, A. C., Dillon, J. R., & Roloff, K. S. (2007). Three Perspectives on Team Learning: Outcome Improvement, Task Mastery and Group Process. *Academy of Management Annals*, 1, p. 269-314.
- Fernandez, D. J., & Fernandez, J. D. (2008). Agile Project Management: Agilism versus Traditional Approaches. *Journal of Computer Information Systems*, 49(2), p. 10-17.
- Fetters, M. D., Curry, L. A., & Creswell, J. W. (2013). Achieving Integration in Mixed Methods Designs – Principles and Practices. *Health Research and Educational Trust*, 48(6), p. 2134-2156.
- Field, A. (2018). *Discovering Statistics Using IBM SPSS Statistics*, 5. Sage Publications, Inc.
- Hackman, J. R. (2002). *Leading Teams: Setting the Stage for Great Performances*. Boston, MA: Harvard Business Press.
- 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), p. 11-30.
- Hoppock, R. (1935). *Job Satisfaction*. Harper and Brothers, New York, p. 47.
- Jarrett, M., & Liu, F. (2016). Zooming With. *Organizational Research Methods*, 21(2), p. 366-385.
- Joshi, A., Kale, S., Chandel, S., & Pal, D. K. (2015). Likert Scale: Explored and Explained. *British Journal of Applied Science & Technology*, 7(4), p. 396-403.
- McKnight, P. E., & Najab, J. (2010). Mann-Whitney U Test. *Corsini Encyclopedia of Psychology*, 1(1), p. 73.
- Meyer, P. (2015). *Agility Shift, Creating Agile and Effective Leaders, Teams and Organizations*. Taylor & Francis Group, New York, p. 202.
- Hoogetboom, M. A. M. G., Saeed, A., Noordzij, M. L., & Wilderom, C. P. M. (2020). Physiological Arousal Variability Accompanying Relations-Oriented Behaviours of Effective Leaders: Triangulating Skin Conductance, Video-Based Behaviour Coding and Perceived Effectiveness. *The Leadership Quarterly*, DOI: <https://doi.org/10.1016/j.leaqua.2020.101493>.
- Humphrey, S. E., Nahrgang, J. D., & Morgeson, F. P. (2007). Integrating Motivational, Social, and Contextual Work Design Features: A Meta-Analytic Summary and Theoretical Extension of the Work Design Literature. *Journal of Applied Psychology*, 92, p. 1332-1356.
- Kaliski, B.S. (2007). *Encyclopedia of Business and Finance*, 2. Thompson Gale, Detroit, p. 446.
- Khalid, A., Butt, S. A., Jamal, T., & Gochhait, S. (2020). Agile Scrum Issues at Large-Scale Distributed Projects: Scrum Project Development at Large. *International Journal of Software Innovation*, 8(2), p. 85-94.
- Kim, M. J., & Kim, B. J. (2020). Analysis of the Importance of Job Insecurity, Psychological Safety and Job Satisfaction in the CSR-Performance Link. *Journal of Sustainability*, 12, p. 3514.
- Lê, J. K., & Schmid, T. (2022). The Practice of Innovating Research Methods. *Organizational Research Methods*, 25(2), p. 308-336.

Lehmann-Willenbrock, N., Chiu, M. M., Lei, Z., & Kauffeld, S. (2017). Understanding Positivity Within Dynamic Team Interactions: A Statistical Discourse Analysis. *Group & Organization Management*, 42(1), p. 39-78.

Leroy, H., Dierynck, B., Anseel, F., Simons, T., Halbesleben, J. R. B., & McCaughey, D. (2012). Behavioural Integrity for Safety, Priority of Safety, Psychological Safety and Patient Safety: A Team-Level Study. *Journal of Applied Psychology*, 97, p. 1273-1281.

Liu, S., Hu, J., Li, Y., Wang, Z., & Lin, X. (2014). Examining the Cross-Level Relationship between Shared Leadership and Learning in Teams: Evidence from China. *The Leadership Quarterly*, 25, p. 282-295.

Millikin, J. P., Hom, P. W., & Manz, C. C. (2010). Self-Management Competencies in Self-Managing Teams: Their Impact on Multi-Team System Productivity. *The Leadership Quarterly*, 21, p. 687-702.

Newman, A., Donohue, R., & Eva, N. (2017). Psychological Safety: A Systematic Review of the Literature. *Human Resource Management Review*, 27(3), p. 521-535.

O'Donovan, R., Van Dun, D. H., & McAuliffe, E. (2020). Measuring Psychological Safety in Healthcare Teams: Developing an Observational Measure to Complement Survey Methods. *BMC Medical Research Methodology*, 20(1), p. 1-36.

Schein, E. H., & Bennis, W. (1965). *Personal and Organizational Change Through Group Methods*. New York: 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), p. 114-148.

Spector, P.E. (1997). *Job Satisfaction: Application, Assessment, Causes and Consequences*. Thousand Oaks, CA, Sage Publications, Inc.

Srivastava, A., Bhardwaj, S., & Saraswat, S. (2017). Scrum Model for Agile Methodology. *International Conference on Computing, Communication and Automation*, p. 864-869.

Terrell, S. (2011). Mixed Methods Research Methodologies. *The Qualitative Report*, 17(1), p. 254-280.

Thompson, E. R., & Phua, F. T. T. (2012). A Brief Index of Affective Job Satisfaction. *Group & Organization Management*, 16, p. 201-344.

Tilahun, N., & Levinson, D. (2017). Contacts and Meetings: Location, Duration and Distance Traveled. *Travel Behaviour and Society*, 6, p. 64-74.

10. Appendix

Appendix 1: Items of Psychological Safety Scale by Detert et al. (2007)

Individual Psychological Safety

During this past meeting, it felt safe for me to:

- 1 make suggestions

- 2 give my opinions
- 3 speak up

Appendix 2: Meeting Psychological Safety Scale & Cronbach's Alpha

Scale: Meeting Psychological Safety Scale

Case Processing Summary

		N	%
Cases	Valid	51	100,0
	Excluded ^a	0	,0
	Total	51	100,0

a. Listwise deletion based on all variables in the procedure.

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
,961	,961	3

Appendix 3: Items of Job Satisfaction (BIAJS) Scale by Thompson et al. (2012)

Job Satisfaction

- 1 I find real enjoyment in my job
- 2 I like my job better than the average person
- 3 Most days I am enthusiastic about my job
- 4 I feel fairly well satisfied with my job

Appendix 4: Job Satisfaction Scale & Cronbach's Alpha

Scale: Job Satisfaction Scale

Case Processing Summary

		N	%
Cases	Valid	51	100,0
	Excluded ^a	0	,0
	Total	51	100,0

a. Listwise deletion based on all variables in the procedure.

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
,862	,871	4

Appendix 5: SPSS Output on Check for Normality (Shapiro-Wilk Test)

Tests of Normality

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Perceived Psychological Safety	,261	51	<,001	,761	51	<,001
Job Satisfaction	,126	51	,042	,957	51	,065
Voice Behaviours	,119	51	,068	,919	51	,002
Silence Behaviours	,237	51	<,001	,797	51	<,001
Collaboration Behaviours	,078	51	,200 [*]	,957	51	,063
Observed Psychological Safety	,093	51	,200 [*]	,961	51	,088

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Appendix 6: SPSS Output on Spearman's Rho Correlation Test for Hypothesis 1

Correlations

			Perceived Psychological Safety	Observed Psychological Safety
Spearman's rho	Perceived Psychological Safety	Correlation Coefficient	1,000	-,003
		Sig. (2-tailed)	.	,985
		N	51	51
	Observed Psychological Safety	Correlation Coefficient	-,003	1,000
		Sig. (2-tailed)	,985	.
		N	51	51

Appendix 7: SPSS Output on Pearson's R Correlation Test for Hypothesis 2

Correlations			
		Observed Psychological Safety	Job Satisfaction
Observed Psychological Safety	Pearson Correlation	1	,100
	Sig. (2-tailed)		,486
	N	51	51
Job Satisfaction	Pearson Correlation	,100	1
	Sig. (2-tailed)	,486	
	N	51	51

Appendix 8: SPSS Output on Spearman's Rho Correlation Test for Hypothesis 2

Correlations				
			Perceived Psychological Safety	Job Satisfaction
Spearman's rho	Perceived Psychological Safety	Correlation Coefficient	1,000	,210
		Sig. (2-tailed)	.	,139
		N	51	51
	Job Satisfaction	Correlation Coefficient	,210	1,000
		Sig. (2-tailed)	,139	.
		N	51	51

Appendix 9: SPSS Output on Pearson's R Correlation Test for Hypothesis 3

Correlations			
		Collaboration Behaviours	Job Satisfaction
Collaboration Behaviours	Pearson Correlation	1	,061
	Sig. (2-tailed)		,671
	N	51	51
Job Satisfaction	Pearson Correlation	,061	1
	Sig. (2-tailed)	,671	
	N	51	51

Appendix 10: SPSS Output on Spearman's Rho Correlation Test for Hypothesis 4a

Correlations

			Voice Behaviours	Job Satisfaction
Spearman's rho	Voice Behaviours	Correlation Coefficient	1,000	-,037
		Sig. (2-tailed)	.	,796
		N	51	51
	Job Satisfaction	Correlation Coefficient	-,037	1,000
		Sig. (2-tailed)	,796	.
		N	51	51

Appendix 11: SPSS Output on Mann-Whitney U Comparative Test for Hypothesis 4b

Mann-Whitney Test

Ranks

	PPS - High versus Low(er)	N	Mean Rank	Sum of Ranks
Voice Behaviours	High	40	24,05	962,00
	Low(er)	11	33,09	364,00
	Total	51		

Test Statistics^a

	Voice Behaviours
Mann-Whitney U	142,000
Wilcoxon W	962,000
Z	-1,787
Asymp. Sig. (2-tailed)	,074

a. Grouping Variable: PPS - High versus Low(er)

Appendix 12: SPSS Output on Mann-Whitney U Comparative Test for Hypothesis 4c

Mann-Whitney Test

Ranks

PPS - High versus Low(er)		N	Mean Rank	Sum of Ranks
Silence Behaviours	High	40	24,08	963,00
	Low(er)	11	33,00	363,00
	Total	51		

Test Statistics^a

Silence Behaviours	
Mann-Whitney U	143,000
Wilcoxon W	963,000
Z	-1,778
Asymp. Sig. (2-tailed)	,075

a. Grouping Variable: PPS - High versus Low(er)