Organisational characteristics' influence on psychological safety

A mixed method study about organisational characteristics' influence on psychological safety within agile software teams.

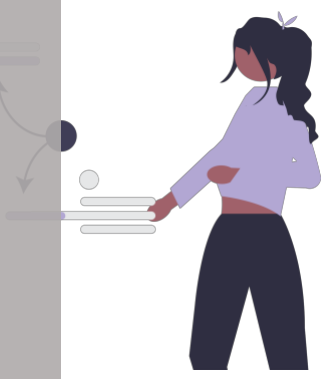
Master thesis

February – June 2023
IDA 10 – cs-23-ida-10-03

Group members:
Casper Anisimow 20186191
Emilie Louise de Bang 20186012
Magnus Damgaard Pagh 20183275

Supervisor:
Adam Alami

Thesis: 83 pages
Appendix: 19 pages
Repository: GitHub



Title:

Organisational characteristics' influence on psychological safety – A mixed method study about organisational characteristics' influence on psychological safety within agile software teams.

Theme: PS within agile software teams

Focus: Theoretical analysis

Project Group: cs-23-ida-10-03

Project Semester: 10th semester, Spring

Students:

Casper Baymler Anisimow

canisi18@student.aau.dk (20186191)

Emilie Louise de Bang

edeban18@student.aau.dk (20186012)

Magnus Damgaard Pagh

Mpagh18@student.aau.dk (20183275)

Supervisor: Adam Alami

Number of pages: 83 pages

Date of completion: 9th of June 2023

Abstract:

Psychological safety (PS) in the workplace is essential for a well-functioning agile software team. Agile software teams have common organisational characteristics that influence PS; however, it is unknown how.

Utilising a mixed method approach with 12 interviewees and 468 survey participants, we found that organisational characteristics guided by social interactions, have an influence on PS compared to structural aspects which do not. PS cannot be reduced to a single concept; it encapsulates actions performed by both managers and the team. Team autonomy was found to be subsumed by variables also influencing PS, which indicates autonomy's position in an organisational context.

Our findings show the importance of team members and managers collaborating, in the process of improving their work environment towards higher PS. We propose multiple implications in regard to maintaining a psychologically safe agile software team. These include proactive leadership, establishing guidelines for good argumentation, as well as collaboration to create a culture of openness in

Summary

Motivation: Over the past two decades, the approach to work and collaboration within software development has experienced significant changes, transforming from a plan-driven approach to mainly adapting methods from the agile paradigm (Oyibo & Gabriel, 2020). The agile paradigm introduces the value of *Individuals and interactions over processes and tools* establishing a more human centred approach, to achieve well-functioning teams. Psychological safety (PS) is relevant when aiming to achieve well-functioning agile software teams.

Aim: The use of social agile practices influences PS with a bidirectional relation, where PS also improves the use of set practices (Hennel & Rosenkranz, 2020). Agile practices are connected to a specific framework such as Scrum and Kanban, while the agile paradigm itself consists of values and principles to establish a culture with an agile mindset. Organisational characteristics focus on structural and social aspects of an organisation representing set values and principles. However, it is unknown how organisational characteristics affect PS in an agile work environment. The objective of this thesis is to identify relevant team's organisational characteristics and how they affect PS. Based on these results, we will present salient suggestions to achieve higher PS by improving organisational characteristics. We propose the following research question:

> How do team's organisational characteristics influence psychological safety in an agile software team?

Methods: In answering our thesis we utilised a mixed method approach with 12 interviewees and 468 survey participants. The interviewees were analysed through a thematic analysis, where they were asked to share experiences from their work life, focusing on team and management collaboration as well as psychological safety. The survey questions were analysed with linear regression based on the qualitative results elicited from the interviews, and includes themes and questions represented in the survey, such as: enduring teams, openness in communication, management, and clear decision process.

Results: We found that organisational characteristics guided by social interactions, have an influence on PS compared to structural aspects which do not. Openness in communication,

supportive management, clear decision process and clear feedback structures have a positive influence on PS when improved. Our study also shows that PS cannot be reduced to a single concept, it encapsulates actions performed by both managers and the team. Organisational characteristics should be built upon social interaction and constitute the creation of a culture that accommodates mistakes and ideas within the team. Lastly, we found that team autonomy was subsumed by variables also influencing PS, which indicates autonomy's position in an organisational context. Our findings show the importance of team members and managers collaborating, in the process of improving their work environment towards higher PS.

Conclusion: We propose multiple implications in regard to maintaining a psychologically safe agile software team. Management should be proactive and lead by example. Management and the team should collaborate in establishing guidelines for what makes an argument good and how to evaluate and discuss ideas within the team. They should also, in collaboration, create a culture that promotes openness in communication, and create the foundation for a workplace where team members can voice their opinion and admit to their mistakes without blame.

Table of content

1.0 Introduction

Over the past two centuries, the approach to work and collaboration has experienced significant changes (Oyibo & Gabriel, 2020). In firms, the traditional tayloristic model of organising work has become demoted and no longer aligns with the norms of the current paradigm. Due to an increasingly globalised world combined with an accelerating evolution of technology, organisations are facing more frequent and intense changes, where the teams need to adapt and respond accordingly (Oyibo & Gabriel, 2020). As a result, there is a growing emphasis on the physical and mental health of employees, both in practical and theoretical terms (Hastuti & Timming, 2021; Guest, 2017).

One theoretical concept that has garnered increased attention in the past two decades is psychological safety (PS), which has evolved through three distinct waves since 1965 (Newman et al., 2017). The origins of this concept can be traced back to the field of organisational change, where Edgar Schein and Warren Bennis (1965) were the first to describe PS. Specifically, they defined it as the individual's experience of confidence in managing change and the feeling of security in a changing environment (Bennis & Schein, 1965).

After Schein and Bennis' description, PS had received little to no attention in the realm of organisational literature. However, with the evolving nature of work and the growing emphasis on employees' well-being, PS has re-emerged as a pertinent concept. The second wave emerged in the early 1990s, where William Kahn revitalised the concept by focusing on how an individual perceives their ability to express themselves without fear of negative repercussions to their self-image, status, or career (Kahn, 1990, p. 708). Additionally, Kahn emphasised the relationship between PS and an individual's interpersonal connections (Kahn, 1990).

Kahn's work marked a turning point in the conceptualisation and understanding of PS. Since then, scholars have explored the role of PS in various aspects of organisational studies, including team performance, innovation, and employees' well-being (Newman et al., 2017, p. 522). As such, PS has become a crucial area of research for organisations seeking to create supportive and effective work environments (Newman et al., 2017).

In the late 1990s, a third wave of research on PS emerged. Amy Edmondson advocated for the examination of PS on a social or team level. Through her extensive work on the topic, Edmondson developed a definition for PS as "*a shared belief held by members of a team that the team is safe for interpersonal risk taking*" (Edmondson, 1999, p. 354). In addition to providing a new definition for PS, Edmondson also introduced a 7-item instrument to measure a team's PS, which is widely used when researching the topic today (Newman et al., 2017).

In recent decades, significant changes have occurred not only in the approach to organising work within organisations, but also in the manner in which software development processes are organised. In the 1960's the software development process was heavily influenced by the methods used in physical product development (Haraty & Hu, 2018). This approach is known as the plan-driven paradigm. The plan-driven paradigm had limitations when used in a changing and complex environment. This led to the development of the agile paradigm, which emerged in the early 2000s with the publication of the Agile Manifesto (Beck et al., 2001). The agile paradigm is based on an iterative approach that emphasises the core values of working software, customer collaboration, responsiveness to change, and individual and interactions (Beck et al., 2001).

The agile approach to organising software development processes has gained widespread embracement across numerous companies (Marnada et al., 2022). When comparing the values and principles of agile software development with the concept of PS, several parallels can be drawn. The first value of the agile manifesto encompasses individuals and interactions over processes and tools, which corresponds with PS's goal of minimising interpersonal risk taking among team members (Edmondson, 1999). PS provides a means for agile software teams to prioritise individuals. The 12th principle highlights the importance of team reflexivity and the ability to adapt team behaviour as needed, which aligns with Edmondson's definition of PS at the team level. It also addresses the possibility of team characteristics and leaders to influence the interpersonal risk taking while feeling safe (Edmondson, 1999).

Agile methods have proven to have a positive effect on PS in a work environment, but the agile principles have multiple implementations through frameworks as Kanban and Scrum (Peeters et al., 2022). The agile manifesto values empowerment of the team, and the ability to trust the team in their performance. The agile manifesto contains principles that do not dictate a predetermined route, and thus the interpretation may also differ (Biehler et al., 2022).

Based on our initial research it is unknown what effect, among others, team size, years of experience, and the number of years working together, communication and decisions processes have on the PS in an agile team. To address this issue, we propose the following research question, trying to understand how a team's organisational characteristics impact PS in a given environment, as well as how management and the team can foster a more psychological safe work environment:

> How do team's organisational characteristics influence psychological safety in an agile software team?

The research questions will be answered based on a two-phased data collection done by Adam Alami and his colleagues (Alami et al., 2023a). Phase one consists of written data collected by email, which creates a guideline for a qualitative interview based on practitioner experience. The interviews focus on the interviewees' experiences in a team in regard to PS's influence on software quality. The analysis will be sorted according to themes extracted from the phase one data. These themes will then be translated into testable hypotheses, to test whether the theme has an influence on an agile software team's PS, provided empirical data exists to support the test of the hypotheses. Based on the data from phase two collected by Alami and his colleagues, the hypothesis will be tested utilising linear regression modelling. The quantitative data will provide the possibility of generalising the results from the qualitative part of the analysis. The survey includes different questions to capture the different aspects of PS, and the participants' conceptions of the current team they work in (Alami et al., 2023a).

We have found by reviewing the data that multiple interviewees focus on the promotion and demotion of PS within the team. They related their experiences of PS in regard to how their leaders or team handled missteps, engagement and pressure when working in the team. When reviewing the qualitative data we identified different implementations of the agile principles, where the team composition also varied. The interviewees mentioned that they experienced a drop in collaboration when team members changed teams. The element of collaboration is a part of Edmondson's scale for PS, and it therefore suggests that the organisational characteristics of a team are relevant to investigate further. The quantitative data includes a variety of different variables in relation to a team's organisational characteristics and allows

for a more generalised analysis of the research question. An overview of the utilised control variables can be seen in section *3.4.2 Control variables*.

2.0 Related work

The following section will synthesise existing literature within the field of PS and agile software development teams. The selection of literature was decided upon through searching different academic databases: Elicit, Web of Science, Google Scholar, EBSCOHost, Science Direct, and ProQuest. The keywords utilised when searching for papers were: *Agile Software Development*, *Psychological Safety* and *Team Characteristics.* These were combined during searches in different combinations. We also utilised the method of snowballing, which was accomplished by looking at previously identified relevant articles, and then including their relevant sources in the literature review.

Recent research articles are interested in the interaction between agile software development and PS. Studies generally showcase that agile software development practices have a positive effect upon the team members PS (Buvik & Tkalich, 2021, Marder et al., 2021, p. 8f). Marder et al. (2021) looked at using an agile work environment to facilitate increased PS, by staging an intervention with a group of students taking advanced management courses. They found that PS was increased through the agile intervention, where higher levels of PS positively influenced team performance and team learning (Marder et al., 2021, p. 8f). For these positive effects to be experienced, a leader figure and framework must be in place (Marder et al., 2021, p. 8f). Similarly, as Marder et al.'s research, Gren et al. (2019) evaluated the effects of psychological training upon agile software developers. Twelve teams were given psychological training, and a follow up interview was conducted after 1,5 months. Generally, the results showed a positive effect on both psychological wellbeing as well as effectiveness (Gren et al., 2019, p. 4).

Lenberg & Feldt (2018) provided research into the area of norm clarity, psychological safety, and their effects on job satisfaction and performance. Utilising survey data from 38 different software teams, consisting of 217 different individuals, linear regression was performed to test the effects of team norm clarity on PS (Lenberg & Feldt, 2018, p. 1). Results indicated that the effects of team norm clarity upon job satisfaction and performance were more powerful than that of PS, supporting Marder et al.'s hypotheses (Lenberg & Feldt, 2018, p. 5ff). Further research was conducted into the area of agile team performance.

In contradiction to Marder et al.'s findings of needing clearly delineated roles in the framework to experience the positive effects of PS, Buvik & Tkalich (2021) investigated how work characteristics influence PS and performance in agile software teams. The data was collected through surveys, with 236 respondents, representing 43 different development teams in Norway. The variables measured were team autonomy, task interdependence, role clarity, psychological safety, team reflexivity, team performance, and a number of control variables. Team autonomy was found to have a positive influence on PS; however, task interdependence and role clarity were not statistically significant, contradicting the findings of Marder et al. In addition, PS was found to positively affect team performance (Buvik & Tkalich, 2021, p. 7).

Supporting Buvik and Tkalich findings, Peeters et al. (2022) examined the effects of working agile on team performance and engagement through surveying 97 different agile development teams (p. 61). Confirming the findings of Buvik & Tkalich; an agile work environment was confirmed by statistical analysis to positively correlate with PS. PS was in turn found to positively correlate with team performance, as well as team engagement (Peeters et al., 2022, p. 70f). These results were partially supported by Gustavsson (2022) who conducted a survey with 201 members of agile teams working for three different companies, examining the effects of PS upon team performance. They also explored how the variables for team performance and inter-team coordination interacted numerically (Gustavsson, 2022, p. 1). Utilising structural equation modelling, team performance was found to positively correlate with PS. However, no correlation was found between inter-team coordination and team performance (Gustavsson, 2022, p. 10f).

Besides Gustavsson's (2022) research of PS's influence on team performance, Hennel & Rosenkranz (2021) also propose a model, based on PS and agile software practices effect upon team performance. The study was conducted with a mixed-methods approach, where semi-structured interviews were combined with field notes and internal documentation, studying three different cases (Hennel & Rosenkranz, 2021, p. 16, p. 22f). Hennel & Rosenkranz found three main results, firstly, the usage of social agile practices positively affects the performance for a team. Secondly, increased usage of social agile practices positively affects PS in a dynamic context, where an increase in PS also improves the use of social agile practices. Thirdly, PS enables and enforces a positive effect on the performance of social agile practices (Hennel & Rosenkranz, 2021, p. 15f). In short, social agile practices, despite their temporary

reduction in performance, ultimately results in greater performance (Hennel & Rosenkranz, 2021, p. 21).

Kakar (2018) examined the interaction between team cohesion and PS, and their effects upon knowledge sharing within the context of software teams (Kakar, 2018, p. 258). The study ran over a four-year period and involved 332 respondents, who had worked on 34 different software projects. Team cohesion and PS were used as independent variables, while knowledge sharing was dependent. Furthermore, control variables were employed. Knowledge sharing and PS had a positive relationship, while the relationship between team cohesion and knowledge sharing was parabolic, both very low and very high levels of team cohesion were found to impact knowledge sharing negatively (Kakar, 2018, p. 264).

Certain areas of PS have not yet been researched thoroughly. One of these areas is the PS of online teams. Khanna & Wang studied PS in the context of agile software teams, and their use of scrum retrospectives when working remotely. They conducted research based on a recording of a software company's research and development team. The conclusion of the study was that PS was entirely possible, provided that the participants utilised the online tools available for them to communicate (Khanna & Wang, 2022, p. 47f). Prior meetings conducted in-person improved the well-being of the participants by allowing for feelings of connection. In addition, the tools that may influence PS included video, audio, checkmarks, polls and votes, emojis, digital boards, and breakout rooms (Khanna & Wang, 2022, p. 39f). A deeper understanding of the relation between PS and online meeting, are however still unknown and as such further research was suggested.

As Khanna and Wang; Holten et al., (2015) evaluated the communication methods utilised in the context of agile software development by utilising a mixed-methods approach (Holten et al., 2015, p. 273f.). Methods of importance were found to be daily stand-up meetings, colocated offices, pair programming, and sprint reviews & retrospectives. Ideally, a compromise should be reached between indirect and direct communication (Holten et al., 2015, p. 288). As such, it would be relevant to investigate the effect of PS when working colocated and remotely in an agile software team.

Not only communication methods, but also organisational culture, may influence PS. Thorgren & Caiman (2019) conducted research into the implementation of agile environments in two different companies. They localised three different challenges: Attitudes towards

inclusiveness, openness in communication, and perceptions of and trust in collective responsibility (Thorgren & Caiman, 2019, p. 34ff). This was coupled with the concept of PS, where a psychologically safe team had an easier time transitioning to an agile work environment, compared to one lacking PS. Furthermore, it also implies the importance of committed management, if one wishes to get his team to work agile, one should be willing to embrace agile completely, without hanging on to previous management practices (Thorgren & Caiman, 2019, p. 34ff). The effects of organisational cultures were also researched by Gupta et al., (2019), with focus on the culture within IT departments and its effect on agile practices. This was accomplished by surveying 189 IT department managers (Gupta et al., 2019, p. 18f). Their findings supported those presented by Thorgren & Caiman; the specific cultures of certain cultures could either help or hinder agile practices. For example, hierarchical culture was negatively related to both social and technical agile practices (Gupta et al., 2019, p. 21f).

One possible avenue of interest in the context of agile teams is the team's diversity. Verwijs & Russo (2023) analysed the impact of diversity upon a team, its PS and performance. Using a quantitative model and over 1000 respondents, they concluded that diversity in age had a positive impact upon teamwork within the context of agile teams, unlike gender, role, and cultural diversity. In fact, it was found that gender diversity increased possible conflicts in the workplace. PS was found to result in more effective teamwork and less conflict, confirming prior expectations. However, high PS did not mediate the effects of increased conflict from gender diversity, showing that even psychologically safe environments may struggle with these challenges (Verwijs & Russo, 2023, p. 12f). Diversity might also exist with a group maturity level. Gren et al., (2017) conducted ten interviews with coaches from different companies and received survey data from 66 group members from four companies to examine the relationship between group maturity and agile teams. Results showcased great overlap between agile principles and group maturity, in fact, well-functioning agile groups could be considered psychologically mature groups (Gren et al., 2017, p. 28f). Group maturity is a measurement based on the Group Development Questionnaire, designed by Wheelan in 1994, to measure a groups maturity based on dependencies, collaboration, effectiveness, and conflicts (Buzaglo and Wheelan, 1999, p. 110f).

So far, while much of the research has been conducted into specific areas, such as either diversity or remote teams, a few researchers have explored team size effects on agile software teams. Dingsøyr et al., (2022) conducted research to create a unified model (ATEM) for

teamwork effectiveness. ATEM is composed of five components: Shared Leadership, Peer Feedback, Redundancy, Adaptability, and Team orientation. Along with three coordinating mechanisms: Shared Mental Models, Mutual Trust, and Communication (Dingsøyr et al., 2022, p. 36). The model was developed as a universal theory to improve teamwork effectiveness on a team level, among other with focus on team size. While it is an extensive model for teamwork effectiveness, it does not have PS as one of its main, isolated features. Instead, indirect elements of PS are included under the coordinating mechanism Mutual Trust. It does not consider PS as a dependent variable but elaborates for elements possibly related to PS through their understanding of Mutual Trust.

Based on the literature found in the related work section, little empirical research has been conducted into the organisational characteristics and their implications for an agile software team's PS. Researching PS has the potential to contribute to better organised workplace practices, and thereby improve the PS within agile software teams. The related work section opens for the opportunity to research multiple different organisational characteristics' influence on PS, such as team size, team maturity, communication methods, as well as team norms and values.

3.0 Methods

The current thesis was based on data collected by, and given to us by Adam Alami our supervisor, and his colleagues, in relation to their studies about PS's influence on software quality in agile teams. The study was performed together with Mansooreh Zahedi and Oliver Krancher. This thesis's research question was therefore based on relevant topics found in the qualitative data. The quantitative data consisted of a variety of control variables as well as other aspects, which made it applicable for answering our research question.

3.1 Philosophy of science

By combining social constructivism, by Berger & Luckmann (2004) with critical rationalism, by Karl Popper (1996), our ontology combined both the individual elements as well as the more structural guidelines. We understood reality as a dualistic perspective between an individual understanding of it, but where set reality also included common elements outside the individual's awareness (Popper, 1996, p. 36; Berger & Luckmann, 2004, p. 87). Social constructivism contributed to an understanding of the individual viewpoint when at work, and their opinion regarding their workplace, hence on a micro level. Critical rationalism accommodated the structural and organisational perspectives on teams and companies, and therefore included a meso and macro level understanding of the structures and PS at the interviewees' workplace.

Our epistemological stance was founded in a combination of social constructivism and critical rationalism as well. We understood science and theory as a part of a context and related to the current social creation of reality. The individual opinions were therefore understood in the settings it was experienced (Berger & Luckmann, 2004, 105). The research of teams' organisational characteristics and its influence on PS was evaluated and either falsified or not, based on the acceptance of tested hypotheses (Popper, 1996, 73f). A hypothesis which could not be falsified would be accepted as a creation of knowledge until proven otherwise.

3.1.1 Sensemaking

When combining two different perspectives on ontology and epistemology it seemed relevant to offer a third perspective which could serve as a mediator between them and provided a

perspective on how to understand organisations. The perspective was built upon Karl E. Weick, who argued the need for investigating an organisation with a combination of social constructivism and realism. He gave the following argumentation based on the subjects of study:

> *"People who study sensemaking oscillate ontologically because that is what helps them understand the actions of people in everyday life who could care less about ontology.* [...] *If people have multiple identities and deal with multiple realities, why should we expect them to be ontological purists?"*
>
> (Weick, 1995, p. 35).

In this thesis we adapted Weick's (1995) view upon studying people in organisations. We understood organisations as more than just structure, hierarchy, and united goals, but as processes, relations and sensemaking (Weick, 1995). The main part of our understanding relied on the sensemaking of people within the organisation. The concept consisted of, at least, seven elements. 1. Sensemaking is influenced by a social process which impacts and is impacted by the way individuals make sense (Weick, 1995, p. 38-43). 2. Every sensemaking process is based on our self and perception of others' identities (Weick, 1995, p. 18-24). 3. Sensemaking happens retrospectively which implies that the way individuals assess an experience is influenced by the individual's understanding of the experience (Weick, 1995, p. 24-30). 4. Sensemaking is focused on the clues which involve perceived sensations leading to sensemaking (Weick, 1995, p. 49-55). 5. Sensemaking is an ongoing process, meaning it involves a dynamic evaluation of perceived senses and individuals constantly thrown into new processes (Weick, 1995, p. 43-49). 6. Individuals' way of sensemaking is driven by plausibility more than accuracy, meaning that individuals are searching for the more plausible truth within a context (Weick, 1995, p. 55-61). 7. Enactment is a central concept in sensemaking that means whenever individuals do something, they are creating their surroundings (Weick, 1995, p. 30-38).

Thereby the concept of sensemaking provided a perspective to understand more complex phenomena of organisational characteristics and how individuals were sensemaking of their experienced reality, while it also facilitated both social constructivism and critical rationalism.

3.2 Data collection

Prior to conducting the interviews, Alami gathered 20 emails from the interviewees to assess their level of PS in their workplace. This was achieved by utilising Edmondson's 7-item scale, which was designed to quantitatively measure an individual's perception of the degree to which their team allows for the tolerance of mistakes, openness, rejection, risk taking, inclusivity, and recognition of unique skills (Edmondson, 1999). Figure 1 illustrates the scale constructed by Edmondson for her research.

ID	Edmondson's Scale to Measure Psychological Safety	Direction
PS1	"If you make mistakes on my team, is it often held against you"	Negative
PS2	"Members of my team can bring up problems and tough issues"	Positive
PS3	"People on my team sometimes reject others based on the ideas they propose"	Negative
PS4	"It is safe to take a risk (e.g., experiment with a new technology, propose initiatives, raise difficult issues, disclose own knowledge gaps) on my team"	Positive
PS5	"It is difficult to ask other members of this team for help"	Negative
PS6	"No one on my team would deliberately act in a way that undermine my efforts"	Positive
PS7	"Working with members of my team, my unique skills and talent are valued and utilised"	Positive

Figure 1 – Depicts the 7-item scale constructed by Edmondson, to measure PS (Edmondson, 1999). The column direction is dependent on the question's phrasing, and suggests the effect it would have on PS.

Alami et al.'s (2023a) research was based on a mixed methods approach where data was collected both qualitatively and quantitatively in two distinct phases. By employing a mixed methods approach, they aimed to leverage the strengths of both qualitative and quantitative research paradigms. The qualitative phase enabled a nuanced understanding of complex phenomena, such as team members' perceived assessment of PS in their work environment (Creswell, 1999). The quantitative phase provided a statistical framework to further explicate the findings, complementing an in-depth understanding (Creswell, 1999).

The selection of a mixed methods approach in this thesis was underpinned by its underlying ontology and epistemology. By adapting a partially social constructivist ontology and utilising qualitative analysis, this thesis aimed to provide a comprehensive understanding of individuals' perceptions and sensemaking within their working environment in relation to PS. The qualitative findings were subsequently translated into a quantitative framework by synthesised

hypotheses, thereby enabling statistical testing of these hypotheses in accordance with the principles of falsification in critical rationalism's epistemology.

The data collected from the interviews in the first phase were subsequently processed and transformed into a questionnaire, used in the data collection for the second phase.

3.3 The qualitative phase

The goal of the qualitative phase was to gain an understanding of the interviewees perceptions of their work environment and how that influences their PS. Industry relationships were used to establish contact with 12 of the interviewees where the last 8 interviewees were found by snowballing from the already collected interviewees, see Table 1 (Alami et al., 2023b, p. 111:5f). Snowballing means that the collected interviewees had agreed to forward the interview invitation in their relevant network, for the researchers to gain more interviewees outside their intermediate network (Bryman, 2014, p. 424). The interviews were conducted as semi structured interviews, and the questions were separated into three sections. Firstly, an introduction was used to set the stage for the following questions, with focus on the interviewee, their team, and their level of PS. Secondly, the core section which was used to gain information in regard to the research question. Lastly, the probing section, a section with questions to dig deeper into the responses. (Alami et al., 2023b, p. 111:5f; see Appendix 1). The interviewees were encouraged to explain examples and their thoughts gained through their own experience working in agile teams, this provided higher data quality and a deeper understanding of the interviewees' experiences (Alami et al., 2023b, p. 111:6).

The interviews were conducted online due to geographic reasons and were collected in 2022 from January till November. The interviews were transcribed using otter.ai, and afterwards adjusted by Alami (Alami et al., 2023b, p. 111:6; See GitHub Repository). Table 1 shows the characteristics of each interviewee. This includes how they were sampled, their role when working, the years of experience, what method of agile they utilise, the type of projects they work on, whether they have high or low PS at work, and lastly their country. For the current thesis we will only be using 12 interviews, the interviewees marked with grey will not be included in this thesis. The interviews were selected after our preliminary reading, where we found specific interviews that better fit the RQ. This was decided since we did not conduct the interviews ourselves, and therefore the original research goal differed from our thesis goal.

#	Sampling	Role	Exp	Method	Project type	PS	Country
P1	Convenience	Lead Software Developer	8	XP	CRUD Apps	High	Netherlands
P2	Convenience	Snr Software Engineer	14	Scrum	Asset Mgmt. software	High	UK
P3	Convenience	Software Tech Lead	15	Kanban	Web applications	High	Poland
P4	Convenience	Snr Software Engineer	14	Scrum	Custom software	High	India
P5	Snowballing	Snr Software Engineer	14	Scrum	Digitalization of services	High	India
P6	Snowballing	Snr Software Engineer	8	Scrum	Custom software	High	India
P7	Convenience	Lead Software Engineer	14	Scrum	Medical data analysis	High	UK
P8	Snowballing	Snr Software Engineer	8	Scrum	Marketing Mgmt. software	High	UK
P9	Convenience	Snr Software Engineer	5	Scrum	Cryptocurrency platform	High	India
P10	Convenience	Snr Software Engineer	13	Scrum	Communication software	High	Germany
P11	Snowballing	Software Engineer	5	Kanban	Asset Mgmt. software	High	UK
P12	Snowballing	Tech Lead	11	Scrum	Insurance products	Low	India
P13	Snowballing	Software Engineer	7	Scrum	Online insurance services	High	Germany
P14	Convenience	Software Engineer	7	Scrum	Robotics software	High	Sweden
P15	Convenience	QA Analyst	5	Scrum	Telecom Software	High	Australia
P16	Convenience	Snr QA Engineer	15	Scrum	Web applications	High	Germany
P17	Convenience	Snr QA Engineer	10	Scrum	Banking application	High	UK
P18	Convenience	QA Analyst	10	Scrum	Digitalization of services	High	Germany
P19	Snowballing	Snr QA Engineer	6	Scrum	Robotics software	High	Sweden
P20	Snowballing	Snr QA Engineer	8	Scrum	Ecommerce platform	High	UK

Table 1 – Overview of the Interviewees from phase 1. For the current thesis we will only be using 12 interviews, the interviewees marked with grey will not be included in this thesis (Alami et al., 2023b, p. 111:6)

3.3.1 Analysis of the qualitative data

Jonny Saldaña's two cycles of coding qualitative data method were utilised to analyse the qualitative data (Saldaña, 2021). The first cycle consisted of a categorising process, to highlight chunks of the data relevant for the research question. To be concise, when we were coding the data, we chose three strategies within which the data could be coded: In vivo, process, and emotional.

The In Vivo coding strategy was used to capture the interviewees own wording in the codes, and to understand the meaning inherent in the experiences the interviewees share (Saldaña, 2021, p. 105f). When we used the In Vivo coding strategy, we gained an individual insight and adapted that understanding to a more generalised, meso/macro level context. Through utilising

the process strategy, we could highlight the interviewees' expressed everyday interactions between their surrounding colleagues (Saldaña, 2021, p. 143f). This would contribute to the interpersonal understanding of the interviewees PS in their work environment. Emotional coding would emphasise the emotional responses of the interviewees when recounting their workplace experiences (Saldaña, 2021, p. 125). Emotions were influenced by the social context and social experience and should therefore not be understood as an individual element but as an expression for an individual's perception of the social context it was a part of (Bo & Jacobsen, 2015, p. 21; Saldaña, 2021 p. 125). Understanding an interviewees emotion in the different examples would provide the research with an understanding of the individuals perception of their workplace, and also an insight into their examples that would not be given by just utilising In Vivo coding.

This thesis would employ the qualitative analysis software program, Nvivo, to organise and categorise relevant phrases for the research question (See GitHub Repository). The initial coding of the interview was conducted through collaborative effort among the thesis group members to promote open discussion and streamline interpretation. Each phrase was assigned a descriptive but concise title and a longer description to capture the thought processes and reflections. At the first cycle, the codes were then sorted into three categories, In Vivo, Process, and Emotion, based on the differentiation criteria provided by Saldaña (2021). The second cycle of coding was also carried out in collaboration, wherein the In Vivo and Process codes were inspected to determine their relevance to the research question and meaningfulness of their titles. The In Vivo and Process codes were then grouped based on selected organisational characteristics, and subsequently, themes were formed by grouping the elements derived from the code. The Emotion codes were likewise reviewed to ensure relevance and meaningful titles and were discussed in relation to the derived elements. For an example of how elements were formed, see Image 1, which highlights an element made up of multiple codes of qualitative data. In Appendix 2 it is possible to see the construction of every element.

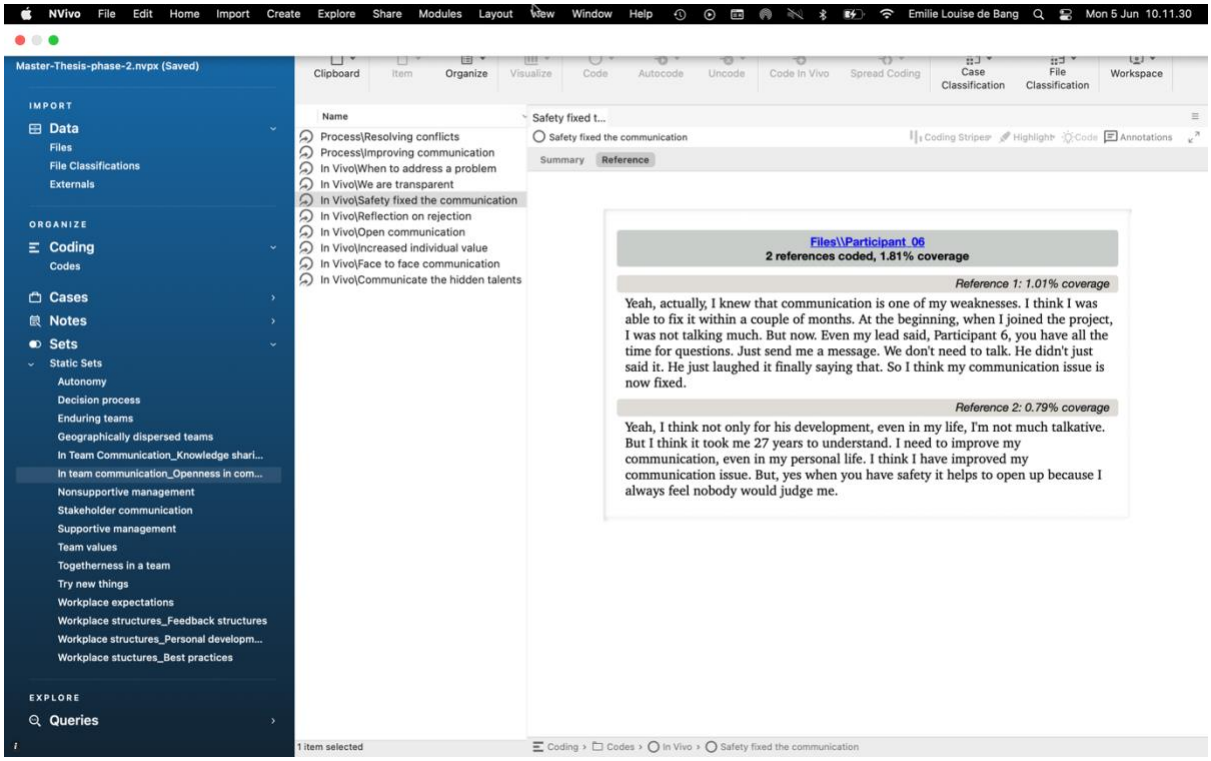


Image 1 – The Nvivo interface. The image illustrates how Nvivo was utilised to code the qualitative data. Displayed here is the element “Openness in communication”, with its associated codes.

To establish the hypotheses, we utilised the Causation Coding to highlight causations within the newly constructed groups. The causation coding built upon the results in the qualitative analysis, and the findings of causal beliefs between the different attributions (Saldaña, 2021). An overview of the hypotheses, and their relation to themes and the problem area of PS, can be found in Figure 2.

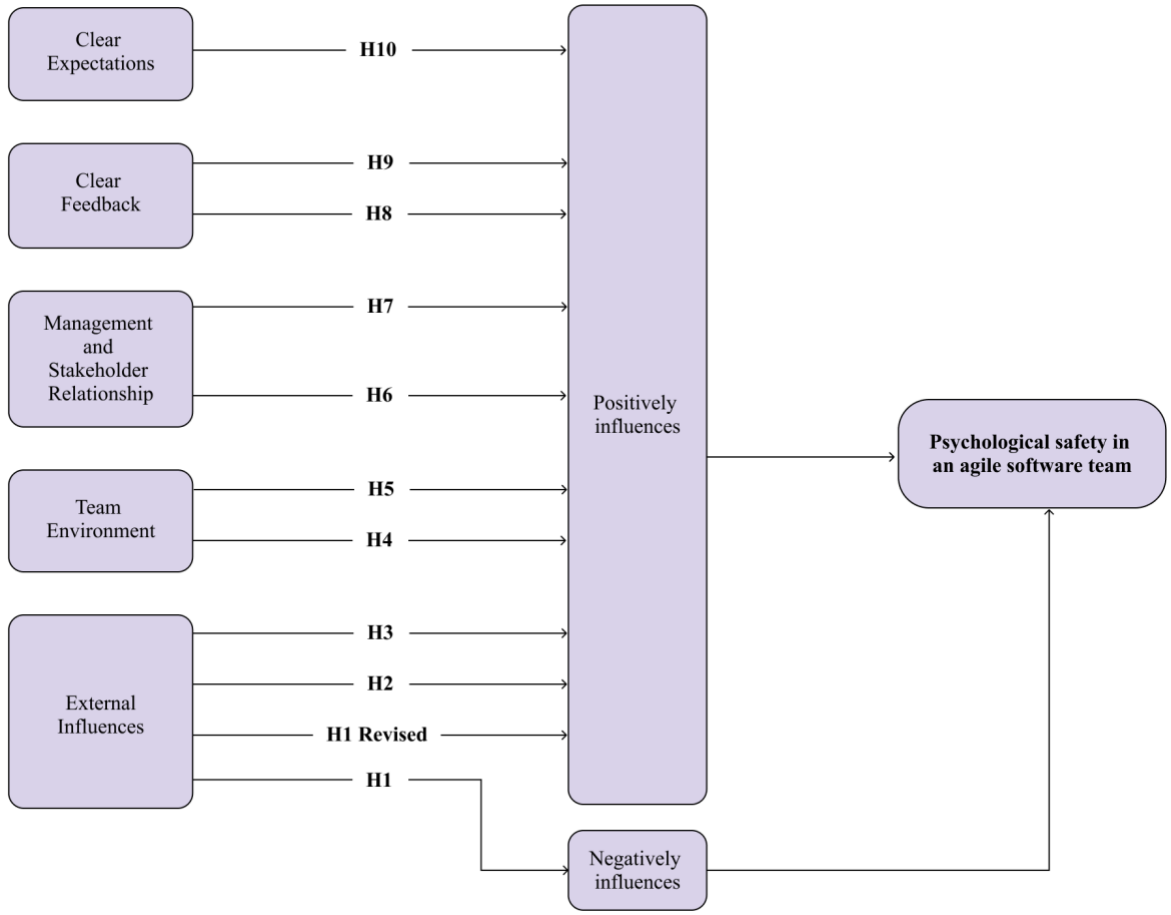


Figure 2 – Displays the direction of our causation coding, depending on the individual hypothesis and its relation to PS.

An overview of a few of the main organisational characteristics in regard to our interviewees, can be seen in Table 2. The characteristics were found within the interviews, searching for context specific variables related to the interviewees current team. The characteristics are related to the control variables found in the quantitative data, as well as built upon patterns found in the analysis of the interviews.

In Table 2 the different characteristics of the interviewees can be seen. The characteristics were found in the interviews. The characteristics were chosen to highlight the interviewees organisational characteristics of their team. Meaning that the characteristics are context dependent on the specific team.

#	Role type	Exp	Method	Project type grouped	Enduring teams	Country	Size	In-house or Outsourced	Multi teams	Org. structure	Autonomy
P1	Lead	8	XP	CRUD Apps	Unknown	Netherlands	4 - 5	Outsourced	Unknown	Unknown	Very self-managed
P5	Senior	14	Scrum	Digitalization of services	1,5 years	India	4	Outsourced	Yes	Top-down	Partially managed
P6	Senior	8	Scrum	Custom software	5 Months	India	9	Outsourced	Yes	Top-down	Partially self-managed
P7	Lead	14	Scrum	Medical data analysis	2,5 years	UK	3 - 6	Outsourced	Yes	Top-down	Partially self-managed
P8	Senior	8	Scrum	Marketing Mgmt. software	1 year	UK	10	In-house	Yes	Top-down	Partially self-managed
P9	Senior	5	Scrum	Cryptocurrency platform	6 months	India	10	In-house	Yes	Top-down	Very self-managed
P10	Senior	13	Scrum	Communication software	2 years	Germany	9	Outsourced	Yes	Independent	Very self-managed
P12	Lead	11	Scrum	Insurance products	9 months	India	16 - 18	In-house	Yes	Top-down	Very managed
P13	Engineer	7	Scrum	Online insurance services	8 months	Germany	20 (13)	In-house / outsourced	Yes	Top-down	Partially self-managed
P15	Analyst	5	Scrum & Kanban	Telecom Software	Over 2 years	Australia	12	In-house	Yes	Top-down	Very self-managed
P17	Senior	10	Scrum	Banking application	2 years	UK	22	Outsourced	Unknown	Flat	Partially self-managed
P18	Analyst	10	Scrum	Digitalization of services	1 > year	Germany	12	Outsourced	Yes	Top-down	Partially self-managed

Table 2 – Overview of the organisational characteristics within the interviewees' team, extracted during the analysis of the interviews.

3.4 The quantitative phase

Phase two had the goal of collecting data that could be used to test the hypothesis developed based on the qualitative data from phase one, done by Alami and his colleagues. The survey was also performed to expand the empirical scope (Alami et al., 2023b, p. 111:7). The survey consisted of 43 questions with multiple sub-questions to measure, among others, PS (See GitHub Repository). They performed a pilot test on 20 participants to test the validity of the survey questions and the planned scales, this was done by conducting an exploratory factor analysis, where both the Cronbach's alpha and the factor loading was evaluated. They made some modest changes to the questions that had low standard deviations, Cronbach's alpha, or factor loadings (Alami et al., 2023b, p. 111:8). The respondents were sampled through the research marketplace Prolific, which allowed for purposive sampling. To accomplish the purposive sampling, additional screening questions were utilised, to filter out irrelevant respondents (Alami et al., 2023b, p. 111:8). This resulted in reducing the number of respondents from 914 to 480. Afterwards they conducted further quality control before starting the analysis. The quality control was utilised to counter possible problems with survey marketplace with bots and dishonest respondents by implementing different mitigation strategies which allowed for a filtering of respondents.

We used Python in the analysis of the data, where different libraries will be used to perform the linear regression and scales, among other Pandas and Pingouin (See GitHub Repository). After reviewing the raw data, multiple variables were checked to ensure the quality of the data. The sample ended up with 468 participants after removing participants that did not complete the questionnaire, answered the attention questions wrong, or people that did not stay consistent in their answers. The question regarding their role was asked both in the beginning and end of the survey, where participants who did not answer the same at both questions were removed. Table 3 shows an overview based on some of the variables included in the questionnaire, the questions provide some characteristics of the data. Table 3 also depicts some of the control variables used in the linear regression tests. This will be elaborated further below.

Survey characteristics						
Country	UK: 136	US: 104	Portugal: 60	Italy: 28	Other: 140	
Gender	Male: 378	Female: 90				
Education	Bachelor's degree: 248	Master's degree: 192	PhD: 20	Other: 8		
Role	SW Engineer: 144	QA: 87	Snr. SW Engineer: 83	Developer: 54	Tech Lead: 51	Other: 49
SW Dev. Exp	< 3 years: 82	3 - 5 years: 147	6 - 8 years: 86	9 - 11 years: 44	> 11 years: 109	
Team Size	< 5 members: 63	5 - 7 members: 213	8 - 9 members: 77	10 - 12 members: 67	> 12 members: 48	
Multiple Teams	Yes: 309	No: 159				
Sourcing Model	In-house: 302	Outsourced: 125	Other: 3	Mix: 38		

Table 3 – Data characteristics for Phase 2. Our sample consisted of 468 participants.

3.4.1 Construction of scales

To answer the hypothesis, multiple scales were constructed to capture the different measurements. The following section would show the construction of the PS scale, as well as the open communication scale, while the rest of the scale's composition could be found in Appendix 3. The composition of the scales would be based on their correlation values, factor loadings and the scales Cronbach's alpha value.

When composing a scale, it was relevant to check the correlation between the different variables, to check whether they could be used to describe the same phenomenon. In Figure 3 it was possible to see a heat map displaying the correlation between the variables for measuring PS, while Table 4 displays the factor loadings for the variables. It was obvious on both the heat map and based on the factor loadings that variable *PS_18_2* were not correlating highly with the other variables, with a factor loading of only 0,280. However, it was important to consider the theory when establishing a scale to measure a theory-based phenomenon. The Cronbach's alpha was therefore calculated for a scale both including and excluding the variable *PS_18_2*. With all seven variables included, Cronbach's alpha was calculated to have a value of 0,6894 while the value was a bit higher without the variable and reached 0,6929. However, we deemed the increase to be insufficient compared to deviating from the theory. The PS scale will therefore include all 7 variables as intended. When creating the PS scale, the values for the variables were standardised between 0 and 1, and afterward multiplied with 100 to get more intuitive numbers when analysing the results. This was done for all the created scales.

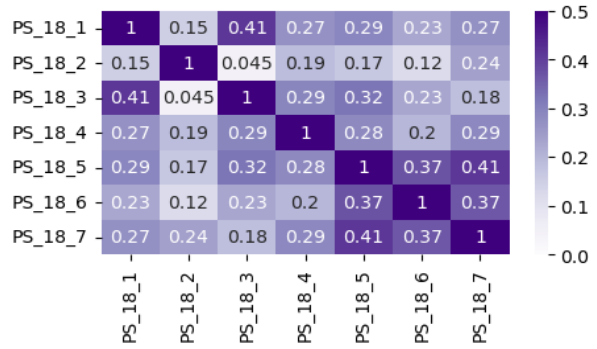


Figure 3 – Correlation values for the PS variables. The correlations values are displayed in a heat map, scaling from 0,0 to 0,5

Variable name	Factor loading
PS_18_1	0,531734
PS_18_2	0,280752
PS_18_3	0,498684
PS_18_4	0,488080
PS_18_5	0,642982
PS_18_6	0,511529
PS_18_7	0,586383

Table 4 – Illustration of the factor loadings for the PS variables

When creating the scale for Openness in communication, both the category *Openness* and *SpeakingUp* seemed to be able to explain the same outcome. Performing a correlation on those variables made it possible to see that they mainly correlated well within their own category, however variable *Openness_29_2* correlated very well with all 9 variables, see Figure 4. In Table 5, the factor loadings depicted high numbers, for all variables when tested in one factor. The Cronbach's alpha was 0,856 when all variables were included, which is assessed as a good Cronbach's alpha value. The scale was therefore constructed of all variables, despite the correlation test suggesting the possibility of two factors.

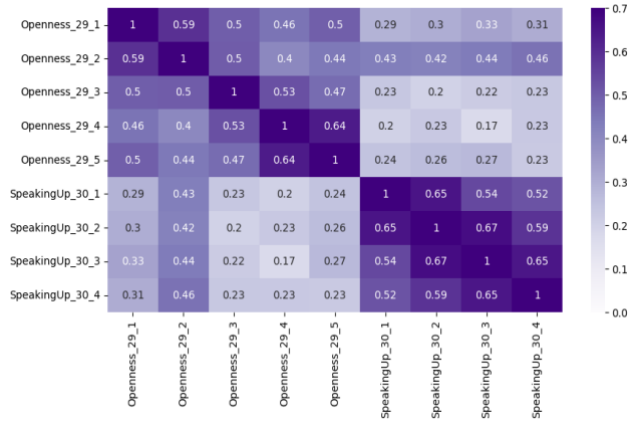


Figure 4 – The correlations values between the variables *Openness* and *SpeakingUp* are displayed in a heat map, scaling from 0,0 to 0,7

Variable name	Factor loading
Openness_29_1	0,639432
Openness_29_2	0,743430
Openness_29_3	0,551390
Openness_29_4	0,540722
Openness_29_5	0,583411
SpeakingUp_30_1	0,629194
SpeakingUp_30_2	0,675210
SpeakingUp_30_3	0,671055
SpeakingUp_30_4	0,649045

Table 5– Illustrates the factor loadings for the variables *Openness* and *SpeakingUp*

Table 6 showed an overview of what variables have been used to test each hypothesis and what element the variables should try to measure. Hypotheses 1 and 2 were tested based on only one variable each, separated into dummies where each category consisted of a Boolean variable. The rest of the hypotheses are tested based on scales, including from 6 variables up to 15 variables, depending on the scale. Table 6 also showed what type of survey instrument was used when designing the questions, whether it was a multiple-choice question, or if the Likert scale was used.

H	X value	Q Category	Questions	Survey instrument	Alpha
H1R	Colocated	Colocated	How often are the members of your team colocated?	Multiple-choice	X
H2	Enduring teams	Working Together	How long has your team been working together in your current team?	Multiple-choice	X
H3	Autonomy	Autonomy	In our team are responsible for deciding how to organise our work.	Likert Scale	0,7704
			In our team we decide how to achieve our goals.	Likert Scale	
			In our team we make the decisions regarding the technical solutions with no interferences from management or our stakeholders	Likert Scale	
			In our team we make decisions regarding the tasks' estimation	Likert Scale	
			In our team we make the decisions for changing our processes in order to improve our performance.	Likert Scale	
			In our team, we have the freedom to make decisions on architectural design decisions, choice of technology and tools	Likert Scale	
H5	Openness	Openness	People in our team are open to criticism and feedback from their peers	Likert Scale	0,8562
			People in our team welcome new ideas and initiatives put forward by their peers	Likert Scale	
			People in our team do not reject ideas based on the individual who proposed it but based on the strength and the soundness of the idea	Likert Scale	
			People in our team accept the rejection of new ideas when the rejection is based on strong arguments	Likert Scale	
			People in our team accept the rejection of ideas when they fail to convince team members with their arguments	Likert Scale	
		Speaking Up	In my team, people raise their concerns.	Likert Scale	
			In my team, people talk about problems.	Likert Scale	
			In my team, people share their opinions	Likert Scale	
			In my team, people point out quality problems	Likert Scale	
H6	Management	Ownership of psychological safety	Our leadership is resolute about psychological safety in our team.	Likert Scale	0,9499
			Our leadership is determined to promote a work environment where people dare to take risks.	Likert Scale	
			Our leadership accepts that failure can occur when we try out new things	Likert Scale	
		Management listening to the team	Our leadership listens to our needs.	Likert Scale	
			Our leadership wants to hear about our concerns.	Likert Scale	
			Our leadership is willing to listen to our suggestions.	Likert Scale	
		Management supporting the team	Our leadership is supportive of us.	Likert Scale	
			Our leadership provides help with everything we need to deliver our current project.	Likert Scale	
			Our leadership treats us with respect.	Likert Scale	
			Our leadership supports us doing our work	Likert Scale	
		Leadership integrity (leadership behaviour)	Our leadership “walk the talk” when it comes to taking risk and accepting failure	Likert Scale	
			Our leadership “practice what they preach” when it comes to psychological safety.	Likert Scale	
			Our leadership follows through on the values of psychological safety.	Likert Scale	
			Our leadership are role models in terms of taking risks and accepting failures.	Likert Scale	
			Our leadership words are well aligned with their actions when it comes to admitting mistakes.	Likert Scale	
H8	Clear decision process	Autonomy	In our team are responsible for deciding how to organize our work.	Likert Scale	0,8264
			In our team we decide how to achieve our goals.	Likert Scale	
			In our team we make the decisions regarding the technical solutions with no interferences from management or our stakeholders	Likert Scale	
			In our team we make decisions regarding the tasks' estimation	Likert Scale	
			In our team we make the decisions for changing our processes in order to improve our performance.	Likert Scale	
			In our team, we have the freedom to make decisions on architectural design decisions, choice of technology and tools	Likert Scale	
		Collective decision making	In our team we engage in constructive discussions to make our decisions	Likert Scale	
			In our team each team member's voice counts when decisions are made	Likert Scale	
			In our team we make decisions based on the best arguments that team members contribute	Likert Scale	
			In our team we aim to reach consensus when make our decisions	Likert Scale	
H9	Clear feedback	No blame	We do not blame each other for mistakes but see them as an opportunity for improvement	Likert Scale	0,8200
			In our team, we do not blame each other for underperforming, instead we coach each other to improve	Likert Scale	
			Our leadership, including team leader and mid management, do not blame us for mistakes in our team	Likert Scale	
		Learning from mistakes	As a team, when we admit mistakes, we learn from our mistakes	Likert Scale	
			Past mistakes usually become a point of reference in our team	Likert Scale	
			When past mistakes become a point of reference in our team, we avoid similar mistakes in the future	Likert Scale	

Table 6 – An overview of the questions used to capture the x-variables, and their Cronbach’s alpha value.

3.4.2 Control variables

When we chose our control variables, this thesis focused on variables from the workplace, and limited focus on variables such as age and education. Age and education described the person outside one's work, where *Experience* and *Role* characterised a person within their workspace. The control variables were also inspired by previous literature in the field. Buvik & Tkalich (2021) similarly used team size, as well as role tenure to account for variance when trying to measure team autonomy, and other variables to explain PS (p. 5). As such, this thesis found it prudent to include similar control variables, such as *TeamSize*, *Experience*, as well as *Role,* to potentially explain some of the variance within the linear regression. In addition, gender was added as a variable, to try to confirm whether Verwijs & Russo's (2023) findings were accurate; as they found that conflict may increase with increased gender diversity, it would be interesting to check if a person's gender affects a person's level of PS. In addition, based upon the analysis in the interviews, it was also decided to include the variable *sourcing model* which indicates whether a team works with in-house software, mixed, or outsourced, software developed for others. Often conflicts occurred due to interactions with stakeholders. It was relevant to find out whether these interactions occur less frequently with in-house teams, or outsourced teams and their possible influence on PS. Finally, the number of other teams they collaborated with, within the organisation, was also included as an independent variable.

The variable *TeamSize* had a single outlier, which was removed to minimise measurement uncertainty. The variable *TeamSize* ranges from 2 – 35 members. Some of the variables utilised were gathered as categorical, meaning that utilising them unchanged in a linear regression would be inappropriate. To accomplish their use, the variables *Gender*, *Sourcing model, Experience, Role, and Multiple teams*, were recoded to be Boolean. For example, a *Gender* value of 0 represented a man, while a value of 1 represented a woman. In addition, background variables with more than one category, such as *Role*, were split into multiple sub-variables, with 0 all other categories, where 1 was representing the value of the current role. This allowed the project to utilise variables which would otherwise not be usable in the context of linear regression, where each category consists of a Boolean variable. In the linear regression models the Boolean variable should then be compared to the reference group, indicated in the tables in the analysis by (RG).

3.4.3 Linear regression

We found it suitable to perform linear regressions based on the dependent variable being interval scaled. There were five conditions that should be fulfilled before being able to consider causality: linear connection between X and Y, homoscedasticity, random sampling, no multicollinearity, and exogeneity (Dreyer & Hussain 2016, p. 297).

The variables were tested for signs of linear connection before performing the linear regression models. *Colocated* and *Enduring teams* showed signs of not fitting the linear regression modelling. However, the other X variables fit the linear regression. Multivariate linear regression was used for each hypothesis to be able to test the hypothesis more thoroughly. Multiple control variables were included in the test to overcome possible heterogeneous data, and to be surer about the causality of the relationship between X and Y. The X values were tested for heteroscedasticity, where the scale for *Openness, Management* and *Clear Decision* were found to contain heteroscedasticity; they were therefore tested with robust standard errors (Treiman, 2009, p. 237). The result for heteroscedasticity can be seen in Appendix 4. The X values for *Colocated, Enduring teams, Autonomy,* and *Clear feedback* did not show signs of heteroscedasticity and were therefore implied to showcase homoscedasticity. The data was collected by invitation, which might have affected the randomness of the sample, and therefore also the representativeness of the data. All control variables were tested for multicollinearity by performing a VIF-test. Multicollinearity indicates a problem if the VIF-test scored above the value 3 (Dreyer & Hussain 2016:297). All variables scored under the value 3 except the variable *TeamSize,* to assess the problem we conducted a correlation test between the control variables. See Appendix 5. The correlation test shows a tendency for a team's size to correlate with the categories 'QA' and 'Other' from the variable *Role*. Indicating these roles might have specific team sizes. We acknowledge the problem of multicollinearity but decided to keep *TeamSize* as a control variable. See the result for the VIF-test in Appendix 5. The following project was based on the understanding that the X variables come before the Y variable in time, and therefore that there was exogeneity. A way to ensure this could be to work with panel data, however this was not an option at the present time.

All the hypotheses were tested based on the H0, meaning that the linear regression showed whether H0 should be verified or falsified, and based on this it was possible to draw a conclusion for the shown hypothesis. Each hypothesis tested had an accompanying table,

displaying the regression values extracted from the data. The coefficients were accompanied by their standard error deviation, noted in parentheses. Some of the tables would display the robust standard errors, based on problems with heteroskedasticity. In addition, their degree of statistical significance, would be noted with *'s; a lack of a star indicated no level of significance, one star (*), indicated a significance level of 0,1, two stars (**) indicated a level of 0,05, and three stars (***) indicated a high level of 0,01.

4.0 Findings

In this section, we present the qualitative analysis of relevant elements derived from the data. Elements that are represented in the quantitative data will be followed by a testable hypothesis derived from the qualitative data. The section will be structured in five themes, constructed of correlated elements. The themes are: External influence, Team environment, Management and stakeholder communication, Clear feedback, and Company expectations. The themes and elements can be seen in Figure 5. To see the transcribed interviews, a complete list of the survey questions, our Nvivo file and the Jupyter Notebook, go to our GitHub repository: https://github.com/Emiliedebang/MasterThesis

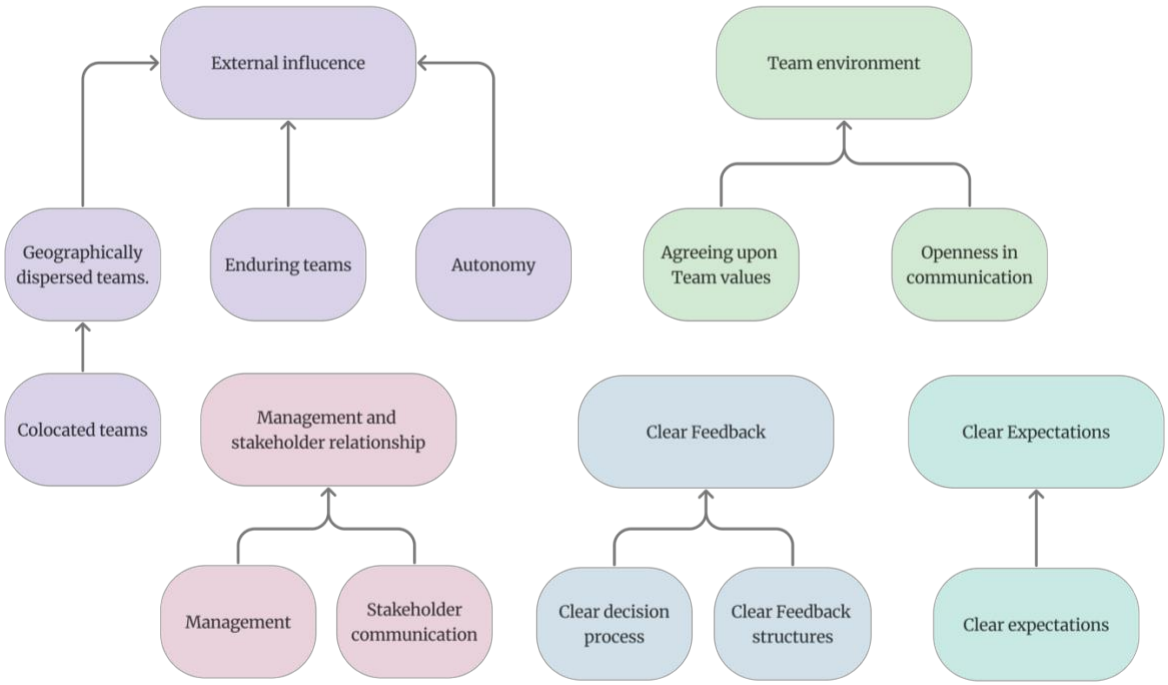


Figure 5 – Overview of themes and their correlated elements

To give an overview on how the themes and patterns codes are related to the generated hypothesis, we have composed Table 7 to show the relationship and whether it is assessed testable or not with the quantitative data.

Themes	Patterns codes	Hypothesis	Testable
External influences	Geographically dispersed teams	**H1:** Geographically dispersed teams are negatively associated with psychological safety in an agile software team.	No
		H1 revised: Colocated teams are positively associated with psychological safety in an agile software team.	Yes
	Enduring teams	**H2:** Being an enduring team is positively associated with greater psychological safety in an agile software team	Yes
	Autonomy	**H3:** Autonomy of a team is positively associated with psychological safety in an agile software team	Yes
Team environment	Agreeing upon team values	**H4:** The existence of shared team values increases psychological safety in an agile software team	No
	Openness in communication	**H5:** Openness in communication is positively associated with psychological safety in an agile software team	Yes
Management and stakeholder relationship	Management	**H6:** Supportive management is positively associated with psychological safety in an agile software team	Yes
	Stakeholder communication	**H7:** Open communication with stakeholders is positively associated with psychological safety in an agile software team	No
Clear feedback	Clear decision process	**H8:** Clear decision processes results in higher psychological safety in an agile software team	Yes
	Clear feedback structures	**H9:** Clear feedback is positively associated with psychological safety in an agile software team	Yes
Clear expectations	Clear expectations	**H10:** Clearer expectations in the organisation is associated with improving psychological safety in an agile software team	No

Table 7 – An overview of the proposed hypotheses in relation to their pattern codes and themes, as well as their testability

4.1 External influences

The present theme encompasses external factors that were identified during the first cycle of analysis. External influences are defined as decisions or actions originating outside the team, with minimal to no influence from the team, yet they still have an impact on the team's environment. The three elements that have been identified in this theme are: Geographically dispersed teams, Enduring teams, and Autonomy.

4.1.1 Geographically dispersed teams

Geographically dispersed teams refer to teams whose members are situated in different geographical locations. When asked whether team members could bring up tough issues, participant 7 stated yes, however, how they handled them could differ. Participant 7's account exemplifies how face-to-face communication resolved a problem that was insurmountable through virtual communication, and thereby improved PS:

> *"*[...] *there were some issues related to communication in certain regards, between colleagues. And it seemed that doing remote calls between those colleagues did not help much. And usually that is resolved by having face to face conversation. Or in a group situation, where we get to know each other better by trying to understand each other what each team member is passionate about, what his or her goals are. Because if we are just having calls over the internet meetings online, it does not help with the communication skills. We also need to do it face to face"*
>
> (Participant 7)

According to the above statement, it is evident that remote communication is not always efficient in resolving certain issues, prompting participant 7 to emphasise the importance of face-to-face communication within their team. As such, it is suggested that non-remote communication may have positive effects on PS. Consequently, this statement raises concerns about the implications of a geographically dispersed team when dealing with complicated work situations. The potential impact of a geographically dispersed team is also addressed by participant 8 and participant 12. Participant 8 provides insight into their team's experience with collaborating across different time zones, as illustrated in the following quote:

> "[...] *if something went wrong, and since* [...] *most of the team work in the Indian time zone and two other people work in the UK time zone.* [...] *I think it's not that great, but at least like one or two people from our time zone should work at midnight just to make sure that, we are still working on that* (problem, red.) *and it just take time"*
>
> (Participant 8)

Participant 8 raises a critique of working across significantly different time zones, arguing that members of participant 8's team in the UK are required to work during midnight to collaborate with colleagues in India, which is a time-consuming process. Notably, participant 8's perspective is representative of a European experience of time zones.
Furthermore, participant 12 shares their experience of collaborating across time zones while working with European or American clients from India.

> *"*[…] *I can tell you the whole truth that people actually don't tell you. So what's happening is it is a colonial environment. That has not moved on. And I'll give you a short example. What happens is that the US clients or the UK clients, they work from 9am to 5pm. And people are working in Southeast Asia, we have to work according to their timezone. So we work from evening to morning 6am. That is to*

> *match with them, and we don't get paid for those extras. In some cases, when we are getting paid, we are paid $5 for the whole night"*
>
> (Participant 12)

As evidenced in the quote, participant 12 draws a comparison between their work with clients in the US or UK and a colonial environment, whereby the western clients dictate when workers in Southeast Asia are expected to work. Participant 12 highlights the example of workers in Southeast Asia having to work at night without proper compensation, reflecting a sense of inequality compared to workers in the western part of the world. Moreover, this example illustrates participant 12's perceived lack of recognition from their clients. Lack of recognition is a part of Edmondson's scale for measuring PS, where they are asked whether their unique skills are valued in their team, see Figure 1 in section *3.2 Data collection*. Both participants 8 and 12 provides insights into the challenges faced by individuals situated in geographically distant locations when working with teams in different time zones

To summarise the Geographically dispersed teams element, we identify challenges related to solving problems virtually and working across widely distributed time zones. Firstly, participant 7 asserts that physical communication is essential for a team to address complex issues that may be difficult to comprehend through virtual means. Secondly, participants 8 and 12 give their opinions on how it is to work with others in different timezones, where participant 12 expresses a feeling of not being recognized and not viewed as an equal. Hence, we propose a hypothesis for investigating the influence of Geographically dispersed teams on PS in an agile software team, formulated as follows:

> **H1:** Geographically dispersed teams are negatively associated with psychological safety in an agile software team

Hypothesis 1 aims to measure whether teams not working from the same country or in the same time zone, negatively influences the PS of the team. However due to the quantitative data not being able to capture this specific element we have chosen to revise the hypotheses and it will be based on participants 7's description of needing face to face communication, as well as the findings by Holten et al. (2015). Their study examined the ways of communication that exist within an agile team, and how communication is accomplished. Their findings assume both direct, and indirect methods of communication suggesting that a remote team may be

insufficient in its ability to communicate, due to missing a physically shared working space (Holten et al., 2015). The revised hypotheses is as follows:

> **H1 revised:** Colocated teams are positively associated with psychological safety in an agile software team

The following linear regression will be based on H1 revised. Five linear regression models were constructed and evaluated upon the variable *Colocated* teams, including control variables, see Table 8. In Model 1, the independent variable *Colocated* shows a negative influence on PS when compared to the reference group of working colocated 'Less than 1 day per week'. As an example, when a person goes from working colocated 'Less than 1 day per week', to working colocated '3-4 days a week', the PS drops with 4,04 points. This connection is however the only medium significant category of the *Colocated* variable. The control variables *Gender* and *TeamSize* are insignificant in the analysis and with little to no effect on PS. In Model 1 the adjusted R^2 value is 0,0002, meaning *Colocated*, *Gender* and *TeamSize* only are able to explain 0,02% of the deviation in PS.

In Model 2, when adding the control variable *Sourcing model*, small changes happen to the variable *Colocated* influence on PS. *Sourcing model* is significant in both categories, 'Outsourced' and 'Mixed' when referenced to the category 'In-house'. The categories have a negative impact on PS when compared to 'In-house'. Working with outsourced software rather than in-house software lowers the PS with 4,22 points. Model 2 has an adj. R^2 value of 1,9%.

In Model 3, the control variable *Experience* is added, and affects the *Colocated* variable 'Always colocated' to move from having a negative effect to having a small positive effect when stable. *Experience* itself is also significant and shows a clear pattern of how more experience influences PS positively. However, its effect on PS stabilises around 8 years where it is almost unchanged. All categories for *Experience* are significant, however the level of significance changes when all control variables are added. Model 3 is able to explain 5,6% of the deviations in PS.

	Model 1	**Model 2**	**Model 3**	**Model 4**	**Model 5**
Colocated: Less than 1 day per week – Remotely (RG)					
Colocated: 1 - 2 days per week – Hybrid	-0,7885 (1,5432)	-1,0499 (1,5316)	-0,9812 (1,5126)	-1,3547 (1,5126)	-1,3567(1,5146)
Colocated: 3 - 4 days per week – Hybrid	-4,0473** (1,8163)	-4,1506** (1,8009)	-3,9293** (1,788)	-4,1048** (1,7900)	-4,0960** (1,7972)
Colocated: Always – Colocated	-1,1433 (1,8763)	-0,7868 (1,8617)	0,0709 (1,8442)	0,1670 (1,8490)	0,1578 (1,8565)
Gender (Man = 0)	0,2789 (1,5895)	0,1086 (1,5754)	0,5425 (1,5588)	0,5136 (1,5708)	0,5202 (1,5759)
TeamSize	-0,0260 (0,1365)	0,0217 (0,1360)	0,0085 (0,1347)	0,0824 (1,1405)	0,0815 (0,1414)
Sourcing model: In-house (RG)					
Sourcing model: Outsourced		-4,2280*** (1,4288)	-3,4305** (1,4161)	-3,1244** (1,4232)	-3,1103** (1,4414)
Sourcing model: Mixed		-4,4453** (2,2242)	-4,3052* (2,1906)	-4,0320* (2,1931)	-4,0244* (2,1987)
Experience: Less than 3 years (RG)					
Experience: 3 - 5 Years			3,3774* (1,8267)	3,2426* (2,0609)	3,24661* (1,8347)
Experience: 6 - 8 Years			7,6110*** (2,0479)	6,9968*** (2,0609)	6,9953*** (2,0634)
Experience: 9 - 11 Years			6,9788*** (2,4602)	6,1303** (2,5020)	6,1261** (2,5056)
Experience: More than 12 years			7,5084*** (1,9487)	6,2485*** (2,0389)	6,2522*** (2,0420)
Role: Software developer (RG)					
Role: Software engineer				2,5229 (2,1029)	2,5168 (2,1073)
Role: Senior software engineer				5,5618** (2,4097)	5,5484** (2,412)
Role: Tech Lead				3,1693 (2,6261)	3,1502 (2,6456)
Role: QA				0,4541 (2,3385)	0,4431 (2,3472)
Role: Other				0,2081 (2,7012)	0,1966 (2,7101)
Multiple teams					0,0853 (1,3194)
Const.	79,8593	81,0684	75,8156	73,5611	73,5150
Adj. R^2	0,0002 (0,02%)	0,019 (1,9%)	0,056 (5,6%)	0,0643 (6,4%)	0,0622 (6,2%)
N	468	468	468	468	468

Table 8 – An overview of the five models tested in regard to *Colocated*. * indicates low-significant (P-value < 0,10), ** indicate medium significant (P-value < 0,05), and *** indicate high-significant (P-value < 0,01).

In Model 4, the control variable *Role* is added, which again has small changes for the *Colocated* variable, where the category '1-2 days a week' goes from having a negative impact on PS of 0,98 points, to be changes by 0,37 points, meaning it is then negative affecting PS with 1,35 points. The only category that is deemed significant for the variable *Role* is 'Senior software engineer', when moving from 'Software developer' to 'Senior software engineer ', the PS is positively influenced with 5,56 points. The model is now able to explain 6,4% of the deviation in PS. When in Model 5, the control variable *Multiple teams* are added, there are no relevant changes in the numbers, and *Multiple teams* does not influence PS in any significant way. The adjusted R^2 value remained 6,2%.

In summary, the category *Colocated* '3-4 days a week' is of medium significance, with a negative impact on PS compared to the reference group. However none of the other categories is significant in their influence on PS. Based on the linear regression models, we can falsify H0 and can also falsify hypothesis H1 revised. This means that being colocated does not have a positive influence on PS, but rather a negative influence, and that it alone can describe 6,2% of the deviations in PS.

4.1.2 Enduring teams

This element focuses on the ways in which agile software teams establish the steps towards achieving an enduring team. Based on participant 5's relatively safe work environment, they described their first-hand experience of transitioning from a newly-formed team to an enduring one, and the importance enduring teams had for ensuring PS.

> *"I think it takes a long time to adjust with a team. Everyone thinks in different ways. Everyone has their own opinion. It takes some time to adjust with each other. But after having a good understanding among each other, I think it is a very safe environment. And I feel confident to work in it"*
>
> (Participant 5)

Participant 5 highlights the difference in capabilities that individual members of a newly-formed team bring to the table. They also emphasise the necessity of mutual adaptation and comprehension among team members for the team to evolve into an enduring one, which is a time consuming process. Additionally, participant 5 establishes a link between enduring teams and a safe working environment as a natural outcome of the team's development.
In the following quote, participant 10 outlines their team's perspective on the necessity of having some level of independence, when making decisions about organising collaborative efforts, as well as having time to help each other:

> *"Because if we are at the same level or something, I mean, technically, at the same level with the older developers, that will help a lot on the long run. That's the mentality behind this. And we are always trying to help each other like peering together. Or maybe if sometimes someone is stuck with a technical difficulty, and we will jump in and just help each other"*
>
> (Participant 10)

The quote highlights the significance of the organisational environment that managers create in influencing a sense of cohesion and longevity within the team. According to participant 10, the team they are a part of has been afforded a suitable level of autonomy in organising their collaborative efforts, with all members being at the same technical level. This approach has resulted in a supportive atmosphere and fostered a sense of peering together as a team. This also re-emphasises the importance of enduring teams, in that they facilitate an equal technical level of expertise, which allows for better work performance and safety. Participant 7 elaborates on the importance of promoting mutual assistance among team members.

> "[...] *If you take people, team members out of the project, and move them to another project, and then another team member comes in, and then you need to start again, with the knowledge sessions* [...]. *But if a member remains on the team, it definitely helps even better"*
>
> (Participant 7)

Participant 7 raises a concern regarding the inevitable change of team members, which can potentially lead to time-consuming onboarding and knowledge transfer for both the established team and the new members. Participant 7 explains that knowledge-sharing is essential in changing teams. They propose that some team members should stay on the team and in their current position, which will aid in the transition of team members and the establishment of knowledge sharing. Eliminating possible time consumption related to changing teams.

To summarise participant 5 highlights how a team's sense of enduringness can create a more supportive and safe work environment. Additionally, participant 10 identifies that giving teams the freedom to collaborate in their own way can enhance mutual assistance and foster a sense of peering together. Lastly, participant 7 emphasises the importance of facilitating knowledge transfer during changes in team membership. These findings provide an impetus to investigate whether enduring teams are associated with higher levels of PS in agile software teams. Therefore, the following hypothesis is proposed:

> **H2**: Being an enduring team is positively associated with greater psychological safety in an agile software team

The five linear regression models conducted can be seen in Table 9. In Model 1, *Enduring teams* do not have a significant connection with PS. The category 'More than 4 years' is of medium significance and has a positive impact on PS with a rise of 4,53 points, in comparison

to the reference group, see (RG) in Table 9. *TeamSize* and *Gender* does not influence the relationship between *Enduring teams* and PS, neither of them are significant in their influence of PS. Model 1 can describe 0,51% of the deviation in PS.

	Model 1	**Model 2**	**Model 3**	**Model 4**	**Model 5**
Enduring teams: Less than 1 year (RG)					
Enduring teams: 1 - 2 years	0,7611 (16097)	0,9221 (1,6035)	0,8517 (1,5833)	1,1055 (1,5921)	1,1442 (1,5950)
Enduring teams: 3 - 4 years	-0,1836 (2,0312)	0,0596 (2,0245)	-1,4353 (2,0311)	-1,2562 (2,0298)	-1,3170 (2,0345)
Enduring teams: More than 4 years	4,5317** (1,9353)	4,5864** (1,9224)	2,3410 (1,9982)	3,3371 (2,0156)	3,3844* (2,0191)
Gender (Man = 0)	-0,1465 (1,5766)	-0,3206 (1,5637)	0,0627 (1,5538)	0,0722 (1,5654)	0,1308 (1,5704)
TeamSize	-0,0545 (0,1379)	-0,0120 (0,1373)	0,0071 (0,1361)	0,0803 (0,1415)	0,0733 (0,1422)
Sourcing model: In-house (RG)					
Sourcing model: Outsourced		-4,0269*** (1,4191)	-3,2373** (1,4131)	-2,9119** (1,4201)	-2,7971* (1,4370)
Sourcing model: Mixed		-4,5977** (2,2318)	-4,2763* (2,2075)	-4,0238* (2,2069)	-3,9617* (2,2117)
Experience: Less than 3 years (RG)					
Experience: 3 - 5 Years			3,1477* (1,8182)	2,8778 (1,8196)	2,9230 (1,8229)
Experience: 6 - 8 Years			7,2577*** (2,0592)	6,4279*** (2,0714)	6,4376*** (2,0731)
Experience: 9 - 11 Years			6,2118** (2,5361)	4,9679* (2,5796)	4,9450* (2,5819)
Experience: More than 12 years			7,3737*** (2,0298)	5,6661*** (2,1443)	5,6967*** (2,1468)
Role: Software developer (RG)					
Role: Software engineer				2,2748 (2,1100)	2,2283 (2,1134)
Role: Senior software engineer				5,6549** (2,4138)	5,5607** (2,4220)
Role: Tech Lead				2,7644 (2,6306)	2,6299 (2,6444)
Role: QA				-0,2176 (2,,3361)	-0,3016 (2,3431)
Role: Other				-0,0203 (2,7145)	-0,1076 (2,7214)
Multiple teams					0,7114 (1,3174)
Const.	77,9582	79,0163	74,6282	72,4041	71,9763
Adj. R^2	0,0051 (0,51%)	0,0229 (2,29%)	0,0526 (5,26%)	0,0954 (9,54%)	0,096 (9,60%)
N	468	468	468	468	468

Table 9 – An overview of the five models tested in regard to *Enduring teams*. * indicates low-significant (P-value < 0,10), ** indicate medium significant (P-value < 0,05), and *** indicate high-significant (P-value < 0,01).

Model 2 includes the control variable *Sourcing model*, when the variable *Sourcing model* is kept stable the influence of *Enduring teams* on PS rises positively for every category. As an example, '1-2 years' goes from influencing PS with 0,761 in 'Model 1 to influencing with 0,922 in Model 2. The category 'More than 4 years' is still significant to the level of 0,05. *Sourcing model* is of both medium and high significance in its relation to PS, and compared to 'In-house' software development, both 'Outsourced' and 'Mixed' have a negative influence on PS. In Model 2 the adj. R^2 is 0,0229 meaning 2,29% of the deviation is described.

When the control variable *Experience* is added, as seen in Model 3, *Enduring teams* ceases to be significant, and its influence is minimised. *Experience* is highly significant for every category, except for the category '3-5 years' which is only of low significance. In general it is seen that when experience rises, PS is also higher, the influence stabilises after 8 years of experience. It is even seen that *Experience* in the category '9-11 years' influences PS less positively with 6,21 points compared to the category '6-8 years' that influences with 7,26 points.

Model 4 includes *Role* as a control variable, with 'Software developer' as reference group. The adj. R^2 for Model 4 is 9,54%, compared to Model 3 that had an adj. R^2 of 5,26%. When *Role* is controlled for, *Enduring teams* have a minor rise in its positive influence on PS. *Role* in itself is not significant with PS, except for the category 'Senior software engineer' which when going from being a 'Software developer', to being a 'Software engineer', the PS rises with 5,65 points. This category is of medium significance. The control variable *Multiple teams* is added in Model 5. The control variable is not significant with PS and only raises the adj. R^2 with less than 0,1% point. However, *Multiple teams*, when kept stable, support the *Enduring teams* category 'More than 4 years' to become significant again, although it is only at a 0,1 level.

One interesting aspect of Model 5 is the impact of *Enduring teams* on PS. While an Enduring team of '1-2 years' and '3-4 years' were found to be nonsignificant, an enduring team of 'More than four years' were found to be significant, albeit at a low level of 0,1. The amount of impact that the level of enduringness has on PS is also interesting; a value of 3,38 is a large step on the PS scale. To better capture the concept of *Enduring teams* in a future study, it would be relevant to scale the variable as an interval-scaled variable, as opposed to a categorical variable. This might be able to minimise the difference between the categories and measure a team's enduringness more accurately with a higher granularity.

Based on the linear regression models, we partially falsify H0 and can therefore accept that our hypothesis H2 is valid until proven false. This means that *Enduring teams* partially has a positive influence on PS, as long as the team has been together for more than four years, and that it alone can describe 9,6% of the deviations in PS.

4.1.3 Autonomy

Autonomy was another element that was found interesting in the qualitative analysis. Many of our interviewees show signs of being self-managed to different degrees (Table 2 in section *3.3.1 Analysis of the qualitative data*). This gives them a sense of self-control and empowerment, which is one of the 12 principles in the agile framework (Beck et al., 2023). Participant 7 says the following when the researcher assess that they work in a very safe work environment:

> "*Yes, that is a fair assessment. Because each team member, each of the team members can participate with ideas and improvements. Mostly, if there are well, larger improvements that need to be made, then we create items, stories, improvement tasks, for that. Otherwise, we can just sign off on those changes and allow the colleague to perform those changes.* [...] *That does not mean that we don't see the consequences, it just means that we are allowed to step out of our comfort zone"*
>
> (Participant 7)

Participant 7 explains how autonomy is a key part in everyone taking decisions and acknowledging the work they do together. They indirectly associate being able to take the decisions themselves, and not having to consult any manager or other personal beforehand with a psychological safe work environment. It gives them freedom, and from the interview it is possible to extract that it also results in a feeling of confidence and pride for their work. They are empowered by being autonomous because they can make decisions themselves, they are trusted by the company.

There were also interviewees that showed signs of a higher degree of being managed. Participant 5 explains a situation where a team lead turned down a promising idea to keep their influence. They explain that they killed the idea with the reason being that they were not very confident in that language. Participant 5 explains that the consequence was that "*the performance dropped significantly. And the development costs being much higher than anticipated*" (Participant 5). In addition to this, participant 5 explains that to avoid these consequences, it is important that the people who make the decisions are "*knowledgeable, and*

they should be willing to accept changes, ideas, and initiatives. Because if you don't accept the changes, you go far behind" (Participant 5).

The level of autonomy in a team has shown to influence the decision process, and the feeling of value and empowerment when working. A feeling of value is relevant when measuring PS, while empowerment also fosters a feeling of value. We therefore propose the following hypothesis:

> **H3:** Autonomy of a team is positively associated with psychological safety in an agile software team

When testing *Autonomy's* influence on PS the control variables were added one by one as shown in Table 9. In Model 1 when a person is raised 1 unit on the *Autonomy* scale the PS is raised with 0,347 points. This connection is highly significant to the level of 0,01 in its p-value. *Gender* and *TeamSize* have no influence on PS and are not significant.

Model 2 is including the variable *Sourcing model* which does not influence the relation between *Autonomy* and PS. *Sourcing model* itself is of medium significance to PS where 'Mixed' is more negatively influenced than 'Outsourced'. Model 2 describes 16,5% of the deviation in PS.

Model 3 includes the variable *Experience*, which lowers *Autonomy's* influence on PS, from 0,341 to 0,328. The relation is still highly significant. *Experience* influence on PS stabilises around 8 years of experience, as is also seen with H2. In Model 4 and 5 the variables *Role* and *Multiple teams* are added which does not influence *Autonomy*'s relationship with PS. Model 4 has an adj. R^2 value of 19,5% while Model 5 is of 19,4%.

Based on the linear regression models, Table 10, we falsify H0 and can therefore conclude that our H3 is valid until proven false. In summary, *Autonomy* has a positive relationship with PS, where a higher degree of autonomy leads to a higher degree of PS, and the model can explain 19,5 percent of the variance within PS.

	Model 1	Model 2	Model 3	Model 4	Model 5
Autonomy	0,3470*** (0,0374)	0,3419*** (0,0372)	0,3280*** (0,0368)	0,3263*** (0,0369)	0,3276*** (0,0370)
Gender (Man = 0)	-0,1863 (1,4537)	-0,3375 (1,4431)	0,3280 (1,4354)	-0,0675 (1,4488)	0,0091 (1,4533)
TeamSize	0,0541 (0,1257)	0,0924 (0,1253)	0,0690 (0,1245)	0,1165 (0,1297)	0,1071 (0,1304)
Sourcing model: In-house (RG)					
Sourcing model: Outsourced		-3,3699** (1,3129)	-2,6948** (1,3064	-2,3270* (1,3176)	-2,1830 (1,3329)
Sourcing model: Mixed		-4,5170** (2,0505)	-4,3494** (2,0288)	-4,0339** (2,0328)	-3,9581* (2,0365)
Experience: Less than 3 years (RG)					
Experience: 3 - 5 Years			2,2505 (1,6813)	2,0802 (1,6854)	2,1288 (1,6875)
Experience: 6 - 8 Years			6,1765*** (1,8850)	5,6229*** (1,8967)	5,6264*** (1,8977)
Experience: 9 - 11 Years			6,0984*** (2,2738)	5,3368** (2,3124)	5,3076** (2,3139)
Experience: More than 12 years			5,7916*** (1,7985)	4,4786** (1,8913)	4,5085** (1,8927)
Role: Software developer (RG)					
Role: Software engineer				2,2605 (1,9452)	2,2081 (1,9475)
Role: Senior software engineer				5,3752** (2,2162)	5,2520** (2,2237)
Role: Tech Lead				2,8008 (2,4129)	2,6312 (2,4252)
Role: QA				0,5234 (2,4129)	0,4296 (2,1628)
Role: Other				1,2968 (2,4946)	1,2022 (2,4992)
Multiple teams					0,8897 (1,2141)
Const.	52,4376	53,8402	51,0347	49,0158	48,4150
Adj. R^2	0,151 (15,1%)	0,165 (16,5%)	0,189 (18,9%)	0,195 (19,5%)	0,194 (19,4%)
N	468	468	468	468	468

Table 9 – An overview of the five models tested in regard to Autonomy. * indicates low-significant (P-value < 0,10), ** indicate medium significant (P-value < 0,05), and *** indicate high-significant (P-value < 0,01).

4.2 Team environment

This section will analyse the Team environment theme, which includes the elements of Agreeing upon team values and Openness in communication.

4.2.1 Agreeing upon team values

The Agreeing upon team values element encompasses the impact of different values on team collaboration. Within this element, four primary values have been identified: no blame, recognition for work, trust, and unity. The importance of implementing a no blame culture within a team has been highlighted by several participants, namely participants 1, 9, 15, 17, and 18. For instance, participant 17 explains how their team views the inevitability of mistakes and the value of not assigning blame.

> *"*[...] *you won't own any mistakes or in a single person, I guess it's because we don't do that.* [...] *we don't blame we own the issue together as a team. How we*

> *handle this by having lessons learned retrospective meetings* [...]. *These are dedicated to what we have learned and how we can avoid it in the future"*
>
> (Participant 17)

Participant 17 emphasises the notion that mistakes are not the sole responsibility of an individual team member, but rather a shared responsibility within the team. This team value can be regarded as a means to enhance PS by reducing the personal risk-taking of each team member while performing a task (Edmondson, 1999). Participant 17 further describes the utilisation of the Scrum ritual of retrospectives to analyse mistakes and extract valuable insights from them.

Additionally, participant 17 advocates for the team value, recognition of work, which is embraced in participant 17's team. Participant 17 highlights, "*we are being recognized unconditionally.* [...] *everyone will be recognized in my game based on their efforts"* (Participant 17). The use of the term *unconditionally* implies a high level of trust, which is also a team value discussed by participant 18 in response to the researcher's inquiry.

> Researcher: *"Yes, so when you feel like you are not judged, how does it influence your attitude at work and the way you do things?"*
>
> Participant 18: *"I think you open up very well to the team because you start to trust them. Otherwise you work alone. You don't ask for help and you do more mistakes. You don't admit your problems, your mistakes and it is not good for the whole team, not just myself"*
>
> (Participant 18)

The above discourse presents participant 18's team's interpretation of trust and the repercussions of not practising it. One such consequence is being perceived as an individual working in isolation by the team, which results in reluctance to seek help and, consequently, more errors.

Furthermore, participant 10 emphasises the significance of having team unity, which aids team members in progressing and supporting each other. The aforementioned examples of team values can be distilled to how a team selects a value and translates it into a functional tool for their work practices. As illustrated by the examples, there is a correlation between values and an individual's perception of PS in the work environment. Values such as no blame, recognition

for work, unity, and trust can influence an individual's willingness to take personal risks while working in a team. Thus, the following hypothesis is proposed:

> **H4:** The existence of shared team values increases psychological safety in an agile software team

While this theme was not in the scope of the quantitative data collection, this hypothesis and possible future work will be discussed in section *5.6 Future work*.

4.2.2 Openness in communication

The focus of this theme revolves around the extent to which sincere and open communication about difficult topics or issues can impact a team's collaboration. Participant 6 elaborates upon how a lack of communication regarding the competencies of individual team members can result in negative assumptions about their abilities:

> *"In my previous project, I was working with someone, to be honest, I thought he has not much talent. But when I had calls with him, when we went through the coding, I really knew he is really more talented than me"*
>
> (Participant 6)

The above quotation highlights the cognitive biases harboured by participant 6 concerning a colleague's attitude. However, upon engaging in collaborative activities and reviewing the colleague's code, participant 6's perception underwent a transformation, leading to a newfound appreciation of the colleague's talents. This highlights that individuals' proficiencies and expertise may not always be overtly discernible or readily accessible. Participant 6 relates their own personal experiences of being less communicative than their colleagues, and how PS affects this:

> *"*[...] *I'm not much talkative. But I think it took me 27 years to understand. I need to improve my communication, even in my personal life. I think I have improved my communication issue. But, yes when you have safety it helps to open up because I always feel nobody would judge me"*
>
> (Participant 6)

As is evident from the aforementioned quotation, participant 6 acknowledges experiencing some communication challenges, which prompted a need for improvement. Furthermore, participant 6 asserts that creating a sense of safety is conducive to effective communication and

mitigates the fear of being evaluated negatively by others. This assertion aligns with Edmondson's research on the salutary impact of PS in team environments (Edmondson, 1999).

When asked to provide an illustration of a scenario in which a team member made a mistake and how the team responded, participant 13 encountered difficulty in recalling a specific example. Instead, participant 13 offered a rationale for this memory lapse.

> *"*[...] *we talk as a team, we talk to each other very often. And if there is a problem, then we catch it before it gets bigger* [...]. *Therefore, maybe that's one of the reasons we don't face such difficulties as often. It doesn't indicate that it won't happen, but so far, either I don't remember or maybe it hasn't happened"*
>
> (Participant 13)

Participant 13 suggests that frequent communication within their team precludes the proliferation of issues, which may escalate into more significant problems. An environment that fosters open communication, free from judgement or blame, is a hallmark of PS (Edmondson, 1999). This statement suggests that participant 13 perceives their work environment as psychologically safe.

Although open and sincere communication cannot be assumed in all team collaborations, as illustrated by participant 6's account of an unacknowledged talented colleague, participant 6's personal experience emphasises the value of a non-judgmental and safe environment in promoting open communication. Similarly, participant 13 highlights the importance of communication to quickly detect mistakes and problems. This outcome may be attributed to the presence of PS in their work environment. Based on the positive impact that open communication had upon the work environment of both participant 6 and participant 13, we propose the following hypothesis:

> **H5:** Openness in communication is positively associated with psychological safety in an agile software team

Five linear regressions were conducted where the control variables were added incrementally, see Table 10. As seen in Model 1, when *Openness* in communication moves one unit, it enhances PS with 0,538 points. The control variables *TeamSize* and *Gender* both have a negative coefficient for PS, meaning that when *TeamSize* is raised with one unit, it lowers the PS with 0,093 points. *Gender* lowers PS with 0,723 points if *Gender* changes from a man to a woman. However, neither *TeamSize* nor *Gender* is significant.

	Model 1	Model 2	Model 3	Model 4	Model 5
Openness	0,5388*** (0,041)	0,5323*** (0,041)	0,5281*** (0,042)	0,5322*** (0,040)	0,5322*** (0,040)
Gender (Man = 0)	-0,7230 (1,295)	-0,8318 (1,275)	-0,5352 (1,276)	-0,4760(1,1232)	-0,4404 (1,237)
TeamSize	-0,0929 (0,124)	-0,0596 (0,121)	-0,0693 (0,123)	-0,0148(0,121)	-0,0192 (0,121)
Sourcing model: In-house (RG)					
Sourcing model: Outsourced		-2,5684** (1,145)	-1,8480* (1,111)	-1,5016 (1,099)	-1,4364* (1,110)
Sourcing model: Mixed		-4,0669** (1,835)	-3,9900** (1,865)	-3,5651* (1,850)	-3,5301* (1,855)
Experience: Less than 3 years (RG)					
Experience: 3 - 5 Years			3,9924*** (1,470)	3,8047*** (1,452)	3,8284*** (1,452)
Experience: 6 - 8 Years			6,0937*** (1,571)	5,5769*** (1,599)	5,5802*** (1,600)
Experience: 9 - 11 Years			7,5258*** (2,124)	6,9287*** (2,104)	6,9159*** (2,105)
Experience: More than 12 years			7,2829*** (1,557)	6,1311*** (1,656)	6,1483*** (1,655)
Role: Software developer (RG)					
Role: Software engineer				3,2073* (1,786)	3,1831* (1,770)
Role: Senior software engineer				5,4462*** (1,952)	5,3891*** (1,931)
Role: Tech Lead				2,2870 (2,111)	2,2083 (2,085)
Role: QA				-0,2056 (1,917)	-0,2497 (1,902)
Role: Other				1,8754 (2,258)	1,8295 (2,243)
Multiple teams					0,4094 (0,978)
Const.	37,7676	39,0742	34,4377	31,6667	31,4316
Adj. R^2	0,340 (34,0%)	0,349 (34,9%)	0,380 (38,0%)	0,393 (39,3%)	0,392 (39,2%)
N	468	468	468	468	468

Table 10 – An overview of the five models tested in regard to Openness. * indicates low-significant (P-value < 0,10), ** indicate medium significant (P-value < 0,05), and *** indicate high-significant (P-value < 0,01). The robust standard errors are given in parentheses.

In Model 3, the control variable *Experience* was added, where the reference group is 'Less than 3 years'. When *Experience* is kept stable, it shows a lowering of *Openness* influence on PS, and it goes down to 0,532 points. *Experience* impacts PS positively, meaning the years of experience is impacting the amount of PS increasingly. However the largest influence is seen in the beginning of the ladder where '3-5 years' of *Experience* raises PS with 3,99 points, when *Experience* is '6-8 years' PS is raised with 6,09 points, and at '9-11 years' PS is raised with 7,53 points. At 'More than 12 years' PS is raised with 7,28 points. All these numbers should be seen in relation to the reference group 'Less than 3 years', which means that the higher the experience the minor impact it has on PS. Therefore, experience is more important for PS in the first couple of years compared to further in one's career. The connection is highly significant in each group. Model 3 explains 38% of the deviations in PS.

Model 4 has an adjusted R^2 of 39,3% and includes the variable *Role. Role* does not influence *Openness's* connection with PS. *Role* does however influence PS where 'Software developer'

is the reference group. When a person goes from being a 'Software developer' to being a 'Software engineer' the PS is raised with 3,207 points. In addition, this connection is significant at a level of 0,1. When it is 'Senior software engineer' it is raised with 5,446 points, this connection is highly significant. In Model 5, when the control variable *Multiple teams* are added, the category 'Outsource' becomes significant again compared to Model 4.

To summarise, openness in communication has a positive influence on PS, with a high level of significance, when moved one unit on the *Openness* scale, PS is raised with 0,53 points. *Openness* also describes around 34,0% of the deviations in PS without any significant control variables. Based on the linear regression models we are therefore able to falsify H0 and accept that H5 is valid until proven otherwise.

4.3 Management and stakeholder relationship

The relationship between the software team and their management and stakeholders can often be of vital importance, and at the centre of this is communication. Holten et al. (2015) stated that communication was essential for an agile software environment to succeed. As such, this section will be split into two parts: Management, and Stakeholder communication.

4.3.1 Management

Participant 12 has had many negative experiences with management, one of them being in relation to making a mistake, by misplacing production data while their manager was on vacation:

> "*I got calls, personal calls from my manager, and he was asking, like, what would you do? I was just on, I was on a vacation, how could you do this to me and my team, and you have like, completely ruined my day"*
>
> (Participant 12)

This quote suggests indications of a psychologically unsafe environment. Personal blame is placed on the developer, and instead of constructive criticism, where the issue is hopefully resolved, the blame is placed on a single individual. Afterwards, management was unable to be contacted for the next eleven days, and when they returned, they stated that it was not that big of an issue. This made participant 12 lose job satisfaction and stop wanting to give anything extra in their job, evoking feelings of sadness since participant 12 beforehand enjoyed working. These feelings are expressed in his explanation for how it made him feel:

> "*It definitely affected me in a way that post that it's been almost one and a half year now, I have made sure I don't work more than eight hours in it. What I have learned is that we are just workers for the game and being at the lowest level as a developer*"
>
> (Participant 12)

Here, it is quite apparent that unsupportive management has affected them negatively. Not only do they no longer wish to work more than necessary, it has also affected them negatively on a personal level, resulting in losing job satisfaction. It is as such clear that unsupportive management had a negative effect on participant 12's PS.

Participant 5 discusses the impact that a change in management can have. "*They* [Management, red.] *try to listen to us so, and now we have an open and transparent communication. When I joined the company, I faced a lot of struggles*" (Participant 5) It is clearly shown that management, which is willing to communicate, is a positive force and it also indicates a more psychologically safe work environment. Participant 6 discusses the impact that a manager can have on a team:

> "[…] *he* [Manager, red.] *should earn the trust from us. If he is capable to earn it, everyone, everyone will do something more than he expect. So actually, the project was one year, but deliver it within eight months, we completed the entire project*"
>
> (Participant 6)

The above quote indicates that supportive leadership and management results in a better product, and performance among team members. Supportive management empowers team members to work beyond the expected performance. One cannot simply order their team to perform better; trust is instead earned by being a supportive manager, and what the manager does reflects onto the rest of the team. As participant 13 states: *"*[…] *I think these kinds of approaches start from top to bottom. So higher hierarchy people, if they're welcoming then the lower degrees are also affected by it*" (Participant 13). As such, management is an integral part of the company; how they perform affects how the company performs. The analysis indicated that unsupportive management can have negative effects on their employees performance, while supportive management have a positive effect on the employees performance. We propose the following hypothesis to investigate further:

H6: Supportive management is positively associated with psychological safety in an agile software team

Five different linear regressions were conducted, each one with progressively more included control variables. Model 1 in the Table 11 shows, when *Management* moves one unit, it impacts the PS scale by 0,345 points, indicating that the more supportive the management is the higher the PS. Model 1 can explain 22% of the deviations in PS.

	Model 1	**Model 2**	**Model 3**	**Model 4**	**Model 5**
Management	0,3454*** (0,045)	0,3394*** (0,045)	0,3433*** (0,045)	0,3401*** (0,044)	0,3410*** (0,044)
Gender (Man = 0)	-0,5791 (1,486)	-0,6801 (1,449)	-0,3820 (1,478)	-0,4870 (1,463)	-0,4090 (1,478)
TeamSize	-0,0899 (0,139)	-0,0565 (0,136)	-0,0833 (0,136)	-0,0596 (0,133)	-0,0698 (0,134)
Sourcing model: In-house (RG)					
Sourcing model: Outsourced		-2,4050** (1,224)	-1,14583 (1,160)	-1,1469 (1,172)	-0,9963 (1,181)
Sourcing model: Mixed		-4,3100** (1,976)	-4,0319** (1,988)	-3,6843* (1,967)	-3,6053 (1,973)
Experience: Less than 3 years (RG)					
Experience: 3 - 5 Years			2,7769* (1,499)	2,6880* (1,492)	2,7406* (1,492)
Experience: 6 - 8 Years			6,4345*** (1,716)	6,0978*** (1,757)	6,1039*** (1,757)
Experience: 9 - 11 Years			7,9557*** (2,288)	7,5067*** (2,276)	7,4828*** (2,265)
Experience: More than 12 years			7,8416*** (1,710)	6,9305*** (1,825)	6,9702*** (1,827)
Role: Software developer (RG)					
Role: Software engineer				2,8126 (1,825)	2,7601 (1,881)
Role: Senior software engineer				4,7864** (2,107)	4,6578** (2,089)
Role: Tech Lead				2,6269 (2,335)	2,4516 (2,313)
Role: QA				1,3208 (2,077)	1,2257 (2,066)
Role: Other				2,2092 (2,395)	2,1125 (2,389)
Multiple teams					0,9153 (1,068)
Const.	54,3578	55,5653	50,5273	48,3488	47,7596
Adj. R^2	0,220 (22,0%)	0,229 (22,9%)	0,271 (27,1%)	0,273 (27,3%)	0,272 (27,2%)
N	468	468	468	468	468

Table 11 – An overview of the five models tested in regard to *Management*. * indicates low-significant (P-value < 0,10), ** indicate medium significant (P-value < 0,05), and *** indicate high-significant (P-value < 0,01). The robust standard errors are given in parentheses.

In Model 2, *Sourcing model* is of medium significance in both categories, however in Model 3 when *Experience* is added only the category 'Mixed' for the variable *Sourcing model* is significant. In Model 3, if a person goes from working with 'In-house' software to working with 'Mixed' software development, it lowers the PS with 4,03 points. *Experience* is again highly significant in every category except '3 - 5 years' where it is only significant at a low level of 0,1. It is possible to see that when *Management* holds stable for the control variable

Experience, higher experience results in a higher PS, but again stabilises around 11 years of experience. Model 3 can explain 27,1% of the deviation in PS, according to the R^2 value.

Model 4 and 5 includes the control variables *Role* and *Multiple teams*. Only the category 'Senior software engineer' is significant and has a positive influence on PS compared to the reference group 'Software developer'. Based on the linear regression models, we falsify H0 and can therefore accept that our hypothesis H6 is valid until proven false. This means that supportive management has a positive influence on PS, and that it along with the control variables can describe 27,3% of the deviations in PS.

4.3.2 Stakeholder communication

Stakeholder communication is an essential part of agile software development (Beck et al., 2023). Participant 12 experienced a lack of direct communication with the stakeholder: "[…] *client calls were always taken by the tech lead and the manager. So when you are not there in the meeting, you can never give any idea* [...]" (Participant 12). Participant 12 previously mentioned a work environment lacking in PS and gives an example where the lead and manager do not entrust the team members with direct client communication. This is not in accordance with the agile principle of businesspeople and developers working together to develop working software. It also becomes apparent that a lack of direct communication with the stakeholder between all members of the team resulted in a rigid and less creative working environment. This work approach has changed in recent years. Participant 12 describes it as "[…] *the biggest improvement that over the years has happened*" (Participant 12), highlighting the importance open stakeholder communication has for a team.

The product itself may also benefit from direct stakeholder communication. Participant 6 and their team were able to develop the first application which a penetration tester was unable to break into, and additionally was able to develop a better product:

> "*And we had close relationship with the product owner. And he said, Okay, guys, if you think this can be done, it's up to you just do it and show me. So that's how we proposed and developed a better solution, not only the design but the usability and better features*"
>
> (Participant 6)

Here it is clearly demonstrated that close working relationships feeds trust and allows for a more autonomous and ambitious project to be worked on and developed. As such, management can support a psychologically safe environment for the team and help resolve issues in the open stakeholder communication. Participant 6 states the positive outcomes of stakeholder communication as follows:

> ”[…] *when we get the trust of clients, he know, yeah, there can be best, but these guys are capable* […] *But our confidence to deal with them is better and we don't feel guilty and we fix them fast and make the code even better when we fix”*
>
> (Participant 6).

Here, direct communication with the stakeholder is presented as unequivocally a positive thing, which implies that the stakeholder is also responsible for the connotation of the communication in their relation. This implies that direct communication with the stakeholder can have a positive effect on PS. Based on the analysis conducted so far, we propose the following hypothesis:

> **H7:** Open communication with stakeholders is positively associated with psychological safety in an agile software team

While this theme was not in the scope of the quantitative data collection, this hypothesis and possible future work will be discussed in section *5.6 Future work*.

4.4 Clear Feedback

The theme Clear feedback encapsulates the two elements: Clear decision process and Clear feedback structures. Clear decision process measures the team members' ability to administer their decision process within the team, as well as their level of collective decision making. Clear feedback structures is a theme capturing how the team members address feedback and how they practise it.

To clarify the use of the word *clear* in the themes, we will use our framework adapted by Weick's (1995) concept of sensemaking. When a decision process and feedback structure is clear it should be understood in the social context where the individuals interact. It builds upon clues, which is formative depending on the context and the individual's viewpoint. The clues

provide knowledge about the current context and how to enact based on set knowledge. Clear is therefore a reflective ongoing process, based in the social context, creating a fluid definition.

4.4.1 Clear decision process

When analysing the interviews, many of the interviewees mentioned the importance of having a clear decision process. There were different elements of a decision process that the interviewees mentioned as important. First, they explained that it was important to have shared responsibility in decision processes. Participant 10 explained a situation where they did not have the best arguments when discussing what tactic to use, and they therefore went with a different approach that failed. Even though it was not participant 10's idea that failed, they still saw it as the team failing together. Participant 10 could not convince the team in the decision process because of participant 10's lack of sufficient argumentation. The fact that the team made this decision cohesively made them stronger as a team, and together they found a new solution.

This leads to another important element relevant for a clear decision process: decisions need to be considered as a team, and the decisions should be done collaboratively. Participant 15 explains "[…] *in Agile there is no individual there is only a team. So, if you are making a decision, then you need to consider it as a team. You don't make decisions alone* […]" (Participant 15). Participant 10 elaborates a bit further in his explanation:

> "[…] *we all should agree on that or on something in the team, we have this rule, we shouldn't have different approaches for different things* […]. *Either I should agree with your idea, or you should agree with my idea. We must be consistent"*
>
> (Participant 10)

Participant 10 directs attention to another important aspect of establishing a clear decision process; to decide on common decision structures. Participant 10 mentions that "*in the end, we always have a rule of meeting that we need to have consensus, or we need to have an output of who and how to continue or how to proceed*" (Participant 10,). The quotes state a relevant aspect found by participant 10, when discussing a decision; they need to have consensus in their decisions or otherwise they need to have a clear process regarding how to proceed without consensus. Participant 15 also adds to this approach by saying that "*if you want to change*

something, you have to explain why, and you have to have an approval to go ahead" (Participant 15). This underlines the statements that you need to have a good argument when convincing your team members, and that in a decision process the best argument gets chosen. To bring a sufficient argumentation participant 13 explains how he studied a solution beforehand to be sure he could answer the questions from his team members:

> *"Of course, I made my study before bringing it* [a solution, red.] *up. So I was prepared to answer the questions that may ask. And of course, they asked a lot of questions like, did you think about this? The did you consider that this would happen if you say that, and, and then I answered all of their questions. And then they just accepted it. And then they, they also gave me a chance to do it in one or two sprints. And it went well"*
>
> (Participant 13)

It is clear from the interviewees that it is important to argue for an idea together, and that well articulated arguments stand better in the decision. Decisions should be decided upon collectively, making sure everyone adheres to the chosen decision. If a disagreement would arise the team should solve it together based on their decision process. Based on these findings in the qualitative analysis, we propose the following hypothesis to test in the quantitative data.

> **H8:** Clear decision processes results in higher psychological safety in an agile software team

Table 12 shows a summary of the linear regression conducted to test H8. Clear decision impacts the scale of PS positively. Moving one unit on the *Clear decision* scale impacts PS with 0,548 points in Model 1. The scale continues to be of high significance throughout the models. In Model 2 the control variable *Sourcing model* is added, which is of medium significance in relation to its influence on PS, where it has a negative impact to work with 'Outsourced' or 'Mixed' software, compared to 'In-house' software. However, these numbers are only medium significant in Model 2, when moving to Model 3 the categories are low significant.

	Model 1	**Model 2**	**Model 3**	**Model 4**	**Model 5**
Clear Decision	0,5487*** (0,045	0,5395*** (0,045)	0,5230*** (0,046)	0,5238*** (0,045)	0,5250*** (0,045)
Gender (Man = 0)	-0,2186 (1,364)	-0,3281 (1,348)	-0,0532 (1,372)	-0,1407 (1,319)	-0,0593 (1,331)
TeamSize	0,0521 (0,125)	0,0802 (0,124)	0,0629 (0,127)	0,1125 (0,125)	0,1024 (0,126)
Sourcing model: In-house (RG)					
Sourcing model: Outsourced		-2,4603** (1,167)	-1,8809* (1,133)	-1,5007 (1,131)	-1,3472 (1,164)
Sourcing model: Mixed		-3,4464** (1,727)	-3,3439* (1,748)	-2,9475* (1,738)	-2,8645* (1,734)
Experience: Less than 3 years (RG)					
Experience: 3 - 5 Years			2,3537 (1,516)	2,1451 (1,516)	2,1981 (1,513)
Experience: 6 - 8 Years			5,3379*** (1,692)	4,7527*** (1,708)	4,7561*** (1,708)
Experience: 9 - 11 Years			5,6353*** (2,102)	4,8374** (2,066)	4,8059** (2,058)
Experience: More than 12 years			5,4702*** (1,694)	4,1483** (1,827)	4,18266** (1,827)
Role: Software developer (RG)					
Role: Software engineer				2,8019 (1,756)	2,7475 (1,737)
Role: Senior software engineer				5,8770*** (1,978)	2,7471*** (1,965)
Role: Tech Lead				2,6838 (2,152)	2,5032 (2,129)
Role: QA				0,8185 (1,911)	0,7188 (1,894)
Role: Other				1,4924 (2,260)	1,3901 (2,245)
Multiple teams					0,9430 (1,062)
Const.	36,0240	37,4903	35,1565	32,,6290	31,9973
Adj. R^2	0,294 (29,4%)	0,301 (30,1%)	0,320 (32%)	0,330 (33,3%)	0,329 (32,9%)
N	468	468	468	468	468

Table 12 – An overview of the five models tested in regard to *Clear Decision*. * indicates low-significant (P-value < 0,10), ** indicate medium significant (P-value < 0,05), and *** indicate high-significant (P-value < 0,01). The robust standard errors are displayed in the parentheses.

In Model 3 the variable *Experience* is controlled for. In this test, *Experience* is highly significant in every category in Model 3, except for the category '3-5 years'. Throughout the different models, the categories '9-11 years', and 'More than 12 years', is of high significance. The experience is also most important in the early stages of a person's work career, as it stabilises at around 8 years of experience. Model 3 can explain 32% of the deviations in PS. Model 4 and 5 includes the variables *Role* and *Multiple teams*, where, again, only 'Senior software engineer' is significant, with a positive impact on PS with 5,8770 points compared to 'Software developer'. *Multiple teams* are insignificant in this context. Based on the linear regression models, we falsify H0 and can therefore accept that our hypothesis H8 is valid until proven false.

4.4.2 Clear feedback structures

One of the seven elements for measuring PS is whether the interviewees finds that mistakes are often held against them. During our analysis we found that the interviewees have processes for giving feedback when a mistake has occurred. These feedback structures help to learn and improve the team's understanding of why those mistakes happened and helps them avoid the mistakes in the future. Participant 7 explains how their feedback structures help to improve the quality of their work, but also how this learning situation is positive in facilitating the PS in their team:

> "[…] *each mistake we run into if we have the proper guidance, and the mentor that can explain to us and well, in kind and calm and constructive manner, it will help us improve in the future. And at least think before we run into that mistake again"*
>
> (Participant 7)

Participant 7 explains the positive outcome feedback structures have on their levels of PS. They do not blame anyone for their mistakes, they try to guide them in how they should have done, and what went wrong, so the team hopefully will not make the same mistake again, and will improve their work. As participant 7 also points out "[…] [we, red.] *need to understand that mistakes are part of the learning Path. Yes. If we don't make any mistakes, then we would not learn anything*" (Participant 7). This is also supported by participant 5 who adds it is important to discuss the mistakes that are made, because then the rest of the team have a chance to learn from you. He also says that "*sometimes we'd like to discuss* [the mistakes, red.] *with our management so that they also understand the consequences. This openness has changed our thinking dramatically. We learn better and faster and it show in the quality of our work*" (Participant 5). The clear feedback structures not only seem to improve PS, but might also be relevant for the quality of their work. This was also suggested and tested by Alami et al. (2023b), where they found that high PS, in agile teams, is positively associated with learning from mistakes which improves software quality.

The clear feedback structures also indicate that it should be easier to admit mistakes and ask for help. Participant 8 were assisted in improving own work that contained a mistake:

> *"Yeah, I felt great. But I also felt that why didn't I ask this before? Like, I know, right. I mean, what I observed was like, until unless you do any mistake, no one will teach you.* […] *sometimes the easiest task is the one where we make mistakes"* (Participant 8).

Participant 8 had a positive experience with asking for help and being told they did something wrong, and that made them wonder why they had not done that earlier. The well formulated feedback made participant 8 more likely to admit mistakes or ask for help again. Participant 9 had a similar experience where genuine feedback really helped them improve their work: "*I was open to feedback. And he also gave me genuine feedback on what should have been done.* […] *if we do not get feedback from the scenarios we never improve"* (Participant 9). Participant 9 gives some good pointers to how feedback should be structured with a clear focus on what should have been done instead. Clear and well formulated feedback is necessary to improve at work. Based on our analysis we propose following hypothesis:

> **H9:** Clear feedback is positively associated with psychological safety in an agile software team

In the test of clear feedback, the linear regression models can be seen in Table 13. Model 1 and 2 includes the control variables *Gender*, *TeamSize*, and *Sourcing model*. In Model 2, the *Clear feedback* scale has a positive impact on PS with 0,42 points every time it is raised one unit on the feedback scale. It has a highly significant influence on PS. In Model 2 the adj. R^2 value is 28,38%.

Gender and *TeamSize* are never significant on any level, indicating that they have no effect on PS. This expands the scope of Verwijs & Russo's (2023) findings, even though gender diversity negatively affects the conflict level within a team, gender does not affect an agile team's PS. Also, like Buvik & Tkalich (2021), *TeamSize* was not found to influence PS. While 'Outsourced' and 'Mixed' have a significant negative effect on the PS in Model 2, this effect is no longer significant when both *Role* and *Experience* are controlled for in Model 4.

Model 3 includes the control variable *Experience*, where all the categories have a highly significant connection to PS throughout the models. It is seen that *Experience* stabilises around

8 years of experience, where it no longer improves the PS, but instead keeps it stable. When adding *Experience,* it is possible to explain 3,75% point more of the deviation in PS.

	Model 1	**Model 2**	**Model 3**	**Model 4**	**Model 5**
Clear Feedback	0,4301*** (0,0318)	0,4223*** (0,0319)	0,4213*** (0,0311)	0,4212*** (0,0313)	0,4219*** (0,0314)
Gender (Man = 0)	-0,6316 (1,3413)	-0,7323 (1,3366)	-0,2826 (1,3134)	-0,3001 (1,3267)	-0,3419 (1,3315)
TeamSize	0,0280 (0,1157)	0,0559 (0,1158)	0,0467 (0,1137)	0,0734 (0,1187)	0,0785 (0,1194)
Sourcing model: In-house (RG)					
Sourcing model: Outsourced		-2,4312** (1,2211)	-1,6366 (1,2007)	-1,3674 (1,2103)	-1,4397 (1,2234)
Sourcing model: Mixed		-3,3708* (1,9007)	-3,2564* (1,8575)	-2,9209 (1,8629)	-2,9592 (1,8667)
Experience: Less than 3 years (RG)					
Experience: 3 - 5 Years			4,0400*** (1,5374)	4,0485*** (1,5436)	4,0235*** (1,5461)
Experience: 6 - 8 Years			7,4696*** (1,7207)	7,2919*** (1,7339)	7,2892*** (1,7355)
Experience: 9 - 11 Years			7,1056*** (2,0800)	7,0282*** (2,1192)	7,0455*** (2,1215)
Experience: More than 12 years			7,6812*** (1,6339)	7,1715*** (1,7208)	7,1531*** (1,7229)
Role: Software developer (RG)					
Role: Software engineer				3,7310** (1,7844)	3,7616** (1,7875)
Role: Senior software engineer				4,1798** (2,0301)	4,2438** (2,0376)
Role: Tech Lead				3,0967 (2,2093)	3,1885 (2,2219)
Role: QA				0,9952 (1,9763)	1,0478 (1,9820)
Role: Other				2,4879 (2,2874)	2,5455 (2,2935)
Multiple teams					-0,4714 (1,1128)
Const.	44,9608	46,3108	41,0571	38,2415	38,4471
Adj. R^2	0,2777 (27,77%)	0,2838 (28,38%)	0,3213 (32,13%)	0,3254 (32,54%)	0,3242 (32,42%)
N	468	468	468	468	468

Table 13 – An overview of the five models tested in regard to *Clear Feedback*. * indicates low-significant (P-value > 0,10), ** indicate medium significant (P-value > 0,05), and *** indicate high-significant (P-value > 0,01).

In Model 5 the control variable *Multiple teams* are added, whether multiple teams exist within the company are also irrelevant to the experience of PS, showing no difference between companies with multiple teams or not. In fact, the model including *Multiple teams* lowers the adj. R^2 value, impacting the explanatory power of the model. Model 3, 4 and 5 both have an R^2 value above 32%, implying that feedback has a significant effect on PS.

Based on the linear regression models, we falsify H0 and can therefore accept that our hypothesis H9 is valid until proven false. This means that clear feedback has a positive influence on PS, and that it alone can describe 32,42% of the deviations in PS.

4.5 Clear expectations

Clear expectations was the last identified theme. This theme is composed of one element, with set title. Clear expectations capture the expectations that a company places on the performance of their employees, as well as the structures established by the company for the employees to fulfil these expectations.

One interesting aspect mentioned by the participants was what happens when mistakes are made accidentally. Participant 1 discusses what happens if errors are made, and in which case people are blamed:

> "[…] *let's say that if you know, you do bad things like that if you're not an integral person, basically. Yeah, that's something that's not like allowed but if you just do stuff because you're naive and stupid, like like, you're just don't educate yourself beforehand yet*"
>
> (Participant 1).

It is generally accepted if an error is made because of a lack of knowledge by the individual developer. Participant 1 mentions later that when the participant found out about his mistake, instead of getting blamed, he informed the other members of the team so that the issue could be resolved. This implies a positive influence on PS.

Participant 1 continued discussing the culture, and the impact it had on the PS of the team: "*I think it's quite clear from the example that we set ourselves how the culture is, meaning, if you've actually done your homework, there is no feeling of unsafety because you already know what you can expect*" (Participant 1). Participant 1 views unsafety as synonymous with uncertainty, and vice versa. As such, it would be relevant to consider a link between clear expectations and PS.

Workplace structures, and what happens if they are unclear, is presented by participant 6. The company participant 6 works for had recently changed management style, to a far more psychologically safe workplace. Participant 6 mentions that the pressure previously was too high, that they were "[…] *the only one who didn't resign, the rest of the team members resigned, and left*" (Participant 6). For participant 6, the workplace has become much better after they started working there. In regard to the team, they mention that "[…] *at the moment we have a*

team, we know what we should achieve. So management has changed, we have a lot of freedom, they listen" (Participant 6).

Participant 6 once again corroborates clear expectations in fostering PS and the structure of the company. Management sets clear structures which bring out the best in the developers. Participant 8 mentions a previous mistake they made while developing software. They had not yet learned the coding standards of the company, and as such participant 8's class names were unreasonably long. However, through the use of an internal plugin and a team lead, these issues were fixed during code review. As participant 8 states: "*But that's where I first learned it, because everyone will think that okay, this is easy*" (Participant 8). Structures can as such have a positive impact on both the PS, as well as the quality, of a given software product.

Following the significance placed upon the clear expectations in relation to facilitating a positive impact upon PS, and the importance of structures as a foundation for this, we propose the following hypothesis:

> **H10:** Clearer expectations in the organisation is associated with improving psychological safety in an agile software team

While this theme was not in the scope of the quantitative data collection, this hypothesis and possible future work will be discussed in section *5.6 Future work*.

4.6 Summary of our findings

Between the ten proposed hypotheses, see Table 14, four of the hypotheses were not eligible for testing with the quantitative data in hand. H1, H4, H7 and H10 are still analysed qualitatively and proposed with a hypothesis, to emphasise a lack of research in those areas, based on the related work found before conducting the analysis. Seven of the hypotheses were possible to test with the data collected by Alami and his colleagues (Alami et al., 2023a). H1 was revised to a testable hypothesis, which was falsified based on having a partially negative relation to PS, only a single category was significant. H2 was also partially accepted, where more granularity in the x-variable might open up for more significant results.

Hypotheses	Accepted?
H1 Revised: Colocated teams are positively associated with psychological safety in an agile software team.	No
H2: Being an enduring team is positively associated with greater psychological safety in an agile software team	Partially
H3: Autonomy of a team is positively associated with psychological safety in an agile software team	Yes
H5: Openness in communication is positively associated with psychological safety in an agile software team	Yes
H6: Supportive management is positively associated with psychological safety in an agile software team	Yes
H8: Clear decision processes results in higher psychological safety in an agile software team	Yes
H9: Clear feedback is positively associated with psychological safety in an agile software team	Yes

Table 14 – An overview of this project's hypotheses. The right column showcases which hypotheses have been falsified, and those who have been accepted until disproved.

Hypothesis H3 was accepted, however a separate analysis conducted with all possible x-variable scales, shows the *Autonomy* scale as irrelevant when the data is controlled for deviations in the scales used for H5, H6, H8 and H9. See Appendix 6. It is important to mention that the x-variables correlated highly, and we therefore have a problem with multicollinearity, meaning the coefficient may be inexact, but can be used as a direction. H3 might not be relevant when other elements, such as *Clear feedback, Clear decision* and *Openness* are present within the team. Related work has also shown a difference in whether autonomy is deemed significant for PS, which might suggest that it depends on the control variables included in the test, since other elements of PS seem to make autonomy irrelevant. H6, supportive management, was also accepted, however as with autonomy, supportive leadership is also less relevant in regard to PS when controlled for the other x-variables. This suggests it is difficult to determine the effect of the individual predictors on PS because the effect of each predictor is difficult to distinguish from the others. This may imply that PS is a strategy, which does not rely solely on team characteristics but also on management's actions and the social support within the team.

5.0 Discussion

This section will discuss and detail our findings, in regard to literature reviewed in the related work section as well as other relevant literature. We will suggest perspectives on how the findings could help improve PS within the agile software practice. The discussion will be structured according to the quantifiable themes presented in our findings section.

5.1 External influences

The following section will discuss the findings of hypothesis 1 revised, 2 and 3 in relation to the articles found in the related work section. H1 revised asserts a positive relation between Colocated teams and PS. H2 argues for a positive connection between enduring teams and PS, while H3 states a positive association between Autonomy and PS.

5.1.1 H1 Revised – Colocated

H1 revised states a positive relation between teams being colocated and PS, which was rejected. The topic of colocated teams' influence on PS i agile software teams was found to be limited, through the research for the related work section. The topic was partially discussed in two different articles. Khanna & Wang (2022) suggested the possibility of high PS in online teams, where communicative tools were utilised, at least in the context of scrum retrospectives. Their findings shows that it is a possibility to improve PS through online teams when using online communication tools (Khanna & Wang, 2022, p. 47f). This supports our findings of PS being higher when working remotely, compared to a hybrid workplace. Holten, Hummel & Rosenkranz described the communication tools utilised in agile software development, placing importance on colocated office spaces in relation to creating a better communicative environment. They suggest colocated office spaces, as one of the means to achieve better communication, which they found to be a necessary factor for succeeding in agile software development (Holten et al., 2015, p. 273f). Our results go hand in hand with Holten et al.'s findings, suggesting that a colocated team has a higher PS compared to hybrid workplaces.

The two articles suggest the possibility that establishing communicative practices can help achieve better communication within the team. The communicative practices are different depending on whether it is a remote working team, or a colocated team. Seen in relation to our findings, working in a hybrid workplace influences PS negatively, compared to remote and

colocated. The hypothesis H1 revised is based on the article by Holten, Hummel & Rosenkranz, but our findings that a hybrid workplace has a negative relation to PS, contradicts the hypothesis. However, our findings did suggest that both Holten, Hummel & Rosenkranz, as well as Khanna & Wang's results are correct. Working in a hybrid environment has a negative impact on PS compared to colocated and remote work, however our results were not statistically significant for the *colocated* variable.

Organising an agile software team based on a colocated model can have a negative impact on the software engineering practice if the team combines colocated work and working from home. Thus, it is suggested that a team should either be working 'Less than 1 day per week' or 'Always colocated', so as to not influence the PS negatively. This might suggest that changes in work patterns disrupts a team's PS. Working pattern with a clear structure, and without combining mixed approaches, helps eliminate communicative problems, since the workers are getting familiar with a specific way of working. This might also suggest that it is not the element of working colocated that is the issue, but maybe rather that communication is affected by working without consistent patterns and approaches.

5.1.2 H2 – Enduring teams

H2 states that taking part of an enduring team increases the PS in the agile software team, which after testing with linear regression modelling is considered partially accepted, only one category were significant. Measuring an enduring team is an abstract concept that might not be possible to measure in a single variable containing an aspect of time. As other research shows, many elements might influence a team's level of maturity, among other enduringness (Gren et al., 2017, p. 28f). Bruce W. Tuckman has developed a theory explaining the development sequence in small groups, containing 5 stages of development: *Forming, Storming, Norming, Performing* and *Adjourning* (Tuckman, 1965, p. 396; Tuckman and Jensen, 1977, p. 426). The importance of Tuckman's theory is that it is not measured based on how much time a team has worked together, but rather a measurement of both maturity, abilities, relationships, and required leadership styles (Tuckman, 1965; Tuckman and Jensen, 1977). Our findings show that the measurement of enduring teams cannot be reduced to simply a measurement of time spent together.

Our results impact the practice of agile software engineering by casting clarity on the measurement of enduring teams, to be more complex than a measure of time. We did not find

a statistically significant relation between our measure of enduring teams and PS. However, Tuckman's measurement of enduringness, and team maturity encapsulates elements of the team not captured in our research (Tuckman, 1965, p. 396). Based on Tuckman's theory of team evaluation, we suggest that the measurement of enduring teams should contain elements of social interaction to influence PS. This is also supported by our concept of sensemaking, where social interaction and the identity of the individual team members influence the team development both structurally and individually (Weick, 1995).

5.1.3 H3 – Autonomy

H3 states a positive association between team autonomy and the level of PS, which was accepted during the quantitative analysis with a slight increase in PS and a high significance throughout all models. Buvik & Tkalich (2021) showed that team autonomy positively influenced a team's PS, which our analysis confirms. These findings imply that agile software teams with higher sense of autonomy also have higher PS which are corresponding with one of the principles of the agile manifesto about giving team members self-control and empowerment (Beck et al., 2001). Therefore, these findings can have an impact on the software engineering practice by underlining the benefits of following the agile principles and thereby get the side effect of a higher PS in a team. Our findings, despite being based on the same data, did find positive connections between PS and autonomy, whereas Alami and his colleagues found no statistically significant connection.

This may be because of Alami and his colleagues' autonomy scale, which contained a question less (Alami et al., 2023a). Alami and his colleagues omitted the question due to the question failing the AVE-test. We decided to keep the question due to its level of correlation within the autonomy scale. The removal of the question from the scale would also reduce the Cronbach's alpha value from 0,77 to 0,74. However, Alami and his colleagues did find connections in the qualitative phase of the project, equal to our qualitative findings.

Self-determination theory could be an additional consideration for the reason behind the positive effects of autonomy proven in this project. According to this theory, individuals are influenced by a need to feel like one's choices are meaningful, gained through basic psychological needs, which includes autonomy, competences, and relatedness (Ryan & Deci, 2017, p. 10). In this theory, autonomy is defined as feeling freedom over one's own decisions but without complete independence from others. This combination leads to a feeling of

motivation for the team members. Based on Ryan & Deci's theory, a low degree of autonomy could be understood to correspond to a low degree of PS, as PS, among other, is an expression of the ability to enact decisions without being blamed for them and feeling heard in one's opinions and ideas.

It is however worth noticing that other elements, not controlled for in the linear regression, might influence autonomy's association with PS. The other elements include among others the independent variables, which despite multicollinearity removes the positive influence provided by autonomy for PS (Appendix 6).

5.1.4 Implications

The theme external influences is constructed of three hypotheses focusing on whether structural elements affect a team's PS. The hypotheses are generally understood not to consist of social interactions in a direct context. H1 revised was falsified while H2 was partially verified, however we do not deem them significantly relevant for PS. H3 encapsulates autonomy, which is highly significant, but it also shows signs of being deemed irrelevant when other factors influencing PS are controlled for, see Appendix 6. When conducting an analysis including all the X variables, autonomy's influence changes and has a negative impact on PS. Because of multicollinearity the coefficients cannot fully be trusted, it is however possible to see patterns and directions in this analysis, which clearly shows autonomy to have a negative impact on PS (Appendix 6). Based on the following theme it is possible to assume the idea that; PS is influenced by social interactions, and is not, in general, affected by structures in the organisation, that does not influence communication between the team's individuals or management. This is assumed since every independent variable that contains clear elements of social interaction, H5, H6, H8 and H9, are highly significant in their influence on PS. H1 revised, H2 and H3 does not contain direct elements of social interaction in their measurement, H1 revised was falsified, while H2 is partially significant, while H3 is significant, but deemed irrelevant compared to the other independent variables. Our findings assess that social interaction is important for elements influencing PS, and a team should mainly focus on variables influenced by social interaction, to improve their PS.

The variable *Colocated* can partially be understood as an element of social interactions, with the understanding of working colocated to be affected by consistent communication and work structures which influences a team's social interactions. If focus should be on colocated or

virtual teams, it is recommended to establish clear structures, to secure knowledge and pattern in how team members communicate within the team, as to not disrupt a team members ability to work communicatively with their colleagues. Hence, we propose the following recommendations:

> **Recommendation #1:** To enhance PS in agile software teams, organisations should focus on establishing, supporting, and improving social interactions within the team and its managers.
>
> **Recommendation #2:** To enhance PS in agile software teams, focus should be on creating consistent work patterns either fully colocated, or fully virtual, as to secure clear structures for communication within the team.

5.2 Team environment

5.2.1 H5 – Openness in communication

H5 states that openness in communication has a positive relation with PS in an agile software team, which was found to be accepted during the quantitative analysis. Throughout the research for the related work section, openness was a key part of different research papers. Thorgren & Caiman (2019) analysed different implications in the implementation of agile principles. They found that openness in communication did not just emerge but needed to be fostered in the transition to an agile work environment (Thorgren & Caiman, 2019, p. 34ff). Bienefeld & Grote (2014) investigated openness and speaking up in a multiteam context. Within a team they found that PS was important to develop an open communication and the will to speak up, also outside the team. They also found that a leader's inclusiveness was important to gain this ability to speak up. Bienefeld and Grote (2014) found that having a psychologically safe environment allows for the benefits of openness to be realised, benefits such as being able to mention mistakes in high-risk contexts and speak up across the organisational lines (p. 930f).

Niklas Luhmann (2000), a system theorist within the field of sociology, has researched social systems. He understands social systems to be created by communication within the system and not by the individuals taking part of set social system. Social systems need communication about how to communicate, before they can develop and reproduce (Luhmann, 2000, p. 453). To develop openness in a team, it is relevant to have well established communication structures.

Through communication a team can establish a culture of openness, where they can speak openly in regard to one's ideas and be free from repercussions. Openness fosters a possibility for everyone to participate equally in the decisions and collaborate without hesitation.

Luhmann distinguishes between three main social system types: *Interactions*, *Organisations* and *Society* (Luhmann, 2000, p. 453). Based on our research the social system *Interactions* is connected well to agile teams, with the understanding that a team is formed by those members present in the team, everyone should participate in the collaboration within the team (Luhmann, 2000, p. 453). The understanding of it being communication that develops the social systems and its ability to accept openness and address different opinions, supports our findings that; PS in the agile teams are improved when the level of openness in communication improves. Our findings show that openness in communication has a positive relation to PS. Openness fosters an environment where people feel able to communicate and speak up about concerns or ideas, this is in direct contact to PS where they feel accepted, and their opinions are heard. However, based on Bienefeld & Grote research as well as Edmondson's definition of PS, it suggests that PS's relation to speaking up and openness have a bidirectional influence. Higher degree of openness in communication and the ability to speak up foster PS while PS also promotes openness in communication and improves the culture of speaking up.

5.2.2 Implications

By adapting a culture of openness in communication and speaking up within an agile software team, it benefits the engineering practices by minimising the team members fear of interpersonal risk taking and giving them an environment to speak freely. This could contribute to preventing problems getting too complex to handle efficiently.

Based on Luhmann's theory regarding social systems as well as related research, we suggest agile software teams establish common guidelines and structures about how to achieve open communication, and an environment where team members can speak up, and influence the process. These guidelines and structures can be gained through mutual communication about how best to achieve this in the specific team context. Every team is different, and it might require different approaches for each team to implement this culture.

To create a culture of openness in communication within a team, there are four main elements we deem relevant to partake in this culture. Firstly, it is relevant to consider how to handle

criticism and feedback from one's peers. The team should embrace constructive feedback, as a means to accept criticism from team members. Secondly, the culture should welcome new ideas and initiatives as an opportunity to develop the team and the final product. The team members should therefore embrace new ideas and initiatives as a valuable contribution when proposed and assess them as any other contribution. Thirdly, the team should focus on a well formulated and constructive argument when consulting an initiative or a rejection. The individual team assesses the guideline for a constructive and well formulated argument. Lastly, the culture should establish a ground for team members to voice their opinion and raise concerns and potential problems, without feeling afraid and vulnerable. The team members should feel empowered by the team and their leaders. These four elements, based on the quantitative data, creates the guideline for a culture that fosters openness in communication. Hence, we propose the following recommendation:

Recommendation #3: To enhance PS in agile software teams, managers and the team should create a culture of openness in communication based on the following guidelines:

1. Embrace constructive feedback to accept criticism.
2. Welcome ideas and initiatives.
3. Provide well formulated argument for initiatives and rejection.
4. Establish ground for team members to voice opinions and concerns.

Management should show the team how to communicate by setting examples and they should admit their own failures. By showing commitment, they would signal their own adherence to these important guidelines.

5.3 Management and stakeholder relationship

5.3.1 H6 – Management

H6 states that supportive management is positively associated with PS in an agile software team. This hypothesis was accepted. As such, supportive leadership is associated with a positive effect on PS. These results, in many ways, confirmed the prior scientific understanding of the impact and effects of management and leadership in the context of PS. As earlier noted by Thorgren & Caiman (2019), committed management is essential when creating an agile work environment with psychologically safety

(Thorgren & Caiman, 2019, p. 34ff). This was also captured in this project's management scale, as demonstrated in the question "*Our leadership 'practice what they preach' when it comes to psychological safety*". In a literature review, Frazier et al. (2017) summarises the primary tendencies within the field of PS. They identified leadership as a variable often understood within literature as an antecedent to PS (Frazier et al., 2017, p. 118f).

To discuss the impact of different leadership styles it is relevant to look at studies made by sociologist Daniel Goleman (2002). He suggests that the leader's way of acting can emotionally be transferred to the employees. Thereby Goleman is talking about two types of outcome produced by six identified styles of leadership; resonance and dissonance within a team (Goleman et al., 2002). Goleman suggests that certain styles create a more supportive environment than others, resulting in resonance. The coaching style focuses on how to support individual team members' development, whereas the affiliative style focuses on the whole team's well-being and strives towards creating harmony (Goleman et. al., 2002). These styles can be seen as a means to creating a more supportive management which increases the feeling of PS in a team environment. On the other hand, Goleman connects certain leadership styles that can easily promote dissonance and an unsupportive leader which is the commanding and the pacesetting styles. Where the commanding style is about the leader as the one who is giving commands, the pacesetting style is heavily focusing on performance and result. Thereby in these styles failures are a matter of placing the blame and not on learning and creating knowledge (Goleman et al., 2002). Based on Goleman's theory, we suggest utilising the leadership styles of coaching and affiliative, to enhance a team's PS.

Our results for H6 echo the results of earlier projects: Management should be willing to support their developers through thick and thin, and help to foster a culture of acceptance, and where risk-taking is rewarded, and punishment is minimised (Thorgren & Caiman, 2019).

5.3.2 Implications

Edgar Schein (2004) states the connection between culture and leadership as *"These dynamic processes of culture creation and management are the essence of leadership and make one realize that leadership and culture are two sides of the same coin"* (Schein, 2004, p. 1). This implies that to get a more psychological safe work environment, a leader should focus on

creating a more supportive culture that rewards risk-taking and practices no blame. To create such a culture, a manager should reflect upon their own practice of leadership and adjust their leadership style according to their current team's environment. Hence, we propose the following recommendation:

> **Recommendation #4:** To enhance PS in agile software teams, managers and leaders should evolve a leadership style which provides frames to support employees. Leaders should show PS by actions, i.e., admitting their own failures, rewarding risk-taking and seeking feedback in their own decisions.

5.4 Clear feedback

The following section will include a discussion of hypothesis 8 and 9, regarding clear decision process and clear feedback structures.

5.4.1 H8 – Clear decision process

H8 states that clear decision processes result in higher PS in agile software teams. H8 was also accepted based on the linear regression models. The findings of this project indicates that clear decision processes should be integrated and thought into the work process. Hennel & Rosenkranz's (2021) earlier study suggested that agile practices are self-reinforcing, in that agile practices improve PS, and PS improves agile practices (Hennel & Rosenkranz, 2021, p. 16, p. 22ff). As such, if a clear decision process is accepted as an agile practice, well-functioning decision processes will improve PS, and in turn, PS will improve the decision processes. Uncertainty in the structure of decision making should thus be rejected, in favour of clear and obvious structures and presences.

Our clear decision process scale is based on questions that assess if the decisions are founded in the development teams themselves and on an equitable basis. This implies a decision process founded in communicative rationality, which places the responsibility of a decision on the common understanding of a team or group (Jacobsen & Thorvik, 2014, p. 302). By communicative rationality, clear decision processes should be understood to facilitate a common ground for discussions, an equal communicative relationship between the team members, and well-formed arguments and rejections. These elements are relevant in the

implementation of a clear decision process within a team. Clear decision processes build upon an environment with low levels of hierarchy as well as leadership that empowers the workers to argue and evaluate the different suggestions equally. Clear decision process is created by many elements also relevant to a PS environment, such as the ability to express opinions without discouragement and the inclusion of new initiatives.

5.4.2 H9 - Clear feedback structures

H9 states that clear feedback structures is positively associated with PS in agile software teams and was accepted. The scale encapsulating clear feedback consists of questions regarding a no blame culture within the team, as well as the ability to learn from one's mistakes, and improve one's work in the future. The *Clear feedback* scale aims to describe how team members provide feedback, as well as how they handle set feedback to improve. Hennel & Rosenkranz finds that a culture that practises *no blame* and *helping others learn* improves the PS of a team, and refines the agile practices performed (Hennel & Rosenkranz, 2021, p. 16, p. 21ff). H9 supports the findings of Hennel & Rosenkranz, based on the assumption that focusing on not blaming others and learning from mistakes, is an indicator for a team's ability to provide sufficient feedback without misunderstandings, blame and not improving (Edmondson, 1999, p. 370f s).

A team should implement clear feedback structures as a way to institutionalise learning with guidelines and a mutual understanding, as a tool to further improve their PS. When analysing one's feedback, a possibility would be to utilise the models of single and double loop learning, developed by Chris Argyris, which encapsulates the process of learning within an organisational context. Single loop learning is using feedback as a means to regulate *what* is learned from an experience, where double loop enables a reflection of *why* we learned set learning (Argyris, 1976). The difference between single loop learning and double loop learning is in the reflection of why we learned this. To establish clear feedback structures, it is relevant to remember the reflection process, this will let the team learn from their mistakes, and reduce repetitive mistakes within the team.

5.4.3 Implications

By implementing a clear decision process with focus on communicative rationality, the engineering practice will gain more self-managed teams and increase the number of initiative within the teams. The clear decision process will establish mutual ground for decisions and

provide the team with the ability to communicate about decisions in a constructive and tangible approach. The decision process and the team is dependent on a leader, which supports and indulges the team members in the necessary steps needed to establish set processes. The process should build upon collective decision making, it is not the individual team members that decide but the team as a unit. This is based on the questionnaire used to capture Clear decision, where focus is on the collective decisions and the argumentation of one's decisions. If the team is granted the responsibility of making decisions and creates an environment, in collaboration with the leader, where argumentation and equality are the main values, it will improve the PS for the team.

Through establishing clear feedback structures, team members will gain insight into *why* they learn from experience, and they will minimise the number of repetitive mistakes when working. Clear feedback structures within a team should be analysed through double loop learning, where they reflect upon the process of learning, while refining their feedback in a constant loop of improvement. To build clear feedback structures, implementation of a no blame culture is important, as it creates a foundation for assignment-based feedback while eliminating the use of ad hominem defence strategies. Another aspect of clear feedback structures is learning from mistakes, which is implemented through double loop learning, and facilitates the ability to adjust and improve the process of feedback, since it creates the continuity of learning. Clear feedback structures should be implemented cooperatively between managers and team members, with discussions about how feedback is to be contextualised within the team. Hence, we propose the following recommendations:

Recommendation #5: To enhance PS in agile software teams, managers and the team should collectively agree upon the structures of their decision processes. They should establish guidelines for good argumentation, and be measured equally, independent of role.

Recommendation #6: To enhance PS in agile software teams, clear feedback structures should be established in collaboration between managers and team members. The feedback argumentation should reflect upon the initiative and eliminate the use of ad hominem defence strategies.

5.5 Control variables

The following section will discuss the control variables in relation to PS and the related work referenced in previous sections. The control variables will be discussed in the same order as they were controlled for in the linear regression models.

During our linear regression we did not find *Gender* to have a significant relation with PS. In contrast to the findings of Verwijs & Russo's (2023), who found that a higher degree of women in an agile software team, increases the risk of relational conflicts. This might suggest that even though gender diversity increases the risk of relational conflicts, it does not affect their general level of PS. However, our measure of *Gender* does not encapsulate the categories of 'Non-binary', as the category 'Non-binary' only included one individual and was therefore excluded based on low representativity. *TeamSize* did not seem to affect the PS of any team members.

The control variable *Sourcing model* has a varying degree of significance, depending on the independent variable. Agile software team members working with 'In-house' software have a higher degree of PS compared to team members working 'Outsourced' or 'Mixed'. However, its influence on PS is less significant when controlled for *Experience* and *Role. Sourcing model* is not significant, when tested together with the independent variables: *Management, Clear decision,* and *Clear feedback.* This suggests that high levels of management support, as well as clear structures and processes to handle decisions and feedback, lowers the issues related to working with outsourced software.

Experience, which is highly significant for most of the tested hypotheses, showing a connection between PS and experience. Our linear regressions show a rise in PS during the first 8 years of experience. PS is most improved within the first couple of years, later stabilising around 8-11 years of experience. This suggests that *Experience* has a high degree of influence on PS, and that its influence is mainly within the beginning of one's career.

When testing the control variable *Role*. The category 'Senior software engineer' is associated with higher levels of PS compared to the reference group 'Software developer', and is significant in all models. 'Software engineer' is also positively associated with PS in comparison to the reference group but is only significant within a few of the hypotheses. However, the title of 'Senior software engineer', might also suggest a higher level of experience compared to the reference group 'Software developer'.

The control variable *Multiple teams* do not have a significant relation to PS in any of our linear regression models, suggesting that PS is not affected by team members working on projects that involve multiple teams.

5.6 Future work

This thesis presents multiple potential avenues for further scientific research within the field of PS and agile software teams. There were four hypotheses that we were unable to test using the quantitative data provided: H1, H4, H7, and H10. While this thesis was unable to accurately measure the effects of H1, geographically dispersed teams, it would provide ample opportunity to conduct further research. To measure the effect of geographically dispersed teams, the following question could be utilised: *does any of your team members work in different timezones?* It would also be prudent to question whether the team members themselves consider working in different timezones to be an issue, as to best continue the work started by Khanna & Wang (2022) regarding online workplaces.

H4 regarding team values could be analysed and researched utilising a method such as action research. Shared team values are socially created, and quantitative research might miss important parts of the subject, since additional questions cannot be asked. Action research can investigate the value first hand within the team, in their local environment. Values do not have a common understanding, and as such it is necessary to research the individuals' teams understanding of their values.

H7, open stakeholder communication, could be measured utilising the approach of the current thesis, by rephrasing some of the questions in the existing survey as to measure stakeholder communication. Stakeholder communication should be measured based on the relationship between the team members and the stakeholder, as viewed by the team. It should focus on the amount and type of communication exchanged between the two parties.

Hypothesis 10, Clear expectations, could be relevant to research in the context of a company-wide questionnaire, where the questions would be able to capture whether clear expectations were placed upon them by their management. It would also be relevant to question management on whether they felt that they provided clear expectations to their employees and compare the results between managers and employees. The variable clear expectations might be influenced by one's relationship to their managers and the communication established between them. It is

therefore relevant to distinguish between management communication and clear expectations in regard to work and assignments.

There were three hypotheses that we were able to test but which could also benefit from future research. Including H2, H8 and H9. Our suggestion for H2 would be to conduct a longitudinal study, where a newly formed agile team was followed for a longer period, to note their changes in agile practices and behaviours. This would better capture the different periods of an agile team.

Hypothesis H8 and H9 would benefit from being studied through action research. These two hypotheses are mostly organisationally bound, focusing on clear decision and clear feedback respectively. Processes can often be difficult to research and understand, and especially when they interact with multiple parts of an organisation. By conducting action research on agile software teams, an insight could be gained into the process of decisions, finalised with an intervention to establish a clear decision process or clear feedback structures.

When discussing H1 and H2 it is relevant to consider the lack of social interaction included in the independent variables. PS is primarily constructed within a social context and exists in the belief that one can share ideas, and take interpersonal risks (Edmondson, 1999). The idea of PS does not make sense, when measured in a non-social context. However, non-social elements might influence the clues and the social context used by the team members to establish one's perceived reality. Suggesting they have an indirect influence on PS. Lastly, a large possibility exists for further studies and inquiries into the field of PS and the identified themes.

6.0 Threats to validity

The following sections are composed of two subsections, internal validity, and external validity. Internal validity will clarify possible bias in the research and how it was handled as to affect the data the least. External validity will explain the troubles of drawing causal conclusions based on our mixed method studies, and also how we have tried to counter these problems.

6.1 Internal validity

Phase 1 consists of 12 interviews collected by Adam Alami and his colleagues. The interviewees were sourced through word of mouth and referrals rather than a random sample. It is therefore possible that the sample has problems with selection bias. This bias however is somewhat minimised with the mixed method approach, where neither analysis is used alone. The phase 2 data collection was also performed by Adam Alami and his colleagues, where there was focus on sampling a large group of people with various experience, occupation, age, and national background. See survey characteristics in Table 3 in section *3.4 The quantitative phase.* They used an external market platform, Prolific, when collecting their data which might influence the data quality negatively, however they screened for and recruited participants who would provide reliable data. They also reviewed the answers before the payment was accepted on the prolific platform (Alami et al., 2023b, p. 111:30f). Further quality checks were performed by us to ensure that the data quality, as explained in section *3.4 The quantitative phase*.

The participants were chosen restrictively based on certain areas, mainly native English-speaking countries or where English is the second language, e.g. South Africa and Europe. Meaning other areas' perspectives may not be included in this research (Alami et al., 2023b, p. 111:31). However, the data in phase 1 includes interviewees with different backgrounds, and where 8 participants come from countries where English is not the native language. As an example, some of our interviewees compare their work environment with that of European or American countries.

The qualitative study only includes one female (P18), indicating a majority of males in the answers of the phase 1 analysis. However, our qualitative analysis has been performed without

knowledge regarding the participants gender, and no focus has been put on gender during the findings. The variance in gender during the qualitative analysis is still relevant in regard to generalising the answers, however this has been partially improved by utilising mixed methods. The quantitative data from phase 2 includes a higher portion of females, where females covered 19% of the participants. When preparing the data for linear regression analysis, the 'Non-binary' category was removed, and the single participant that had answered set category was removed from the dataset. When controlling for gender, a category with a single participant is not generalisable, and it would also not be accurate to add the non-binary person to the females or males, since it could blur the data results. However, based on the gender variable it was possible to see and control for one's identified gender (Male and Female) influence on PS.

Social desirability bias could be a threat to the validity of the research. In the qualitative analysis these biases were tried minimised by encouraging the participants to give examples from their team experiences. The examples help reduce both memory bias, since it gives the interviewee a possibility to better remember their experience, but it also serves as a trustworthy description in one's perceived reality. The participants were also assured in the qualitative data analysis that their answers were anonymous, and completely sealed (Alami et al., 2023b, p. 111:31).

When working with scales it is important to remember the validity of the scales, whether the scale actually measures the instance that is aimed for. Adding extra questions to already collected data is not possible, and the scales are therefore built based on the variables at hand. We have tried to evaluate the scale as thoroughly as possible, to be sure that the measurement in hand is consistent with the findings of the qualitative data. As an example, the *Clear feedback* scale should capture the will to help others improve and to provide honest feedback without consequences. The scale was therefore built upon the categories: *No Blame* and *Learning from Mistakes.* However, some minor adjustments might have improved the scale, for example questions such as; *Do you feel free to give feedback within your team?* or *Is the feedback you receive from your colleagues constructive and useful?*

Another important aspect is reliability in regard to the performed scales, where a part of it includes interests in whether the variables measure the same concept. This was among other tested with the Cronbach's alpha value, and a correlation test, these tests were performed for all the scales used in the project (Drost, 2011, p. 111). All the scales have a Cronbach's alpha

value of 0,7704 or above, meaning they are assessed as good at measuring the same instance. The scale for PS only has a Cronbach's Alpha value of 0,6894, which is not as good. However, when creating scales, it is important to also remember the theory behind the measurement, which in our case helps establish validity to our PS scale. Alami and his colleagues used Edmondson's well-established scale, which is used and tested in several studies helping establish a higher validity for our PS scale. Alami et al.'s questionnaire adapted the PS scale with a minor change to a question. Question PS5 is phrased positively, but based on Edmondson's questionnaire it should have been negative phrases (Edmondson, 1999), for reference see Figure 1.

6.2 External validity

When working with linear regression it is necessary to understand the connection between the independent and the dependent variables, and to aim for the knowledge of exogeneity. However, it is difficult to ensure this presumption when working with correlational studies. The impact could come from multiple other places, among other, omitted variables bias or from simultaneity bias. When using already collected research it is not possible to add new variables to the data. Alami and his colleagues have however implemented a robust set of control variables to be used in the analysis, helping to eliminate some of the fear from omitted variable bias. We have utilised a mixed methods approach, among others to ensure insight into the causal processes of the interviewees. Combining the qualitative data with the quantitative data have helped us gain an insight into the direction of the causality, and to provide a stronger foundation when giving causal assertions.

7.0 Conclusion

The purpose of our master thesis was to investigate how agile software teams' PS is affected by different characteristics within their organisation. Based on a mixed method approach the following research question has been answered:

> How do team's organisational characteristics influence psychological safety in an agile software team?

We found that organisational characteristics guided by social interactions have a higher influence on PS compared to other elements researched in this thesis, i.e. question about working colocated and a team's enduringness. Common for the characteristics guided by social interactions is a focus on communication within the team, and an openness for the team members to express doubt, concerns, and ideas on equal footing. The practice of management was found to be a powerful antecedent on a team's PS, with influence on both open communication, clear decision processes and clear feedback structures.

Based on our findings we draw implications relevant for the software engineering practice. The implications mainly express the importance of collaboration between the team members and management, as to better facilitate PS. Both parties need to be involved in the process of establishing a psychological safe work environment. Management should be supportive of their team to increase their PS. This entails a focus on the leadership style and for managers to lead by example. Our findings suggest that teams working fully remote or colocated, have a higher level of PS compared to hybrid working teams. Suggesting that changing work patterns influence the team members ability to communicate within the team, resulting in a lower degree of PS.

When analysing autonomy, we found it to have a positive relation to PS. However, when variables also influencing PS are accounted for, autonomy's effect is subsumed. Autonomy cannot exist alone in an organisational context, the relevance of management, openness etc. will always be present. Multiple organisational structures influence PS which cannot be reduced to a single concept and includes actions from management and social support from within the team.

8.0 Bibliography

A. Haraty, R., & Hu, G. (2018). Software process models: a review and analysis. *International Journal of Engineering & Technology, 7(2.28)*, 390. https://doi.org/10.14419/ijet.v7i2.28.14537

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

Alami, A., Zahedi, M., and Krancher, O. (2023b). How does psychological safety influence agile teams' pursuit of achieving software quality?. *J. ACM 37, 4, Article 111*. https: //doi.org/XXXXXXX.XXXXXXX

Argyris, C. (1976). Single-Loop and Double-Loop Models in Research on Decision Making. *Administrative Science Quarterly, 21(3)*, 363–375. https://doi.org/10.2307/2391848

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

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

Berger, P. L., & Luckmann, T. (2004). *Den sociale konstruktion af virkeligheden : en vidensociologisk afhandling* (3rd Edition). Akademisk Forlag.

Biehler, J., & Przybilla, L., & Krcmar, H. (2022). "Primus inter Pares"?—The Perception of Emergent Leadership Behavior in Agile Software Development Teams. 10.24251/HICSS.2022.879.

Bienefeld, N., & Grote, G. (2013). Speaking up in ad hoc multiteam systems: Individual-level effects of psychological safety, status, and leadership within and across teams. *European Journal of Work and Organizational Psychology, 23(6)*, 930–945. https://doi.org/10.1080/1359432x.2013.808398

Bo, I. G., & Jacobsen, M. H. (2015). Hverdagslivets følelser - tilstande, relationer, kulturer. In *Hverdagslivets følelser: tilstande, relationer, strukturer* (1st ed., Vol. 1, pp. 13-46). Hans Reitzels.

Buvik, M., & Tkalich, A. (2022). Psychological Safety in Agile Software Development Teams: Work Design Antecedents and Performance Consequences. *Proceedings of the Annual Hawaii International Conference on System Sciences.* https://doi.org/10.24251/hicss.2022.880

Buzaglo, G., & Wheelan, S. A. (1999). Facilitating Work Team Effectiveness: Case Studies from Central America. *Small Group Research, 30(1),* 108–129. https://doi.org/10.1177/104649649903000106

Creswell, J. W. (1999). Chapter 18 - Mixed-Method Research: Introduction and Application. In Handbook of Educational Policy (pp. 455-472). Academic Press. https://doi.org/10.1016/B978-012174698-8/50045-X

Dag Ingvar Jacobsen, Thorsvik, J., & Jørgensen, S. W. (2014). *Hvordan organisationer fungerer : en indføring i organisation og ledelse*. Hans Reitzel.

Dreyer, J., K. & Hussain, M., A. (2016). "*Multipel lineær regressionsanalyse*". In Kristensen, C. J. & Hussain, M., A. (red.): *Metoder i samfundsvidenskaberne.* København: Samfundslitteratur

Drost, E.A. (2011) Validity and Reliability in Social Science Research. *Education Research and Perspectives*, *38*, 105- 123.

Edmondson, A. (1999). Psychological Safety and Learning Behavior in Work Teams.*Administrative Science Quarterly, 44(2)*, 350–383. https://web.mit.edu/curhan/www/docs/Articles/15341_Readings/Group_Performance/Edmondson%20Psychological%20safety.pdf

Frazier, M. L., Fainshmidt, S., Klinger, R. L., Pezeshkan, A., & Vracheva, V. (2017). Psychological Safety: A Meta-Analytic Review and Extension. *Personnel Psychology, 70(1)*, 113–165. https://doi.org/10.1111/peps.12183

Goleman, D., Boyatzis, R., & McKee, A. (2002). Følelsesmæssig intelligens i lederskab, Børsens Forlag, 2. udgave.

Gren, L., & Goldman, A., & Jacobsson, C. (2019). The Perceived Effects of Group Developmental Psychology Training on Agile Software Development Teams. *IEEE Software.* PP. 10.1109/MS.2019.2955675.

Gren, L., Torkar, R., & Feldt, R. (2017). Group development and group maturity when building agile teams: A qualitative and quantitative investigation at eight large companies. *Journal of Systems and Software, 124,* 104–119. https://doi.org/10.1016/j.jss.2016.11.024

Guest, D. E. (2017) Human resource management and employee well-being: towards a new analytic framework. *Human Resource Management Journal, 27,* 22– 38. doi: 10.1111/1748-8583.12139

Gupta, M., George, J. F., & Xia, W. (2019). Relationships between IT department culture and agile software development practices: An empirical investigation. *International Journal of Information Management, 44,* 13–24. https://doi.org/10.1016/j.ijinfomgt.2018.09.006

Gustavsson, T. (2022). Team Performance in Large-Scale Agile Software Development. In: , et al. Advances in Information Systems Development. *Lecture Notes in Information Systems and Organisation, vol 55.* Springer, Cham. https://doi.org/10.1007/978-3-030-95354-6_14

Hastuti, R., & Timming, A. R. (2021). An inter-disciplinary review of the literature on mental illness disclosure in the workplace: implications for human resource management. *The International Journal of Human Resource Management, 32(15)*, 3302–3338. https://doi.org/10.1080/09585192.2021.1875494

Hennel, P., & Rosenkranz, C. (2020). Investigating the "Socio" in Socio-Technical Development: The Case for Psychological Safety in Agile Information Systems Development. *Project Management Journal, 52(1),* 875697282093305. https://doi.org/10.1177/8756972820933057

Hummel, M., Rosenkranz, C., & Holten, R. (2015). The Role of Social Agile Practices for Direct and Indirect Communication in Information Systems Development Teams. *Communications of the Association for Information Systems, 36.* https://doi.org/10.17705/1cais.03615

Kahn, W. A. (1990). Psychological Conditions of Personal Engagement and Disengagement at Work. *Academy of Management Journal, 33(4)*, 692–724.

Kakar, A. K. (2018). How do team cohesion and psychological safety impact knowledge sharing in software development projects? *Knowledge and Process Management, 25(4),* 258–267. https://doi.org/10.1002/kpm.1584

Khanna, D., & Wang, X. (2022). Are Your Online Agile Retrospectives Psychologically Safe? the Usage of Online Tools. *Lecture Notes in Business Information Processing,* 35–51. https://doi.org/10.1007/978-3-031-08169-9_3

Lenberg, P., & Feldt, R. (2018). Psychological safety and norm clarity in software engineering teams. In *Proceedings of the 11th international workshop on cooperative and human aspects of software engineering* (pp. 79-86).

Luhmann, Niklas. *Sociale systemer: Grundrids til en almen teori.* Hans Reitzels Forlag, 2000.

Marder, B., Ferguson, P., Marchant, C., Brennan, M., Hedler, C., Rossi, M., Black, S., & Doig, R. (2021). "Going agile": Exploring the use of project management tools in fostering psychological safety in group work within management discipline courses. *The International Journal of Management Education, 19(3),* 100519. https://doi.org/10.1016/j.ijme.2021.100519

Marnada, P., Raharjo, T., Hardian, B., & Prasetyo, A. (2022). Agile project management challenge in handling scope and change: A systematic literature review. *Procedia Computer Science, 197*, 290–300. Sciencedirect. https://doi.org/10.1016/j.procs.2021.12.143

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

Oyibo, C. O., & Gabriel, J. M. O. (2020, August). Evolution of Organization Theory: a Snapshot. International Journal of Innovation and Economic Development, 6(3), 46 - 56. 10.18775/ijied.1849-7551-7020.2015.63.2005

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

Popper, K. R. (1996). Kritisk rationalisme : udvalgte essays om videnskab og samfund. (K. Haakonssen & N. C. Stefansen, Trans., S. Hansen, Ed.) (1st ed.). Nyt Nordisk Forlag.

Ryan, R. M., & Deci, E. L. (2017). *Self-determination theory: Basic psychological needs in motivation, development, and wellness.* Guilford Press.

Saldaña, J. (2021). *The Coding Manual for Qualitative Researchers* (4th ed.). SAGE

Schein, E. H. (2004). *Organizational Culture and Leadership* (3rd ed.). Wiley.

Strode, D., Dingsøyr, T., & Lindsjorn, Y. (2022). A teamwork effectiveness model for agile software development. *Empirical Software Engineering, 27(2),* https://doi.org/10.1007/s10664-021-10115-0

Thorgren, S., & Caiman, E. (2019). The Role of Psychological Safety in Implementing Agile Methods across Cultures. *Research-Technology Management, 62(2),* 31–39. https://doi.org/10.1080/08956308.2019.1563436

Treiman, D. J. (2009). *Quantitative Data Analysis: Doing Social Research to Test Ideas.* Jossey-Bass.

Tuckman, B. W. (1965). Developmental sequence in small groups. *Psychological Bulletin, 63(6),* 384–399., https://doi.org/10.1037/h0022100.

Tuckman, B. W., & Jensen, M. A. C. (1977). Stages of Small-Group Development Revisited. *Group & Organization Studies, 2(4),* 419–427. http://faculty.wiu.edu/P-Schlag/articles/Stages_of_Small_Group_Development.pdf

Verwijs, C., & Russo, D. (2023). The Double-Edged Sword of Diversity: How Diversity, Conflict, and Psychological Safety Impact Agile Software Teams. *arXiv preprint arXiv:2301.12954.*

Weick, K. E. (1995). *Sensemaking in organizations.* Sage Publications ; London.