Shared Leadership and Team Performance: Psychological Safety as Main Protagonist?

Korte, Chanda de

Shared Leadership and Team Performance

Psychological Safety as Main Protagonist?

Leiden University, Faculty of Governance and Global Affairs

Master Thesis for Public Administration: Public Management and Leadership

Student name: C.J. de Korte

Student number: S2903679

Date: January 7th, 2022

Supervisor: Prof. dr. J. van der Voet

Word count: 10106

Acknowledgments

I want to thank…

… first and foremost, my best friend/SPSS-magician Esmée, for helping me interpreting the statistics. I honestly would not have been able to finish this thesis without the combination of your help, your coffee and a lot of your chocolate.

… second, Laurens, for being a good big brother doing a spelling-check.

… and last but not least, thank you Dr. Joris van der Voet, for the excellent supervision and your critical yet positive feedback. Your contributions boosted my motivation after every feedback session.

Table of contents

List of tables and figures

Tables

Figures

1. Introduction

1.1. Background

Individual talents win games, but teamwork and combined intelligence win championships. Working in teams is increasingly important in a dynamic organizational environment, which is rapidly changing in this era. Individual perspectives, skills and knowledge, are crucial in teams in order to solve problems or to reach goals set by the organization. In other words, by using horizontal interaction, sharing the leadership within a team, the team's knowledge contributes to the organizational performance (Edmondson, 2012). However, are those individual perspectives and skills fully appreciated if there is no room to feel safe in a team? The belief that a team member will not be punished or humiliated for speaking up is a vital condition for a team to function. After conducting research for two years, Google researchers confirmed this claim by publishing a list of five key dynamics of what makes a perfect and efficient team. The first and foremost dynamic is, without a doubt for the researchers, psychological safety (Duhigg, 2016).

There is a need for sharing responsibilities in the Dutch public sector. According to the Dutch Senior Civil Service, "The public leader puts shared leadership into practice, is focused on the broader context and not exclusively his/her "own" domain, actively seeks collaboration and co-creation and is able to understand various perspectives" (Bureau for the Senior Civil Service, 2017, p. 2). To understand perspectives of other teammates, there must be a safe climate for teammates to share their perspectives.

A team is psychologically safe when the team in question is safe for interpersonal risk-taking. Edmondson argues that a team is not directly psychologically safe, when psychological safety is explicitly being discussed within the team (Edmondson, 1999, p. 354). Moreover, teams with a high level of psychological safety tend to feel respected and are more likely to respect other members of the team. Because of this, all members of the team feel safe to take risks and even make errors. In teams where the degree of psychological safety is low, team members are worried to take risks, as they are afraid to be humiliated or punished should they make an error (Edmondson, 1999). Thus, psychological safety is more than merely feeling safe for interpersonal risk-taking: it is also about team members respecting each other in order for a team to function optimally.

Hand in hand with psychological safety goes shared leadership. Shared leadership is defined by Pearce as the distribution of leadership practices on all levels instead of only at the "heroes" on top (Pearce & Conger, 2003, p. 22). In other words, those heroes on top are being supported by team members on all levels, distributing the leadership practices among the team. Apart from the distribution of leadership, Pearce also names social interaction as a key concept within shared leadership, as shared leadership occurs in and through relationships between team members.

Furthermore, Pearce claims that shared leadership is needed for learning (Pearce & Conger, 2003, p. 23). Here is where the connection between shared leadership and psychological safety becomes clear. Pearce argues that conditions need to be created for collective learning to happen, one of those conditions being a safe space for members (Pearce & Conger, 2003, p. 24). By creating this safe space, individuals within a team feel safe to learn from their mistakes and make their shared leadership stronger.

By enforcing psychological safety and shared leadership separately, many authors wrote about the outcome of team performance. Team performance is defined as a product of team members working together to reach goals, by using their pool of individual and team skills (Salas & et al., 2008, p. 541). Team performance can be measured through assessing team effectiveness. For example, Kim et al. wrote about how psychological safety affects team performance through efficacy and learning behaviour (Kim, Lee, & Connorton, 2020), and Han wrote about the effects of shared leadership on team performance (Han & et al., 2021). However, little is written about the dynamic between psychological safety, shared leadership and team performance.

The lack of literature about the dynamic between those three variables poses an opportunity. Therefore, this thesis builds on the existing literature of shared leadership and team performance. How does shared leadership affect team performance? What exactly is their relationship, and is it a positive one? Does psychological safety have a role as mediator in the relationship between those variables? And if yes, how does psychological safety exactly mediate the relationship?

1.2. Research question

In order to answer those questions and to enrich existing literature, the research question is as follows:

'How does psychological safety mediate the potentially positive relationship between shared leadership and team performance in Dutch welfare teams?'

In this thesis, the focus of shared leadership lies on the supporting-relations behaviour as described by Yukl in his 2002 hierarchical taxonomy (Yukl, 2002). The rationale behind this choice is that shared leadership and psychological safety connect closely to this leadership behaviour, as the leadership behaviour described in this thesis is conducted by the team members as a shared task, for which a psychological safe environment is needed. The aim of this study is to further extend the knowledge of the relationship between shared leadership and team performance.

In order to answer the research question, two datasets will be utilized. First, a dataset made out of a survey with several hundred workers in about 90 welfare teams will be analysed. In this survey, it is tested how these workers perceive, among other matters, the shared leadership of their team as well as psychological safety. Second, another dataset is created on the basis of a survey conducted among the supervisors of those teams. Moreover, from the second survey, information is gained about how supervisors perceive team performance in their teams.

1.3. Scientific relevance

It seems that few researchers have addressed the question of psychological safety's role as a mediator to this relationship. The correlational results arising from researching those variables will complement existing literature, filling a part of the literature-gap discussed in the conclusions of many articles. Kim et al. discuss how psychological safety affects team performance; however, in that research the relationship is the other way around and use efficacy and team learning are used as mediators (Kim, Lee, & Connorton, 2020). Moreover, Parker et al. studied how psychological safety can be increased and how it empowers team performance, therefore using psychological safety as an independent variable, and not as a mediator (Parker & du Plooy, 2021).

All in all, the scientific relevance of this thesis can be found in the gap in literature. There is limited literature about shared psychological safety affecting shared leadership. In general, most literature on this topic includes shared leadership and team performance solely. By including psychological safety as a variable, new theories can be drafted.

Moreover, as this thesis will analyse the relationship between shared leadership, team performance and psychological safety in the setting of Dutch welfare teams, a new door is opened in the field of welfare workers. The dynamic between those three variables in that specific setting can lead to crucial, new insights in the field, making this thesis a building block for further research.

1.4. Societal relevance

Theses should not only be scientifically relevant: supervisors of welfare teams wishing to develop and improve their leadership skills and to understand their team better will be advantaged by understanding this explanatory research. It is important to note that psychological safety is crucial for contemporary organisations in the public sector. The government of West-Australia, for example, acknowledges psychological safety as essential factor in their working teams. They argue that trust is critical for building psychological safety, which in turn can lead to increased capability and performance of the team (The Government of Western Australia, n.d.). Moreover, by finding out the possible mediating effect of psychological safety on the relationship between shared leadership and team performance, team managers can adapt their level and knowledge of psychological safety to, in the end, achieve a high degree of team performance. Also, the results will benefit team members themselves, as they can test the effects of psychological safety in their team as well.

Hence, after reading the results of this thesis, team managers can decide whether their team requires a higher or lower degree of psychological safety.

1.5. Roadmap

The content of this thesis is divided into five chapters, starting after this introduction with the second chapter on theory, which consists of a review of existing literature on the variables, a theoretical framework and the proposed hypotheses. After the theory, the concepts will be operationalized, and the research design used will be explained in the third chapter. In the fourth chapter, the results will be portrayed, after which it is time for the analysis. Lastly, a reflective discussion and conclusion will be given in chapter five, in which the thesis will be summarized, the research question will be answered and limitations as well as suggestions for future research will be given.

2. Theory

This chapter presents the theory from which the hypotheses will be derived. It entails the scientific foundation required to provide understanding into theoretical theories that help answering the research question *How does psychological safety mediate the potentially positive relationship between shared leadership and team performance in Dutch welfare teams?* First, the concept of team performance will be discussed. Additionally, in the same section, the Input, Process, Output framework will be introduced. Then, the concept of shared leadership will be reviewed. Furthermore, the notion of psychological safety will be elaborated on.

2.1. Team performance

Firstly, it is useful to define the concept of a *team*. In 1984, Dyer was among one of the first to conceptualize what a team is. According to Dyer, teams are social units composed of members with high task interdependency and collective, respected goals (Dyer, 1984).

Only few studies have been published on the actual definition of *team performance*. Fortunately, Kozlowski and Klein offer a definition of team performance, that will be used as conceptualization for this thesis. They conceptualize team performance as a complicated function of specific individual and dyadic-networked-contributions, emerging from the behaviours of individual team members (Kozlowski & Klein, 2000).

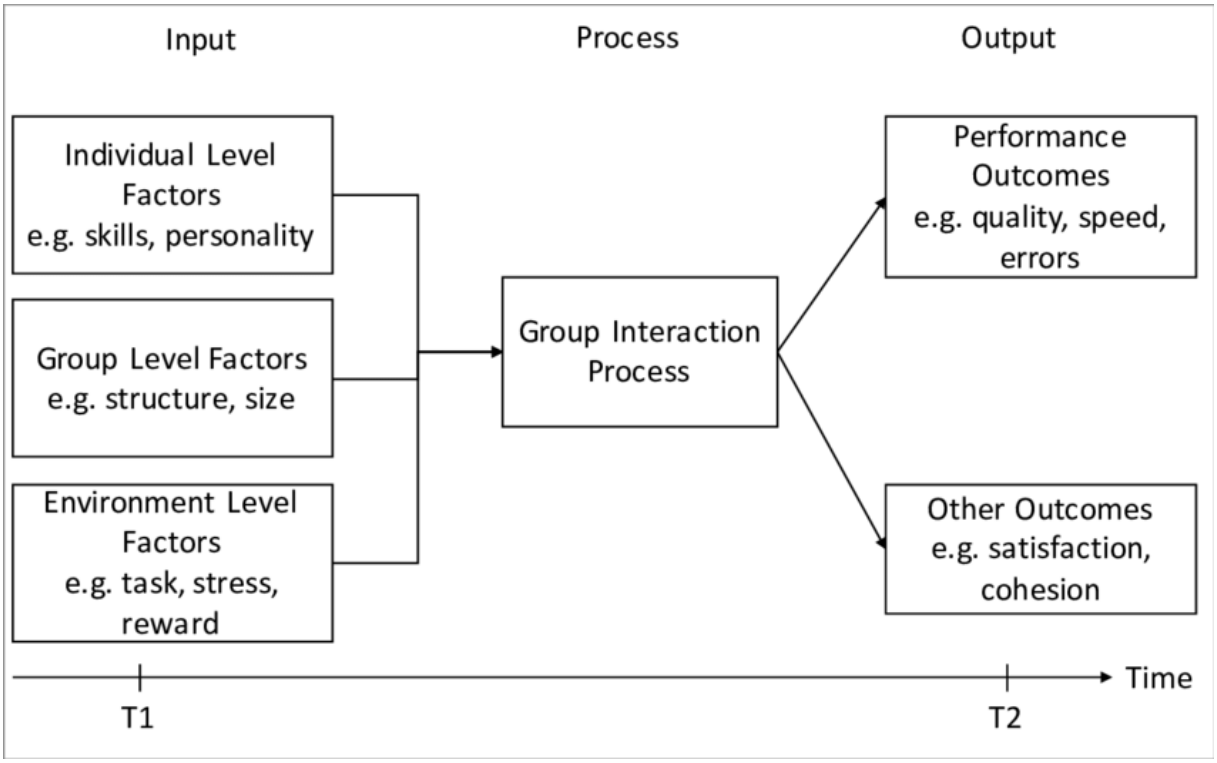
Moreover, Van der Hoek, Groeneveld, and Kuipers refer to performance as to "what is actually accomplished" (van der Hoek, Groeneveld, & Kuipers, 2018, p. 474).

Furthermore, Savelsbergh and colleagues argue that team performance can be measured by focusing on team characteristics or behaviours within the team (Savelsbergh, Van der Heijden, & Poell, 2009, p. 579).

However, Kozlowski and Klein argue that the dimension of interest for team performance, the nature of the team's work-flow interdependence, and the organizational context in which the team operates, among other things, are influencing factors of team performance (Kozlowski & Klein, 2000, p. 12). Thus, the conceptualization used for this thesis is that team performance can be explained as a function of individual and dyadic inputs, which can be measured by testing changing team characteristics and processes as well as behaviours within a team.

A useful framework for understanding team performance, called the input-process-output (I-P-O) framework, was drafted by Hackman in 1987. Figure 1 visualizes the framework (Hackman, 1987).

Figure 1: The I-P-O framework for analysing group behaviour and performance (Hackman, 1987, p. 316).



Both input and output variables are divided into three categories by the framework: those that portray individual team members, those that portray the group as a whole, and those that portray the group's operating environment. One of the framework's main assumptions is that input states influence group outputs through member interaction (Hackman, 1987, p. 317).

However, Ilgen and co-authors argue that the framework poses limitations to analysing team behaviour and performance in two ways. First, many theories offered by academics attempting to invoke the I-P-O model as process are actually emergent cognitive or affective conditions rather than a process. Second, they argue that recent research has moved past the simple notion of a single, direct path with steps from inputs to outcomes (Ilgen, Hollenbeck, Johnson, & Jundt, 2005, p. 520).

Nevertheless, Salas, Cooke, and Rosen explain why the I-P-O framework is relevant in order to understand team performance. They build on Kozlowski and Klein's conceptualization of team performance by adding that team performance is part of the I-P-O framework when

explaining the dynamic nature of teams (Salas, Cooke, & Rosen, 2008, p. 541). Despite the criticism, they draw our attention to the model by saying that team performance is actually a process and not a product, as teamwork is nested within team performance through a set of interconnected cognitions, attitudes, and behaviours that contribute to the dynamic processes that performance is (Salas, Cooke, & Rosen, 2008, p. 541). Besides explaining their view of the I-P-O model in light of team performance, Salas and colleagues explain the relationship between team performance and *team effectiveness*. According to them, team effectiveness is an assessment of the results of team performance processes in terms of a set of criteria (Salas, Cooke, & Rosen, 2008, p. 541). In other words, team effectiveness differs from team performance, as team effectiveness is concerned with the outcomes of team performance. An important takeaway is that in the public sector, team performance concerns a lot of different values, such as effectiveness. Team performance can be measured by using many types of instruments.

For this study, we consider team performance to be an output in this framework, as leadership is positioned as input and psychological safety as process in this thesis.

2.2. Shared leadership

Traditionally, *leadership* is seen as an individual focus based on vertical approaches to organizing work tasks within an organization or team. In order to develop a clear conceptualization of shared leadership, Yukl's (1989) definition of leadership is utilized as: "influence processes involving determination of the group's or organization's objectives, motivating task behaviour in pursuit of these objectives, and influencing group maintenance and culture (Yukl, 1989, p. 5)". Recently, more intensive research has been conducted on a *shared* approach of leadership, practicing leadership on a group level.

Some preliminary work on shared leadership was conducted by Gibb (1954), who was the first to suggest two kinds of team leadership: distributed leadership and focused leadership. Focused leadership occurs when leadership is resided within one individual, whereas distributed leadership resides within two or more individuals. Those multiple individuals then share roles, responsibilities and tasks which leadership brings (Gibb, 1954).

Additionally, Pearce and Conger's definition of shared leadership somewhat overlaps with Gibb's description. According to them, shared leadership entails the development of numerous leaders at the same time, with the purpose of enabling essential talent to emerge in accordance with the job requirements at hand, hence facilitating the achievement of broad common goals (Pearce & Conger, 2003).

Moreover, when looking back at the I-P-O framework, it can be concluded that shared leadership is an input in this model. Also, for this thesis, Pearce's conceptualization of shared leadership will be used.

In 2007, Carson and co-workers defined shared leadership slightly different than both Gibb and Pearce, as they refer to shared leadership as "a team property whereby leadership is distributed among team members rather than focused on a single destinated leader" (Carson, Tesluk, & Marrone, 2007, p. 1217). Thus, it contradicts the notion of a vertical leadership, where emphasis lays on the role of a manager which is hierarchically above a team. Moreover, Carson et al. provide useful information on the distinction between shared leadership and closely related concepts. For example, shared leadership is not to be confused with emergent leadership, as emergent leadership refers to group members exercising influence over other members of their group although no formal authority has been assigned to them (Scheiner & Goktepe, 1983). Furthermore, shared leadership is not to be confused with team processes such as cooperation and/or helping team members, as these do not involve the active influence that shared leadership entails (Carson, Tesluk, & Marrone, 2007, p. 1221).

Moreover, Carson and colleagues propose a theory that shared leadership is facilitated by an environment that consists of three dimensions: shared purpose, social support, and voice (Carson, Tesluk, & Marrone, 2007, p. 1222). The first dimension, shared purpose, exists when all team members agree on the ultimate collective goal and how they are going to work towards it. When a team unanimously agrees on objectives, it is easier to establish a collective goal and undertake action that support other team members' activities in order to reach those objectives. The second dimension, social support, entails a team environment in which team members provide emotional and psychological strength to each other, thereby creating an environment in which team members feel valued, appreciated, and supported. The last dimension, voice, is described by Carson as connoting participation and output. As a result, the presence of a higher degree of voice in a team should foster shared leadership by inspiring members to be committed to and actively participate in helping the team achieve its goals and objectives, as well as constructively challenging one another in pursuit of collective goals (Carson, Tesluk, & Marrone, 2007, p. 1222). Thus, Carson mentions shared leadership being important to help a team achieve its goals and objectives. Therefore, that notion of Carson is utilized in this conceptualization of shared leadership.

Furthermore, Käufer and Fletcher present another model, consisting of three shifts as a result of shared leadership's underlying paradigm change from traditional to distributed

leadership. The three shifts are presented in Table 1 below (Fletcher & Käufer, 2003, pp. 22-24).

Table 1: Three shifts of Fletcher

SHIFT	EXPLANATION
I – Distributed and Interdependent	Figureheads are needed hierarchically above a team. However, those figureheads are supported by the distribution of leadership practices throughout the organization. Moreover, leadership is interdependent in nature, focusing on collective achievement and shared responsibilities.
II – Embedded in Social Interaction	Shared leadership is a social process, described as a "dynamic, multidirectional, collective activity", since leadership occurs in and through relationships.
III – Leadership as Learning	Shared leadership fosters collective learning. Mutual learning, greater shared understanding and positive action are outcomes of shared leadership.

These shifts are needed to move from a traditional leadership to a distributed/shared leadership. The first shift illustrates that an actual figurehead is needed, apart from the existing shared leadership in a team. The second shift shows that social interaction is an antecedent to establish shared leadership. The third and last shift demonstrates that shared leadership fosters collective learning as an outcome of shared leadership. These shifts somehow connect with Pearce and Conger's conceptualization of shared leadership. Apart from the fact that Fletcher does argue for one specific figurehead while Pearce specifies multiple leaders, she agrees with Pearce by saying that shifts in behaviour are needed in order to realize shared leadership.

Moreover, Yukl's (2002) taxonomy of leadership behaviour consists of four categories. One of them will be used for the conceptualization of shared leadership in this thesis, namely the relations behaviour. According to Yukl, "the primary objectives of relations behaviour include strong commitment to the unit and its mission, and a high level of mutual trust and cooperation among members". In other words, the relations behaviour is all about commitment to the mutual goal, while fostering trust and cooperation within a team. The relations behaviour consists of five different kinds of behaviour, namely supporting, developing, recognizing, consulting, and empowering (Yukl, 2002, p. 19). For this study, supporting relations behaviours

are analysed. As explained by Yukl, "Supporting is defined as showing consideration, acceptance, and concern for the needs and feelings of other people (…) Supportive leadership helps to build and maintain effective interpersonal relationships" (Yukl, 2002, p. 20). Supporting behaviours are chosen to be analysed in this study, as shared leadership requires essential talent to emerge, for which a supportive environment is beneficial.

After reviewing the literature and following the I-P-O framework, it becomes clear that there is a relationship between shared leadership and team performance. The reason that shared leadership, and in particular the supporting behaviours, will contribute to team performance according to this conceptualization, is because shared leadership helps a team to achieve mutual goals and/or objectives.

Thus, it can be concluded that shared leadership and team performance are positively related through the relations/supporting behaviour, which fosters effective relationships while being in the pursuit of achieving common goals. Therefore, we can hypothesize the following:

H_1: There is a positive relationship between shared leadership and team performance in Dutch welfare teams.

2.3. Psychological safety

Lastly, linking psychological safety to the I-P-O framework, psychological safety can be identified as a process, which completes the framework for this study. The literature review on psychological safety below will explain why it is conceptualized as a process.

There is a vast amount of literature on psychological safety. However, reviewing psychological safety literature always starts with noting Amy Edmondson. As stated in the introduction, Edmondson names a team psychologically safe, when the team in question is safe for interpersonal risk-taking. In her ground-breaking article *Psychological Safety and Learning Behavior in Work Teams*, Edmondson (1999, p. 354) states that psychological safety is usually taken for granted within a team, even though sometimes psychological safety is being explicitly discussed. The essence of psychological safety is not within making it explicit, it is about mutual respect amongst team members, ensuring a sense of confidence that a team will not embarrass, punish or reject a member for speaking up (Edmondson, 1999, p. 354). Thus, psychological safety is not solely about the level of cohesiveness within a team; it requires additional measures than cohesiveness and making the concept explicit to all team members. What also becomes clear from Edmondson's study, is that psychological safety functions as a

cushion for failure in teams, promoting positive risk-taking and in turn creating creative performance. Hence, Edmondson's conceptualization of psychological safety will be utilized in this thesis. Also, it became clear from Edmondson's conceptualization that psychological safety is a group interaction, and can therefore be identified as a process in the I-P-O model
Amy Edmondson and Per Hugander presented a model on the Harvard Business Review website, which contains four steps to enhance psychological safety at the workplace (Edmondson & Hugander, 2021):

1. *Focus on performance* – encourage the team to share stories about how candor and vulnerability actually assisted effective outcomes.
2. *Train individuals and teams* – individual skills and team practice combined makes a winning team. Individuals must learn how to look from different perspectives and team practices hold those perspectives together.
3. *Incorporate visualization* – "by envisioning and writing down specific, tangible descriptions, people are better able to internalize new skills and practices"
4. *Normalize work vulnerability* – practicing minor acts of vulnerability decreases anxiety.

These four steps are important to understand for the conceptualization of psychological safety in this research, as the level of psychological safety can be tested through these steps.

Even though Edmondson is a pioneer in the field of psychological safety, more authors have made a contribution to psychological safety literature. Some preliminary work was carried out in the early 1990s by Kahn, who described and illustrated three kinds of psychological conditions: meaningfulness, safety and, availability (Kahn, 1990, p. 694). His definition of psychological safety overlaps with Edmondson's, as Kahn states that is a "sense of being able to show and employ self without fear of negative consequences to self-image, status, or career" (Kahn, 1990, p. 705). Kahn mentions that a psychological safe *climate* is needed for workers to feel connected to their work role. He also portrays a theory, in which he argues that there are four aspects that can influence psychological safety, namely interpersonal relationships, group and intergroup dynamics, management style, and process, and organizational norms (Kahn, 1990, p. 705). These points of influence strengthen Edmondson's argument that cohesiveness and explicitness do not carry enough influence to ensure a psychological safe climate within a team; there are many influences on psychological safety, therefore a psychological safe climate does not stem from one simple measure.

The association between shared leadership and psychological safety has been researched in numerous studies. Edmondson, for example, argues that the absence of psychological safety in a team might have negative consequences for team performance (Edmondson, 1999). Based on the literature, it can be concluded that psychological safety is likely to control the relationship between shared leadership and team performance. It is expected that the relationship between those variables takes place through psychological safety, since the level of team performance is predicted to be decreased when psychological safety is absent. Therefore, we can hypothesize the following:

H_2: Psychological safety mediates the relationship between shared leadership and team performance in Dutch welfare teams.

3. Methodology

This research sought to answer the question how psychological safety mediates the relationship between shared leadership and team performance in Dutch welfare teams. In this chapter, the methods for collecting and analysing data are explained and justified. First, the research design is given, in which the case selection background and the experiment are explained, and the empirical setting is discussed. Second, the methods of data collection will be explained. Third, the operationalization of variables is given: it is explained how the variables are measured. Fourth, the analysis strategy will be elaborated on, and the control variables will be introduced. Lastly, the validity and reliability of the research will be argued.

3.1. Research design

The case studied in this research are Dutch social welfare teams. The choice to study Dutch welfare teams as a case in this thesis has multiple reasons. The first and foremost justification is that there was a shift in this governmental body, namely decentralization in 2015. Municipalities have been in charge of youth care, job and income, and long-term sick and elderly care since that year (Rijksoverheid, n.d.). This shift from government to municipal tasks has led to many municipalities forming social welfare teams (called “wijkteams” in Dutch), in which professionals with different expertise work together to support residents (Engbersen, Verweij, Buizer, de Vries, & van Arum, 2021). Because of this shift, there were changes in performance and functioning of those teams, which are interesting factors to observe from an academic perspective. Another reason to study Dutch welfare teams, is because it was expected that shared leadership and psychological safety are topics that are widely used and/or discussed in such welfare teams. Thus, testing these variables in such an environment seems logical, also since team performance can be measured in those teams as well.

The respondents were chosen through non-random sampling. As for the study itself, 87 teams participated in a survey. Those teams stem from five different organizations. The survey itself was based on the Likert-scale, giving the respondents the opportunity to specify the extent to which they strongly disagree (1) or strongly agree (5), and everything in between. Statements were formulated on the basis of five themes, namely teamwork, team learning, leadership, bureaucracy and individual experience.

Moreover, a survey has been filled in by supervisors of the teams. This survey is used to measure the level of team performance in a team. Both surveys are aggregated to teams, and consequently merged together.

3.2. Methods

The data is collected through a survey-method between October and December 2020. As stated before, the survey was modelled according to the Likert-scale, which very commonly used. According to Nemoto and Beglar, the use of a Likert-scale results in fairly large amounts of data, with low effort (Nemoto & Beglar, 2014, p. 1). Moreover, the same authors argue that "they can provide highly reliable person ability estimates" (Nemoto & Beglar, 2014, p. 2). As such, this method of surveying is highly effective. However, the survey does start with some general questions about the respondent himself/herself.

Also, the response rates of officials willing to participate through the survey are an important aspect of the summary statistics. Therefore, teams with at least 30% of the professionals participating in the survey are included in the analysis to guarantee that the responses collected provide a representative view of the team. This resulted in a total of 761 respondents included in the study. Both the survey conducted with supervisors and with team members are used in this thesis, in order to gain insights from various levels in the field. Since both the supervisor dataset and the team member dataset are used in this study, the data of the team members had to be aggregated and the datasets had to be merged. In order to do so, the respondents in the team member dataset were clustered by team name. This resulted in a total of 70 teams, after which the dataset is merged with the supervisor dataset. In this newly combined dataset, twelve teams were removed as they did not indicate answers to items concerning team performance, resulting in 58 teams to be included in this study. Thus, teams are the unit of analysis, not solely team members. Also, of the total number of respondents, 12% identified themselves as male and 88% as female, with a mean age of 42 years old.

The data is collected by posing statements, to which the participants can respond by indicating a number between 1 and 5. The statements were written and categorized in various themes and for each theme other sources are used for justification. The themes used for this study were performance and innovation, teamwork and leadership. First, the only theme stemming from the supervisor survey, the theme *performance and innovation* is used to measure the variable of team performance. Items in this theme were drafted based on Hood's (1991) article on public management. Second, the theme *teamwork* is used to measure psychological safety. The statements included in this theme are based on research done by Amy Edmondson (1999). Third, the theme *leadership* consists of the items concerning measurement of shared leadership, more specifically, the supporting-relations behaviour. The items are based on research done by Gary Yukl (2012).

3.3. Operationalization

The variables that will be operationalized are the dependent variable of team performance, the independent variable of shared leadership and the mediating variable of psychological safety. The data source for all the concepts are the same, namely the survey. Table 2 below demonstrates how these variables are operationalized. The definition is given, as retrieved from the previous chapter. Moreover, the indicators are specified, in order to explain how the variable is measured. Lastly, the reliability of the score is reported, which is calculated with the aggregated data.

Table 2: Operationalization of concepts

CONCEPT	DEFINITION	INDICATORS	CRONBACH'S ALPHA (α)
Team performance	Van der Hoek and colleagues refer to performance as Van der Hoek, Groeneveld and Kuipers refer to performance as to "what is actually accomplished" (van der Hoek, Groeneveld, & Kuipers, 2018, p. 474). "Team performance is a complex function of specific individual and dyadic-networked-contributions (…) Team performance emerges from the behaviours of individual team members" (Kozlowski & Klein, 2000, pp. 11-12).	Measured through the supervisor survey, see Appendix I	α = .834
Shared leadership	Shared leadership entails the development of numerous leaders at the same time, with the purpose of enabling essential talent to emerge in accordance with the job requirements at hand, hence facilitating the achievement of broad common goals (Pearce & Conger, 2003). Here, shared leadership is measured by looking at the supporting	Measured through the team member survey, see Appendix I	α = .817

	behaviours specifically. "Supporting is defined as showing consideration, acceptance, and concern for the needs and feelings of other people (…) Supportive leadership helps to build and maintain effective interpersonal relationships" (Yukl, 2002, p. 20).		
Psychological safety (measured through the team member survey)	"Team psychological safety is defined as a shared belief that the team is safe for interpersonal risk taking. For the most part, this belief tends to be tacit-taken for granted and not given direct attention either by individuals or by the team as a whole" (Edmondson, 1999, p. 354).	Measured through the team member survey, see Appendix 1	α = .823

As the degree of team performance is rather difficult to assess, the dependent variable was measured through statements on team characteristics from the supervisor survey. For example, the following statement is used in order to measure team performance: "the team acts the same in similar cases". Appendix I can be consulted for more example statements; the measuring instrument consists of nine items. The reliability of the scale is α = .834.

Furthermore, the independent variable of shared leadership was measured through three statements in the team member survey, by following the relations behaviour. More specifically, the supporting behaviour is taken into account, by measuring, for example, the following statement: "the members of my team jointly plan the tasks of the team". Further examples can be found in Appendix I; the measuring instrument consists of three items. The reliability of the scale is α = .817.

Moreover, the mediating variable of psychological safety was also measured through employee perceptions in the team member survey. An example statement includes: "in my team, you can bring up problems or difficult issues". More examples of statements can be found in Appendix I; the measuring instrument consists of four items. The reliability of the scale is α = .823.

3.3.1. Control variables

In order to avoid possible confounders, two control variables are added to the models to ensure validity, namely team age and team experience. Team age is chosen as control variable, as Hoch, Pearce, and Welzel argue that age diversity moderates the relationship between shared leadership and team performance, thus having a certain effect on the relationship which should be controlled for (Hoch, Pearce, & Welzel, 2010). Moreover, team experience is chosen as control variable, since experience in the field might influence the relationships between variables as well and should be controlled for.

Team age is measured by asking the respondent's age in the beginning of the survey. Furthermore, team experience is measured by asking the respondent how many years they have been working as professional in the field. The mean age and experience of every aggregated team is used to run analyses including the control variables.

3.4. Analysis strategy

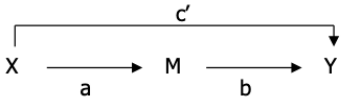
It should also be justified how the data derived from the dataset will be utilized and how the hypotheses will be tested. In order to test the hypotheses, the aggregated team scores will be analysed through SPSS. By doing so, it can be tested whether the relationship between shared leadership and team performance is positive, and it can be tested whether psychological safety plays a mediating role between those variables.

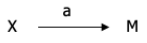
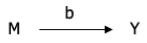
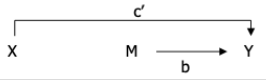
Before going into depth about the results, it is important to explain the steps taken to produce a correlation matrix. First, team responses to all nine items concerning team performance were added up and divided by the total number of variables, resulting in the mean instrument of that series. The same has been done for supporting behaviours of shared leadership and psychological safety. After that, the three instruments were analysed in a bivariate correlation, which resulted in a correlation matrix with a standard Pearson coefficient. Good states that "the standard Pearson Correlation test provides exact significance levels regardless of the distributions from which the data are drawn" (Good, 2009, p. 1). Moreover, Pearson's correlation coefficient fits this study better than, for example, Spearman, as Pearson is used when all variables that are being studied are normally distributed.

Furthermore, to test our first hypothesis, the relation between the independent variable "shared leadership" and the dependent variable "team performance" is tested through a multiple regression analysis.

Then, in order to test our second hypothesis, another model is used. For this thesis, Baron and Kenny's method for mediation is used. This method suited the mediating relationship that is tested in this research. Baron and Kenny's four-step method for testing mediation is tested through three regressions, namely the independent variable predicting the dependent variable, the independent variable predicting the mediator, and the independent variable and mediator collectively predicting the dependent variable. Lastly, the effect of the independent variable on the dependent variable controlling for the mediator should be zero for full mediation (Kenny, 2021). Their strategy is a popular one, as the steps are simple to follow. Moreover, by using this method, both partial and full mediation can be recognized. Figure 2 visualizes the steps in the Kenny and Baron method.

***Figure 2: Baron and Kenny's mediation method* (Newsom, 2020, p. 1)**



	Analysis	*Visual Depiction*
Step 1	Conduct a simple regression analysis with X predicting Y to test for path *c* alone, $Y = B_0 + B_1X + e$	
Step 2	Conduct a simple regression analysis with X predicting M to test for path *a*, $M = B_0 + B_1X + e$.	
Step 3	Conduct a simple regression analysis with M predicting Y to test the significance of path *b* alone, $Y = B_0 + B_1M + e$.	
Step 4	Conduct a multiple regression analysis with X and M predicting Y, $Y = B_0 + B_1X + B_2M + e$	

C stands for the total effect between X and Y. C' stands for the path between X and Y (the direct effect), A is the path between X and M, and lastly, B is the path between M and Y. Complete mediation occurs when path C=0. Partial mediation occurs when path C>C'. However, the relationship between path C and C' should be significant in order to be able to conclude that there is a partial mediation.

3.5. Validity and reliability

As argued by Leung, validity and reliability are both important criteria for quantitative and qualitative research (Leung, 2015). In this part, the reliability and validity will be discussed, and potential problems will be addressed. First, a general definition of validity will be given,

after which both *internal* and *external* validity will be discussed. Second, the concept of reliability will be explained.

Validity is, as stated by Heale and Twycross, "defined as the extent to which a concept is accurately measured in a quantitative study" (Heale & Twycross, 2015, p. 1). In other words, validity signifies the accuracy of a measure. As stated by Neuman, "when we say that an indicator is valid, it is valid for a particular purpose and definition" (Neuman, 2014, p. 215). In this case, we discuss two types of validity: internal and external validity. According to Patino and Carvalho Ferreira, internal validity refers to the degree to which the observed results are really in the population researched and are not due to methodological errors (Patino & Carvalho Ferreira, 2018). Thus, one could test internal validity by asking how well the study is conducted. The level of internal validity is enhanced in this research by avoiding possible confounding variables and using the same measurement instrument for all respondents. However, this study does not have a random selection and randomization. Additionally, the control variables are quite limited in scope; age and experience are rather general variables to control for. Therefore, the internal validity of this thesis might be threatened. Moreover, according to the same authors, external validity refers to whether the results will be the same with similar respondents in a different setting (Patino & Carvalho Ferreira, 2018). So, to ensure external validity, one could ask how applicable the results are to the real world, with for example different respondents, at a different time, in a different place. In order to ensure external validity in this study, sampling bias is avoided by choosing a sample that is representative to the population. The population to which the results of this study are generalized are municipal organizations. Moreover, the results can be generalized, since the sample size is quite large in this field. 761 respondents are included, aggregated in 58 teams, stemming from various public organisations in The Netherlands. Thus, the external validity of this study is, to some extent, ensured.

Reliability is defined by Heale and Twycross as "the accuracy of an instrument" (Heale & Twycross, 2015, p. 1). Moreover, Neuman adds to this definition by adding the word "consistency" to the measure (Neuman, 2014, p. 212). Thus, to put it differently, reliability is the degree to which a particular instrument consistently produces the same results when employed in the same situation over and over again. Neuman explains that there are three types of reliability: stability reliability, representative reliability and equivalence reliability (Neuman, 2014, p. 212). First, in order to assess stability reliability, one could ask whether the measure delivers the same answers when applied in different time periods. Since these surveys are only conducted in one time period, stability cannot be ensured. Second, representative reliability can

be assessed by asking whether the indicator delivers the same answer when applied to different groups, which can be tested with a subpopulation analysis (Neuman, 2014, p. 212). This type of reliability cannot be ensured in this research, since a subpopulation analysis has not been conducted. Third, equivalence reliability can be assessed by asking whether the measure produces consistent results across various indicators (Neuman, 2014, p. 213). In this research, multiple indicators are used to measure a specific construct. For example, four statements are used to measure psychological safety, instead of only one. This ensures equivalence reliability in this research. Moreover, by having a Cronbach's Alpha of more than .8 on every variable, the level of reliability of this study is heightened.

Thus, when considering the validity and reliability of this research, some strengths and weaknesses can be identified. The internal validity is a serious weakness of this study. Threats to internal validity are tried to be reverted by adding two control variables. However, the scope of those control variables remains limited. Additionally, the absence of random selection and randomization could pose a threat to the internal validity of this study. The external validity of this study is ensured by avoiding sampling bias during distribution of both surveys and can be considered a strength of this study.

Moreover, equivalence reliability is a strength of this research design, since for most variables, more than one statement has been specified in the surveys. Unfortunately, representative reliability cannot be ensured and is a possible weakness of this design. Lastly, the high level of Cronbach's Alpha does ensure a greater degree of reliability.

As stated before, this thesis' results are based on analyses from both the supervisor and the team member dataset. The use of more than one dataset has some methodological implications. An advantage of including more than one dataset is that perceptions from both sides can be included, namely from the team member level and the supervisor level. This also means that the sample size is increased, which helps enhancing the level of generalizability of the study. This could strengthen the conclusions made from this study. However, a negative side-effect of using two datasets, is that it costs a lot of time, and a mistake is made more easily when merging the datasets.

4. Results

This chapter of the thesis will focus on testing the actual results. It will be structured as following. First, the summary statistics, or descriptive statistics, will be given. This subchapter focusses on descriptive statistics such as age and years of experience, as well as on the variables of interest. Second, in order to show the correlation coefficients between the variables, a correlation matrix will be presented. Third, the hypotheses are tested**.** Lastly, a small summary of the results is given.

4.1. Descriptive statistics

The first table (Table 3) visualizes the descriptive statistics of team age and team experience. Most teams had a mean age of between 41 and 45 years old when the survey was conducted, accounting for about 31% of all teams. Since the distribution is quite even, no particularities stand out, apart from the fact that no team has a mean age of lower than 31 years. Besides, it is relevant to consider how long the respondents have worked as a professional in the field of welfare teams. Therefore, Table 3 also illustrates how many years the respondents have worked in the field. The mean of most of the teams have worked in the field for eleven to sixteen years at the time the survey was conducted, accounting for 41% of the total teams. It is also interesting to notice that 17% of the teams have worked in the field for six to ten years. This explains why the mean age of the teams is not lower than 31-35 years, as most teams have 11-15 years of experience already. Lastly, having well-distributed answers to these items benefits this research in terms of its generalizability.

Table 3: Team age & years of experience

Team age		N	%
	31-35	6	10.3
	36-40	14	24.1
	41-45	18	31.0
	46-50	9	15.5
	Total	47	81.0
Missing	System	11	19.0
Total		58	100.0

Team years of experience			
	5 or less	1	1.7
	6-10	10	17.2
	11-15	24	41.4
	16-20	10	17.2
	20 or more	3	5.2
	Total	48	82.8
Missing	System	10	17.2

Moreover, Table 4 below visualizes the descriptive statistics for our variables of interest, including the control variables. The mean scores for all main variables exceeded the theoretical average of 3.0 on a 5-point scale. This indicates that teams mostly find that their team has high levels of psychological safety and shared leadership, and that supervisors consider their level of team performance relatively high. What also stands out, is that no teams scale these main variables lower than three out of five, and that some teams even perceive their level of shared leadership and psychological safety as high as possible (five out of five).

Moreover, the mean age in the teams is 41 years old and the mean team experience is 13.5 years.

Table 4: Descriptive statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Team performance	58	3.00	4.89	4.14	.47
Shared leadership	58	3.44	5.00	4.44	.30
Psychological safety	58	3.73	5.00	4.31	.26
Team age	58	30.6	50.3	41.00	4.49
Team experience	58	1.5	22.1	13.5	4

To conclude the section on descriptive statistics, team age and team experience provides us insights in the demographics of our sample set, which includes the rather diverse characteristics of the respondents. Regarding the main variables, a key takeaway is that teams perceive the level of team performance, shared leadership and psychological safety rather high. These statistics give us a general idea about the perception of the measured variables before going into depth about their possible relationships.

4.2. Correlation matrix

Now, since the variable statistics are given, it is useful to see how those variables actually relate to each other. In order to show the correlation coefficients between our three main variables of team performance, shared leadership and psychological safety, and our two control variables of team age and team experience, a correlation matrix is needed. In such a matrix, it becomes visible how two variables are correlated. Thus, it offers a small summary of relationships. Below, Table 5 presents a correlation matrix.

Table 5: Correlation table main- and control variables (N=58)

Variables	1	2	3	4	5
1. Team performance	1				
2. Shared leadership	.093	1			
3. Psychological safety	.179	.667**	1		
4. Team age	-.211	-.176	-.227	1	
5. Team experience	.041	-.017	-.052	.744**	1

** Correlation is significant from the 0.01 level (2-tailed).

The correlation between shared leadership and psychological safety is rather high (.667). However, the relationship between psychological safety and team performance, and shared leadership and team performance, does not seem significant. The relationship between psychological safety and team performance is lower than expected, with a relatively low positive correlation (.179).

Regarding the control variables, it is visible that almost all main and control variables are negatively correlated, except for team performance and team experience (.041). No statistical significance has been identified, apart from the correlation between team age and team experience (.744). Noteworthy are the negative correlations between team age and all of

the main variables. Thus, as the team ages, the rank for team performance, shared leadership and psychological safety becomes lower. On the contrary, the positive correlation between team performance and team experience illustrates that the longer the teams work in the field, the higher they would rank team performance.

What can be interpretated from this matrix, is that it is interesting for these variables to examine the strength and direction of the relationships instead of solely the statistical significance, as most correlations are not statistically significant. Also because of this study's limited N (58) it is wise to pay extra attention to the strength and direction of relationships. Noteworthy is the fact that almost all main variables have a negative direction towards the control variables. The only positive, strong relationships can be identified between shared leadership and psychological safety, and between team age and team experience.

Moreover, team performance is the only variable that is measured from the supervisor dataset. Not one variable has a strong relationship with team performance. The only strong relationships are relationships that are drawn from the team member dataset. Thus, in this case, the source of data can determine the strength of the correlations.

In conclusion, the key takeaways from the correlation matrix are the strong positive relationship between shared leadership and psychological safety and between team age and team experience. Noteworthy is also the unexpected weak positive relationship between psychological safety and team performance, as well as the weak relationships between the main variables and control variables, in a mostly negative direction.

4.3. Hypothesis testing

In this section, our two hypotheses are tested. The hypotheses were formulated as follows: *H_1: "There is a positive relationship between shared leadership and team performance in Dutch welfare teams"* and *H_2: "Psychological safety mediates the relationship between shared leadership and team performance in Dutch welfare teams".*

The four models that are being tested for the hypotheses are the following: Model 1 contains the control variables as independent variables and team performance as dependent variable. Model 2 contains the independent variable of shared leadership, as well as the control variables and the dependent variable of team performance. In Model 3, the dependent variable is replaced by psychological safety. Lastly, Model 4 shows the effects of shared leadership and psychological safety together with the control variables as independent variables on team

performance as dependent variable. Models 1 and 2 are used to test the first hypothesis, and models 2, 3 and 4 are used to test the second hypothesis. Table 6 visualises all models, reporting unstandardised and standardised coefficients.

Table 6: Regression Analyses (N = 58)

	B (95% confidence intervals)	*SE*	β	*t*	*p*
Model 1					
Constant	5.76	.62		9.26	<.001
Team age	-.06	.02	-.54	-2.87	.006*
Team experience	.05	.02	.44	2.35	.022*
Dependent variable: Team performance, R^2 Adjusted = .100					
Model 2					
Constant	5.71	1.24		4.62	<.001
Shared leadership	.01	.21	.01	.05	.962
Team age	-.06	.02	-.54	-2.75	.008*
Team experience	.05	.02	.44	2.29	.026*
Dependent variable: Team performance, R^2 Adjusted = .084					
Model 3					
Constant	2.20	.54		4.06	<.001
Shared leadership	.56	.09	.64	6.18	<.001
Team age	-.01	.01	-.19	-1.22	.228
Team experience	.01	.01	.10	.65	.519
Dependent variable: Psychological safety, R^2 Adjusted = .432					
Model 4					
Constant	5.15	1.42		3.63	<.001
Shared leadership	-.13	.27	-.08	-.49	.626
Psychological safety	.26	.31	.14	.82	.417
Team age	-.05	.02	-.51	-2.57	.013*
Team experience	.05	.02	.43	2.20	.032*
Dependent variable: Team performance, R^2 Adjusted = .078					

Note: Asterisks are statistically significant when P < .05.

As can be derived from Table 6, team age and team experience are statistically significant (p = .006 and p = .022 respectively). Team age has a slightly negative impact on team

performance (B = -.06). The explained variance of this model is relatively low, namely 10% ($R^2_{adjusted}$ = .100).

Model 2 offers a new insight, as it becomes clear that shared leadership is not statistically significant with team performance as dependent variable (p = .962). Team age and team experience remain significant (p = .008 and p = .026 respectively). This indicates that shared leadership does not impact team performance. Moreover, by adding shared leadership to the model, the adjusted R^2 value even decreases to 8.4% ($R^2_{adjusted}$ = .084). Thus, the hypothesised relationship cannot be supported with this data and is therefore rejected.

Moving on, the Baron and Kenny Method will be utilized to test the second hypothesis, for which Model 2, 3 and 4 are needed. Baron and Kenny's four-step method for testing mediation is tested through three regressions. To repeat, Model 2 explains X_1 (shared leadership), X_2 (team age) and X_3 (team experience) predicting Y (team performance). Model 3 explains X_1, X_2 and X_3 predicting M (psychological safety). Lastly, Model 4 explains X_1, X_2, X_3 + M predicting Y. Thus, Model 2, 3 and 4 together represent the first three steps of Baron and Kenny's method. Lastly, the fourth step: the effect of X on Y controlling for M (direct effect = C') should be zero for full mediation. If the first three steps are met, and the fourth is not zero, then partial mediation could nevertheless be present in the relationship. Table 7 sets *B* apart from the other coefficients, to provide a clear picture of the coefficient used in the method.

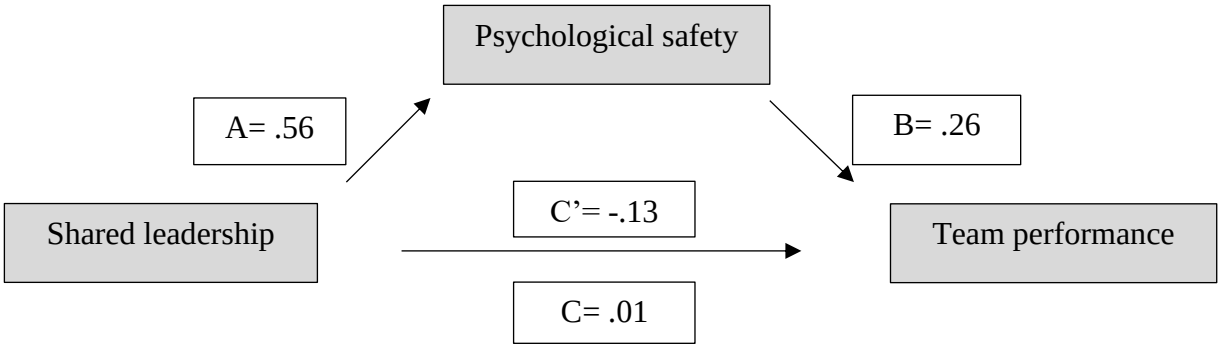
Table 7: Unstandardized Coefficient B per main variable

Model		Unstandardized Coefficient *B*
2	Shared leadership	.01
	Team age	-.06
	Team experience	.05
	Dependent Variable: Team performance	
3	Shared leadership	.56
	Team age	-.01
	Team experience	.01
	Dependent Variable: Psychological safety	
4	Shared leadership	-.13
	Psychological safety	.26

Team age	-.05
Team experience	.05
Dependent Variable: Team performance	

As visualized in Table 7, Model 2 shows that shared leadership barely influences team performance as *B* = .01. Furthermore, Model 3 demonstrates that shared leadership does moderately effect psychological safety when psychological safety is positioned as dependent variable, as *B* = .56. Then, Model 4 shows that shared leadership and psychological safety together as independent variables have a minimal effect on team performance. When controlling psychological safety, shared leadership has a negligible negative effect on team performance, as *B* = -.13. Psychological safety has a minimal positive effect on team performance, since *B* = .26. After completing the first three steps, the paths can be visualized, as can be seen in Figure 3. It is important to note that the variables of team age and team experience were controlled while implementing this method.

Figure 3: Baron and Kenny Method with psychological safety controlled



According to Baron and Kenny, when psychological safety completely mediates the relationship between shared leadership and team performance, path C' should be zero (Kenny & Baron, 2021). Figure 3 illustrates that path C' is not zero, thus psychological safety does not fully mediate the relationship between shared leadership and team performance. We find that the independent variable still has effect on the dependent variable when the mediating variable is controlled, therefore, complete mediation is ruled out. Nonetheless, since the first three steps were met, partial mediation could be indicated, as shared leadership has reduced effect on team performance when psychological safety is controlled for. However, since path C (*B* = .01) is statistically insignificant, the relationship is barely present.

Thus, the second hypothesis is also rejected; the partial mediation that psychological safety offers to the relationship between shared leadership and team performance is insignificant.

To conclude the results section, we can state that both hypotheses are rejected. The positive relationship between shared leadership and team performance is so minimal, that it should be disregarded and therefore rejected. Moreover, after following the Baron and Kenny Method for mediation, it became clear that psychological safety's mediating role is so negligible, that it truly cannot be labelled a mediator.

What does stand out are the overall significant statistics of the control variables team age and team experience, when combined with shared leadership as independent variables and team performance as dependent variable.

5. Discussion, conclusion and limitations

This chapter will focus on analysing and discussing the results in order to conclude with an answer to the research question of this thesis, which reads as follows: "*How does psychological safety mediate the potentially positive relationship between shared leadership and team performance in Dutch welfare teams?"*. Moreover, limitations are specified and recommendations for future research are given.

5.1. Discussion

As became clear earlier, the correlations between our three main variables were positive. However, only one relationship was significant: the relationship between shared leadership and psychological safety. Moreover, no significant relationships were found between our main variables and our control variables team age and team experience. Research had found that age might have an effect on the relationship between shared leadership and team performance (Hoch, Pearce, & Welzel, 2010). However, no significant effect was found in this research for team age, as well as team experience. Noteworthy is that team experience and team performance did correlate positively, although minimal. Also, there is little research done about experience having an effect on the relationship between shared leadership and team performance, which makes explaining the relationship between those variables particularly difficult.

The weak positive relationship between shared leadership and team performance led to a rejection of the first hypothesis. The control variables did not have a significant effect on either the independent or dependent variable. Therefore, the relationship cannot be explained through team age or team experience.

Edmondson (1999) argued that the absence of psychological safety in a team might have negative consequences on team performance, which indicates that psychological safety might mediate the effect of a predicting variable on team performance. This study held shared leadership as predictor of team performance. More specifically, the supportive behaviour of shared leadership was analysed in this study. The results from the Baron and Kenny Method illustrates that the second hypothesis is also rejected, because of statistical insignificance. Multiple regressions were employed to generate that result through the Baron and Kenny Method.

However, in the chapter on methodology, we decided that statistical significance is not the only important factor when analysing the results. Non-existing relationships and insignificant results are important and interesting findings as well. When looking at the strength and direction of the three paths, it can be concluded that the relationships among themselves move towards a positive direction, despite the fact that they are not very strong. We find that shared leadership leans towards a positive direction when explaining team performance, albeit not strongly. Shared leadership and psychological safety are even strongly and positively correlated with each other, which raises new questions for future research. Thus, even though the results are not as expected, they can be implicated for analysing the strength and direction between the variables and are valid results as well.

Moreover, during the multiple regression analyses, it became clear that team age and team experience are statistically significant when predicting team performance, which is a good implication for future research as well.

Lastly, using two datasets, from the team member as well as the supervisor perspective, proved itself useful. By using the perceptions of both groups, a strong test between supervisor-perceived team performance and team member-perceived shared leadership and psychological safety could be conducted.

5.2. Conclusion

This research aimed to test the relationship between shared leadership and team performance in Dutch welfare teams, and to examine the possible mediating role of psychological safety. In sum, by employing the answers to these hypotheses, the research question "*How does psychological safety mediate the potentially positive relationship between shared leadership and team performance in Dutch welfare teams?"* was answered through survey analysis. Team members as well as supervisors presented answers to different statements measuring their perceptions of their own team. This study indicates that the statistically weak and slightly negative relationship between supporting-oriented shared leadership (X) and team performance (Y) in Dutch welfare teams is not mediated by psychological safety (M). Moreover, as psychological safety does not play the role of a mediating protagonist, it cannot be explained how it mediates the relationship. Lastly, this study has modestly contributed to existing research of welfare teams. Its unexpected findings do provide implications for further research on leadership and performance in the public sector.

5.3. Limitations and future research

Reflecting on this study, some limitations can be stated. First, the internal validity is a weakness of this research. Even though control variables have been added in this research, internal validity cannot be ensured, as the control variables were limited in scope. Future researchers are encouraged to test other kinds of control variables with a greater scope.

Second, another limitation of this research is the fact that it is not longitudinal research; the survey is only conducted in one place during one time period. If this study was based on a longitudinal approach, the progress and changes over a certain time period could have been examined, which would have improved stability reliability as well. In sum, for future research, it is highly recommended to use a longitudinal approach in order to make stronger and more reliable assumptions.

Third, even though it is very useful having two datasets from two different perspectives, having to aggregate and merge two datasets can lead to the removal of viable data. In this case, some aggregated teams had to be removed from the dataset as their supervisor did not present answers to the statements of team performance, which resulted in labelling that team as unusable. Future researchers could opt for analysing individuals instead of aggregated teams to avoid this limitation.

Bibliography

Bureau for the Senior Civil Service. (2017, April). *Vision of Public Leadership: The Netherlands Senior Civil Service.* The Hague: Bureau for the Senior Civil Service, Ministry of the Interior and Kingdom Relations.

Carson, J., Tesluk, P., & Marrone, J. (2007). SHARED LEADERSHIP IN TEAMS: AN INVESTIGATION OF ANTECEDENT CONDITIONS AND PERFORMANCE. *Academy of Management Journal* (Vol. 50, No. 5), 1217-1234.

Duhigg, C. (2016, February 25). *What Google Learned From Its Quest to Build the Perfect Team*. Retrieved September 2021, from The New York Times: https://www.nytimes.com/2016/02/28/magazine/what-google-learned-from-its-quest-to-build-the-perfect-team.html?smid=pl-share

Dyer, J. L. (1984). Team research and team training: A state-of-the-art review. *Human Factors*, 285-323.

Edmondson, A. (1999). Psychological Safety and Learning Behavior in Work Teams. *Administrative Science Quarterly,* (Vol. 44, Iss. 2), 350-383.

Edmondson, A. (2012). Teaming: How Organizations Learn, Innovate, and Compete in the Knowledge Economy. In A. C. Edmondson, *Teaming: How Organizations Learn, Innovate, and Compete in the Knowledge Economy.* Hoboken, NJ.

Edmondson, A., & Hugander, P. (2021, June 22). *4 Steps to Boost Psychological Safety at Your Workplace*. Retrieved October 2021, from Harvard Business Review: https://hbr.org/2021/06/4-steps-to-boost-psychological-safety-at-your-workplace

Engbersen, R., Verweij, S., Buizer, D., de Vries, S., & van Arum, S. (2021, March 3). *Wat belemmert en helpt wijkteams?* Retrieved November 2021, from Movisie: https://www.movisie.nl/publicatie/wat-belemmert-helpt-wijkteams

Fletcher, J. K., & Käufer, K. (2003). Shared Leadership. In C. L. Pearce, & J. A. Conger, *Shared Leadership : Reframing the Hows and Whys of Leadership.* Thousand Oaks, CA: SAGE Publications.

Gibb, C. (1954). Leadership. *Handbook of social psychology* (Vol. 2), 877-917.

Good, P. (2009). *Robustness of Pearson correlation.* Huntington Beach CA: Information Research.

Hackman, J. R. (1987). The design of work teams. *The Handbook of Organizational behavior*, 315-342.

Han, & et al. (2021). The effects of shared leadership on team performance. *Leadership & Organization Development Journal* (Vol. 42, Iss. 4), 593-605.

Heale, R., & Twycross, A. (2015, May 15). Validity and reliability in quantitative studies. *Evidence-Based Nursing Online First* (Vol. 0, Iss. 0), 1-1.

Hoch, J. E., Pearce, C. L., & Welzel, L. (2010, January 1). Is the Most Effective Team Leadership Shared? *Journal of Personnel Psychology* (Vol. 9, Iss. 3), 105-116.

Hood, C. (1991, March). A PUBLIC MANAGEMENT FOR ALL SEASONS? *Public Administration* (Vol. 69, Iss. 1), 3-19.

Ilgen, D. R., Hollenbeck, J. R., Johnson, M., & Jundt, J. (2005). TEAMS IN ORGANIZATIONS: From Input-Process-Output Models to IMOI Models. *Annual Review Psychology* (56), 517-543.

Kahn, W. (1990). Psychological conditions of personal engagement and disengagement at work. *Academy of Management Journal* (Vol. 33, Iss. 4), 692-724.

Kenny, D. A. (2021, May 4). *Mediation*. Retrieved November 2021, from David A. Kenny: http://davidakenny.net/cm/mediate.htm

Kenny, D., & Baron, R. (2021, May 4). *MEDIATION*. Retrieved December 2021, from David A Kenny: http://davidakenny.net/cm/mediate.htm

Kim, S., Lee, H., & Connorton, T. P. (2020, July 24). How Psychological Safety Affects Team Performance: Mediating Role of Efficacy and Learning Behavior. *Frontiers in Psychology*(Vol. 11, Art. 1581).

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

Leung, L. (2015). Validity, reliability, and generalizability in qualitative research . *J Family Med Prim Care.* (Vol. 4, Iss. 3), 324–327.

Nemoto, T., & Beglar, D. (2014). Developing Likert-scale questionnaires. *N. Sonda & A. Krause (Eds.)* (pp. 1-8). Tokyo: JALT.

Neuman, W. L. (2014). Social Research Methods: Qualitative and Quantitative Approaches. In W. L. Neuman, *Social Research Methods: Qualitative and Quantitative Approaches* (Vol. 7). Essex, United Kingdom: Pearson.

Newsom. (2020). Testing Mediation with Regression Analysis. *Structural Equation Modeling*.

Parker, H., & du Plooy, E. (2021, March). Team-based games: Catalysts for developing psychological safety, learning and performance. *Journal of Business Research*, (Vol. 125), 45-51.

Patino, C. M., & Carvalho Ferreira, J. (2018). Internal and external validity: can you apply research study results to your patients? *J Bras Pneumol* (Vol. 44, Iss. 3), 183.

Pearce, C., & Conger, J. (2003). Shared leadership reframing the hows and whys of leadership. In C. Pearce, & J. Conger, *Shared leadership reframing the hows and whys of leadership*. Thousend Oaks, CA: SAGE.

Rijksoverheid. (n.d.). *Decentralisatie van overheidstaken naar gemeenten* . Retrieved November 2021, from Rijksoverheid: https://www.rijksoverheid.nl/onderwerpen/gemeenten/decentralisatie-van-overheidstaken-naar-gemeenten

Salas, & et al. (2008). On Teams, Teamwork, and Team Performance: Discoveries and Developments. *Human Factors: The Journal of the Human Factors and Ergonomics Society* (Vol. 50, Iss. 3), 540–547.

Salas, E., Cooke, N. J., & Rosen, M. A. (2008, June). On Teams, Teamwork, and Team Performance: Discoveries and Developments. *Human Factors* (Vol. 50, Iss. 3), 540-547.

Savelsbergh, C., Van der Heijden, B., & Poell, R. (2009). The Development and Empirical Validation of a Multidimensional Measurement Instrument for Team Learning Behaviors. *Small Group Research* (Vol. 40, Iss. 5), 578-607.

Scheiner, C. E., & Goktepe, J. R. (1983). Issues in emergent leadership: The contingency model of leadership, leader sex, leader behavior. *Small groups and social interactions* (Vol. 1), 413-421.

The Government of Western Australia. (n.d.). *Psychologically Safe and Inclusive Workplaces*. Retrieved December 2021, from The Government of Western Australia:

https://www.wa.gov.au/government/multi-step-guides/psychologically-safe-and-inclusive-workplaces

van der Hoek, M., Groeneveld, S., & Kuipers, B. (2018, December 19). Goal Setting in Teams: Goal Clarity and Team Performance in the Public Sector. *Review of Public Personnel Administration* (Vol. 38, Iss. 4), 472-493.

Yukl, G. (1989). Leadership in organizations. In G. Yukl, *Leadership in organizations* (Vol. 2nd edition). Englewood Cliffs, NJ.

Yukl, G. (2002). A Hierarchical Taxonomy of Leadership Behavior: Integrating a Half Century of Behavior Research. *Journal of Leadership & Organizational Studies* (Vol. 9, Iss. 1), 15-31.

Yukl, G. (2012, December 27). Effective Leadership Behavior: What We Know and What Questions Need More Attention. *Academy of Management Perspectives* (Vol. 26, Iss. 4), 66-85.

Appendix I

Code	Operationalization in Dutch	Concept
PR_RECHT_1	handelt bij vergelijkbare casussen hetzelfde.	Team performance
PR_RECHT_2	gaat altijd rechtmatig te werk.	Team performance
PR_RECHT_3	communiceert open en transparant.	Team performance
PR_EFF_1	levert waar voor haar geld.	Team performance
PR_EFF_2	opereert kostenbewust.	Team performance
PR_EFF_3	gaat efficiënt te werk.	Team performance
PR_RESP_1	speelt adequaat in op veranderde omstandigheden.	Team performance
PR_RESP_2	reageert serieus op suggesties voor verbetering.	Team performance
PR_RESP_3	blijft haar werk goed doen in moeilijke omstandigheden.	Team performance
GL_0_1_LG	heeft aandacht voor de behoeftes van individuele teamleden.	Shared leadership
GL_0_2_LG	is betrokken met de teamleden.	Shared leadership
GL_0_3_LG	ondersteunt teamleden indien nodig bij een moeilijke taak .	Shared leadership
PV_1	kan je problemen of lastige kwesties naar voren brengen.	Psychological safety
PV_2	is het gemakkelijk om anderen om hulp te vragen.	Psychological safety
PV_3	is een vergissing maken geoorloofd.	Psychological safety

PV_4	worden ieders unieke vaardigheden en talenten gewaardeerd.	Psychological safety