
Exploring Demographic Diversity in Low-Skilled Jobs: The Effect of Surface-Level Dormant Faultlines on Team Performance and Social Integration

September 2018

Natalia Popova

First supervisor: Dr. Ir. J. de Leede

Second supervisor: MSc, PhD Candidate M. Renkema

ABSTRACT We build on diversity literature by theorizing that teams with strong surface-level faultlines (based on age, gender, nationality or origin) are more likely to be broken into subgroups. To examine this, we conducted research in low-skilled teams as less creative job allows neutralizing the effect of deep-level diversity. Using data collected from 300 employees nested in 32 teams from a Dutch cleaning company, we do not find that dormant faultlines' strength has an effect either on faultlines' activation or on team outcomes. Nevertheless, our results demonstrate that if the team is split into subgroups, it negatively affects its performance and social integration with conflicts as a mediator of this relationship. These findings contribute to the diversity literature by providing strong evidence that it seems impossible to predict team performance by focusing exclusively on teams' demographic composition. Therefore, we call for more field research of subgrouping process in diverse teams. The main implication for diversity management is that subgroups in low-skilled teams worsen its work outcomes, thus, the creation of team unity and cohesion is a key to better performance. It requires skilled line managers who see teams on daily basis and should be able to monitor team dynamics and prevent subgrouping based on demographic or other differences.

Keywords: diversity, faultlines, subgroups, performance, team functioning, conflict

Master of Business Administration

Specialization: Human Resource Management

University of Twente

Master Thesis

Title	Exploring Demographic Diversity in Low-Skilled Jobs: The Effect of Surface-Level Dormant Faultlines on Team Performance and Social Integration
Author	Natalia Popova
Student number	S1890239
Current address	De Grutto 98 7905BD Hoogeveen The Netherlands
E-mail	Natalyp.perm@gmail.com
University	University of Twente P.O Box 217 7500 AE Enschede The Netherlands
Faculty	Behavioural, Management and Social sciences
Master programme	Business Administration
Track/Specialisation	Human Resource Management
First Supervisor	Dr. Ir. J. de Leede j.deleede@utwente.nl +31646012830 University of Twente Faculty of Behavioural, Management and Social Sciences
Second Supervisor	MSc, PhD Candidate M. Renkema m.renkema@utwente.nl +31534894489 University of Twente Faculty of Behavioural, Management and Social Sciences
Organization of Research	Asito B.V. Plesmanweg 24 7602 PE Almelo The Netherlands
Research defence	26-09-2018

PREFACE

This research is the final project of the master in Business Administration at the University of Twente. Past several months were intense and this study is the biggest and most ambitious project of my life. Even though there were many challenges in the process of doing research and writing this report, it gave me an incredible wealth of knowledge and experience that ultimately was worth all that hard work.

I would like to thank my first supervisor Dr. Ir. Jan de Leede for organizing HR Analytics course where I was inspired by the field of diversity and for giving me the opportunity to conduct the research on this topic. I am grateful that you visited the Research Honours events, gave me profound feedback and provided so much help during all the way to my graduation. Without your support I definitely could not complete my master thesis. I would like to thank my second supervisor MSc Maarten Renkema for helping me enormously to organize the data and providing clear recommendations on the better ways of conducting quantitative analysis and reporting my findings. I truly appreciate that you have spent so many hours to help me and gave feedback that allowed me to improve the study very much.

I must express my biggest gratitude to Melissa Veltien, Laura de Wit and Hans van Leeuwen from Asito. My data collection succeeded only through your support. Thank you that you have believed in my idea and helped me to find 32 teams to establish contacts with all team managers and to draw up a team visiting schedule. Moreover, I am overwhelmingly grateful that you adjusted the survey for low-skilled employees and spent so much energy and time on it. Asito is a wonderful company and I am happy that I had the opportunity to acquire some practical knowledge and to meet many diverse people from more than 40 countries.

Big thanks to Dr. Anne Dijkstra, Dr. Petra de Weerd-Nederhof and Molly Waite who have organized Research Honours Programme and, luckily for me, have selected me to take part in it. This unique experience helped me to develop many research skills and improved my scientific writing. Your feedback has helped me to gain new insight and improve this paper.

My special thanks go to my relatives and friends who supported and believed in me during the whole period of my master education. I really appreciate the patience of my Dutch friends who have practised Dutch with me and have helped to advance it during two years. If I was not able to speak Dutch, the execution of this research would be impossible.

This research is not only my success but the combination of efforts of all aforementioned people and unconditional support that I had from all of them. I had a great time at the University of Twente and thank you all! You have made this time unforgettable!

Natalia Popova

September 2018

TABLE OF CONTENTS

INTRODUCTION.....	4
THEORETICAL FOUNDATIONS OF RESEARCH ON TEAM DIVERSITY	6
Faultlines Concept	8
Dormant and Activated Faultlines	11
Surface-level Faultlines in Low-skilled Labour Context	11
HYPOTHESES.....	13
Proposed Effect of Dormant Faultlines' Strength on Team Outcomes.....	13
The Dormant Faultlines' Strength and Faultlines' Activation.....	14
Conflicts as a mediator of the relationship between Activated Faultlines and Team Outcomes.....	15
DATA AND METHODS.....	18
Sampling and Respondents.....	18
Procedure	18
Measures	20
Aggregation.....	22
RESULTS	23
Group-level analysis	23
Dormant Faultlines' Strength and Activated Faultlines	25
Activated Faultlines, Conflicts and Team Outcomes	27
Individual-level analysis	31
DISCUSSION.....	35
Theoretical Contributions.....	35
Practical Implications.....	37
Limitations and Future Research.....	40
CONCLUSION	41
REFERENCES	42
Appendix A.....	48
Appendix B	50
Appendix C.....	51
Appendix D.....	52

INTRODUCTION

The relentless globalization of today's society has not only given a rise to the growth of multinational corporations and has created severe competition but also has changed the nature of countries' demographics. People immigrate to other countries for many possible reasons that, eventually, results in a big change in the labour market. The recognition that the workforce has become increasingly diverse has provided a stimulus for research on the consequences that workforce heterogeneity brings to the organization (Jackson & Joshi, 2011; Rico, Molleman, Sánchez-Manzanares, & Van der Vegt, 2007).

Despite a rich history of research on team diversity, results have been very inconsistent. While some studies claim that more diversity is better for team performance (Gibson & Vermeulen, 2003; Lau & Murnighan, 2005), others find strong support that teams with no diversity outperform those with high diversity (Jackson & Joshi, 2011; Li & Hambrick, 2005; Van Knippenberg & Schippers, 2007). One big shortcoming of the diversity literature was recognized by Lau and Murnighan (1998), who proposed that taking all the demographic attributes in isolation from each other does not represent the diversity in a team correctly. In order to address the structure of diversity Lau and Murnighan developed and introduced new concept called faultlines. Group faultlines result from the alignment of multiple characteristics such as nationality, age or gender. For instance, if a hypothetical team has six members, where three are Polish men older than 40 and the other three are German women younger than 30, the surface-level faultline would be very strong as the team can easily split into two separate subgroups. Faultline theorists argue that this separation can lead to conflicts, low group cohesion and team members' satisfaction and, eventually, to lower team performance (Bezrukova, Thatcher, Jehn, & Spell, 2012; Lau & Murnighan, 2005; Molleman, 2005).

One reason for inconsistent findings in past research may be the lack of understanding that the impact of surface-level diversity faultlines depends on task and context characteristics (Molleman, 2005; Rico et al., 2007). While the faultlines concept was developed in the diversity literature, Harrison, Price, and Bell (1998) drew researchers' attention to the fact that the impact of demographic diversity may be significantly weakened with the length of time that team members work together while the impact of deep-level diversity would be strengthened. Similarly, Molleman (2005) proposed that if work task becomes more creative and complicated, after a short period of time team members may start to ignore their demographic differences and concentrate on their educational or functional background dissimilarity. In other words, deep-level diversity would excel the effect of surface-level diversity in teams performing more complex tasks.

Surprisingly, research on diversity and later on faultlines has mostly ignored teams performing low-skilled jobs and a vast majority of the studies were focused on R&D or top management teams (Joshi & Roh, 2009; Thatcher & Patel, 2012; Webber & Donahue, 2001). However, dynamics of R&D teams is unlikely to depend on demographic faultline while employees' functional and educational

background faultlines would affect their group work. For instance, the exchange of ideas among people who have similar experience and knowledge is more straightforward and, thus, an information-based subgroups may be formed (Jehn, Northcraft, & Neale, 1999; Pelled, Eisenhardt, & Xin, 1999). In contrast, if the task is easier and does not require extensive ideas exchange, employees would be more concentrated on their surface-level diversity because they may not have time or interest in exploring their less visible differences.

We conclude that, in order to find an effect of faultlines that are based on easily observable demographic attributes such as age, gender, race, or nationality, the research should be conducted in low-skilled labour context where the influence of deep-level diversity attributes can be mostly neutralized. Moreover, low-skilled work floors are facing one of the highest levels of surface-level diversity as in a new country, immigrants tend to begin with low-skilled work due to the lack of language skills or certification. It should be pointed out that we do not mean no-skilled jobs as employees in low-skilled jobs are expected “to act appropriately at work and to perceive cues from others correctly” (Maxwell, 2006, p. 2). Often low-skilled jobs have requirements for basic language proficiency, communication skills, mechanical skills and sometimes even educational credentials or professional certifications. Therefore, in low-skilled settings employees have enough task interdependence to form a team that is defined as “a collection of individuals who are interdependent in their tasks, who share responsibility for outcomes, who see themselves and who are seen by others as an intact social entity” (Cohen & Bailey, 1997, p. 241).

The main purpose of this study is to deepen our understanding of diversity by examining whether surface-level faultlines in the group performing low-skilled tasks affect its group process and performance. In addition, following the literature on faultlines as a guide, we take the activation of the faultlines into account, i.e. if team members do actually perceive that their demographic differences caused the subgroup formation (Chrobot-Mason, Ruderman, Weber, & Ernst, 2009; Jehn & Bezrukova, 2010). Foreshadowing our logic, we hypothesize that low-skilled teams with stronger surface-level faultlines would more likely to have subgroups, which would lead to conflicts that, consequently, would worsen the team performance and social integration. Thus, we will try to find the answer to the following research question: *To what extent do surface-level dormant faultlines influence the low-skilled teams’ outcomes?* In order to answer this question, such group processes as *faultlines’ activation, conflicts, teams’ social integration* and *performance* will be examined.

We test our hypotheses based on data derived from questionnaires, filled out by 300 team members, resulting in a unique sample of 32 cleaning teams. Collecting the data from individuals allows us to measure to what extent employees indicate that their teams are divided into subgroups, report having conflicts, are satisfied and attached to their colleagues (i.e. extent of their social integration) and indicate whether their team performs well. Subsequently, the data obtained from individuals can be aggregated to the team-level. In addition, data collected from the 31 managers of

the same teams allows us to compare results for the influence of surface-level dormant faultlines on teams' process and outcomes.

Together, our theory and empirical evidence offer several contributions. First, our study contributes by improving the understanding of the association between diversity and outcomes in teams. Although Webber and Donahue (2001, p. 158) claimed that "researchers should stop treating diversity as a generic concept by suggesting that all types of diversity will have a positive outcome on group performance", still the operationalization of the diversity traits is inconsistent in the literature (Stanciu, 2015). In this study, we were able to collect extensive demographic information for every team that allows us to calculate the faultlines' strength based on all possible easily-observable attributes (age, gender, nationality, country of origin) and demonstrate that making the choice between nationality or country of origin already leads to different diversity and faultlines' values for the same team. As a consequence it explains why current findings in the diversity and faultlines' literature may be inconsistent as authors use different combinations of the demographic attributes.

Second, our research contributes to the faultlines literature by responding to numerous calls for studies conducted in the real teams and in low-skilled context (Joshi & Roh, 2009; Thatcher & Patel, 2012). Usually, researchers have created artificial work groups consisting of students that allowed to control the faultlines' strength (Chrobot-Mason et al., 2009). As mentioned above, the negative effect of demographic differences should be more salient in low-skilled context because employees would more likely choose to collaborate with similar others as such jobs do not stimulate employees to explore more deep-level diversity. Therefore, we propose that if no relationship between demographic composition and team outcomes would be found for the low-skilled team, then it is unlikely to find the relationship between surface-level faultlines' strength and performance or innovativeness in R&D or top management teams while this type of diversity studies is prevalent at this moment.

Thirdly, we challenge the idea that based on information about the demographic composition of the team, managers are able to predict its performance. We suggest that faultlines' activation has to be taken into account as if the team is not actually divided into smaller cliques then we will not be able to find conflicts or low social integration there. Finally, our study provides guidelines for managers on whether it is possible to enhance diversity in their organization without having threats to team performance and process (e.g. conflicts). We detail the contributions in the discussion section.

THEORETICAL FOUNDATIONS OF RESEARCH ON TEAM DIVERSITY

During the last three decades the domain of diversity has been getting more attention from both researchers and organizations as the workforce is continuing to become more and more heterogeneous (Jackson & Joshi, 2011; Jackson, Joshi, & Erhardt, 2003; Van Knippenberg & Schippers, 2007). Increased globalization, higher workforce mobility, changes in national demographics, the rise of multinational enterprises in combination with a sense of moral obligation for

diversity in organizations resulted in the greatest multiplicity of ages, races, nationalities on the work floor than ever before (Buengeler & den Hartog, 2015; Cunningham, 2009). Obviously, this tendency was recognized by researchers and studies on how differences in a team affect group performance and group members' attitudes are in a great request.

Van Knippenberg and Schippers (2007) give the following definition of diversity: "differences between individuals on any attribute that may lead to the perception that another person is different from self" (p. 517). A number of researchers focused on readily observable demographic attributes such as age, gender, race and nationality - this type of diversity was called demographic or surface-level diversity (Lau & Murnighan, 1998; Van Knippenberg & Schippers, 2007). Another type of diversity is less easily discoverable as it includes differences in educational and functional backgrounds or different attitudes, personality traits, values and other social identities. It is called diversity on underlying attributes, information-related or functional diversity.

Studies of diversity are typically focusing on the team-level and are trying to assess if different degrees of diversity affect group outcomes (Jackson & Joshi, 2011). Several theoretical perspectives were guiding diversity research: the social-categorization perspective (Turner, 1982), the information/decision-making perspective (Williams & O'Reilly, 1998), social identity perspective (Tajfel & Turner, 1986), attraction-selection-attrition model (Schneider, 1987) and organization demography perspective (Pfeffer, 1985). Those perspectives have provided many useful insights into the effect of diversity on group performance, however, the empirical results were inconsistent and evidence for both positive and negative effect of diversity can be found (Thatcher & Patel, 2012).

For example, the social-categorization perspective asserts that individuals prefer to work with similar others and would distinguish between in-group members (who are similar to them) and out-group members (who are different). Eventually, the work team breaks into two or more subgroups that may have a detrimental effect on group outcomes as conflicts and competition between the subgroups could take place. Therefore, a group would have better performance outcomes if it is less diverse as the possibility to break into subgroups would be lower. On the other hand, information/decision-making perspective posits that differences in knowledge and skills may help a group to develop more creative ideas, to have more different perspectives and, thus, be more innovative (Jackson & Joshi, 2011; Van Knippenberg & Schippers, 2007).

All the aforementioned perspectives were mostly focused on heterogeneity – "the extent to which differences or inequalities exist among categories of organizational members" (Alexander, Nuchols, Bloom, & Lee, 1995, p. 1458). Blau (1977) developed one of the first heterogeneity measures and assumed that if for any nominal parameter (e.g. race) the number of groups is large and the population is evenly divided among those groups then the diversity in a given population is high. For example, if the nominal parameter is nationality and the number of nationalities in a given group is large, then that team is more diverse than the team with two nationalities only.

However, the dispersion measure does not allow examining multiple demographic attributes simultaneously (Lau & Murnighan, 1998). Let us suppose that we measure the group diversity on the basis of nationality and gender. Firstly, we need to separately calculate the heterogeneity indexes for nationality and for gender and, secondly, make the weighted average of two results. But the final number would not give us the full information about the combination of the nationalities and genders in the group. Thus, the shortcoming of the diversity literature was the assumption that all the demographic attributes are isolated from each other. Bezrukova et al. (2012) provide an example of gender disregard so that the experience of African American men would be erroneously equated to the experience of African American woman in the same group. Thereby, in 1998, Lau and Murnighan introduced the concept of faultlines that included also the configuration of team members' characteristics.

Faultlines Concept

Lau and Murnighan (1998) coined the term faultlines and defined it as “hypothetical dividing lines that may split a group into subgroups based on one or more attributes” (p. 328). The formation of faultlines and its effect on team outcomes were explained by the theoretical mechanisms of self-categorization and social identification perspectives that explicate why individuals classify themselves and others based on salient attributes such as race or gender, and similarity-attraction perspective that explains why individuals prefer to work with similar others that results in subgroup formation (Lau & Murnighan, 1998; Thatcher & Patel, 2012).

According to self-categorization theory (Turner, 1982) and social identification theory (Tajfel & Turner, 1986), people differentiate themselves from others on the basis of observable differences and place themselves and similar others into in-groups and dissimilar others into out-groups. The result of such categorization process may be that less diverse groups function more smoothly (Van Knippenberg & Schippers, 2007). Similarity-attraction perspective (Williams & O'Reilly, 1998) complements the social-categorization perspective and explains that people tend to like and trust in-group members more than out-group members that produces subgroups and rises problems in inter-group relations. Those theories are extended by categorization-elaboration model of Van Knippenberg, De Dreu, and Homan (2004).

The categorization-elaboration model explains how faultlines influence team process and outcomes. Three factors such as comparative fit, normative fit and cognitive accessibility affect the salience of social categorization. “The comparative fit refers to the extent to which the categorization provides a good reflection of similarities and differences between people” (Van Knippenberg et al., 2004, p. 1014). For example, if the team consists of many different nationalities, categorization into same-nationality subgroups would be uneasy and the categorization in this team would be less salient.

Lau and Murnighan (1998) suggest that if employees' differences converge (e.g. young German males vs. older Spanish females), faultline becomes stronger and social categorization is more likely to occur.

“Normative fit reflects the extent to which the categorization makes sense in relation to the individual's cognitive frame of reference like beliefs and stereotypes” (Van Knippenberg et al., 2004, p. 1014). This may depend on individual characteristics or on job settings. Van Knippenberg et al. (2004) give an example that truck driving seems to be more associated with stereotypic beliefs about gender differences than many other jobs. So the belief of members that differences are meaningful within their work context may affect the categorization process.

Cognitive accessibility of a categorization refers to “the readiness of the perceiver to use the categorization” (Van Knippenberg et al., 2004, p. 1014). Throughout their lives individuals learn that some categories like gender and age are often used for the formation of the social groups. Thus, such categories are more obvious to use for subgrouping than, for instance, the colour of people's car (Van Knippenberg et al., 2004). If some difference is strongly emphasised within the organization, it may make categorization on this attribute more accessible. For example, if study is conducted in the university settings and an experiment team is formed from the business administration and biomolecular engineering students, categorization based on the educational background will likely to be salient because universities often emphasize differences between faculties and educational programmes, so a category of task-relevant background can easily come to students' minds (Van Knippenberg & Van Ginkel, 2010). If the three aforementioned categories are present, the more social categorization can be expected. That also means that diversity in itself does not automatically lead to subgrouping.

As the cognitive accessibility of categorization based on differences in demographic characteristics may be more or less given, it does not attract much scholars' attention while the principle of comparative fit was studied most profoundly (Van Knippenberg & Van Ginkel, 2010). Consequently, the need to avoid studying dimensions of diversity in isolation or in additive models, but to consider the extent to which positions of the dimensions converge to form a faultline was recognized by diversity theorists (Van Knippenberg & Van Ginkel, 2010). Extensive evidence is obtained that it is not so much diversity per se that is problematic but rather combinations of diversity dimensions that result in strong faultlines (Thatcher & Patel, 2012; van Knippenberg, Dawson, West, & Homan, 2011).

Lau and Murnighan (1998) have constructed a set of four-person groups that vary in the nature of their demographic characteristics to demonstrate the difference between the concepts of diversity (heterogeneity) and faultlines (see Table I). For instance, in the second team the diversity is low while the faultline is very strong, meaning that it is very likely that members will form two subgroups (members A and B, and members C and D). If more demographic characteristics overlap, the chance is higher that team members would identify themselves with the subgroups than with their team as a

whole (Gibson & Vermeulen, 2003; Lau & Murnighan, 1998). So-called “us vs. them” categorization may result in in-group-favouring bias that impedes communication with the out-group (Pesch & Bouncken, 2018). Li and Hambrick (2005) argue that this way of thinking leads to a lower level or even the total loss of interaction and joint decision making. What is more, not only communication may suffer but the willingness of team members to contribute to the task as a whole may drop down (Thatcher, Jehn, & Zanutto, 2003). Barkema and Shvyrkov (2007) add that in strong faultline settings new initiatives that appear in one subgroup may never reach the entire team as propositions from the out-group might be quietly dismissed behind the scenes.

Table I. The composition of four-person group varying in Diversity and Faultlines’ strength across four Demographic Characteristics: Race, Sex, Age, and Occupation Roles

N	Member A	Member B	Member C	Member D	Diversity ^a	Faultline Strength
1	White Male 20 Sales	White Male 20 Sales	White Male 20 Sales	White Male 20 Sales	None	None
2	White Male 50 Plant Manager	White Male 50 Plant Manager	Black Female 31 Clerical staff	Black Female 35 Clerical staff	Low (4 diff.; 2 levels; 0.23)	Very strong (4 align; 1 way)
3	White Male 60 Plant Manager	Asian Female 30 HR manager	White Male 20 Plant Manager	Black Male 35 HR manager	Moderate (4 diff.; 2-3 levels; 0.29)	Strong (3 align; 2 ways)
4	Native American Female 20 Unskilled	White Male 30 Supervisor	Black Female 65 Executive	Asian Male 50 Machinist	Maximum (4 diff.; 2-4 levels; 0.45)	Very weak (1 align; 4 ways)

Adapted from “Demographic Diversity and Faultlines: The Compositional Dynamics of Organizational Groups,” by D. C. Lau and J. K. Murnighan, 1998, *Academy of Management Review*, 23(2), p. 330.

^aFor an explicit explanation of the calculation of Diversity and Faultlines’ values see Lau and Murnighan (1998, p. 330).

The main consequence of faultlines is their threat to team unity because different subgroups will have little in common as team members may cohere and share opinions more often within the subgroup than with others. That may lead to gossips, irritations and disputes between the subgroups (Gibson & Vermeulen, 2003). Mortensen and Haas (2018) explain that view of a member from one subgroup may differ from the view of a member from another subgroup given their distinct vantage points. Those views may be partial and inaccurate as members tend to interact primarily with their subgroup, so the understanding of the entire team may differ very much and lead to disagreements. It is commonly assumed that subgroups increase conflicts and worsen team performance (Barkema & Shvyrkov, 2007; Gibson & Vermeulen, 2003).

In sum, even though the faultline concept has some similar underlying theoretical foundations as the diversity literature, the differentiation between subgroups provides a better understanding of how

the group composition may influence group process and outcomes (Thatcher et al., 2003). Thatcher and Patel (2012) have reviewed 39 empirical articles that were investigating the effect of faultlines on group outcomes and concluded that: “The quantitative aggregation results show that faultline strength and faultline distance have significant effects on group outcomes above and beyond the effects of team diversity” (p.987). For the aforementioned reasons, the concept of faultlines will be used in this paper.

Dormant and Activated Faultlines

Dormant faultlines are objective demographic alignments across members that potentially may divide a group into subgroups, while active faultlines take place when employees actually perceive their team being divided into two or more subgroups (Jehn & Bezrukova, 2010). The analogy between team faultline and a geological fault was drawn by Lau and Murnighan (1998). The geological fault is a “crack in the earth's crust resulting from the displacement of one side with respect to the other” (“Geological fault,” 2003-2008). Long-dormant faultlines can be unobserved before they actually become active and trigger an earthquake. By parity of reasoning, the same can happen in the work team. Group faultlines can be unnoticed and invisible until some external forces bring the subgroup conflict to the workplace. It is probable that although the team has strong dormant faultlines, the subgroups may never actually appear and team members may perceive themselves as one group ignoring demographic differences.

Chrobot-Mason et al. (2009) studied events that activate the dormant faultlines and argued that in the team with strong faultlines the possibility of a small event to evoke identity threat, anxiety, tension or conflict is high. Authors called such occasions “triggers” and defined them as “events involving two or more people from different social identity groups that ignites a replication of societal-based identity threat in the organization” (Chrobot-Mason et al., 2009, p. 1770). After analyzing more than 130 events, authors defined five types of trigger: differential treatment, different values, assimilation, insult of humiliating action and simple contact. It was the first field study examining the activation of faultlines as all previous researches were conducted in the laboratory settings or with the artificially made groups consisting of students. Chrobot-Mason et al. (2009) have provided strong evidence in their work that demographic faultlines if being activated bring conflict and negatively impact group performance.

Surface-level Faultlines in Low-skilled Labour Context

While the advantages of diversity are often promoted and managers believe that more diversity will have more benefits than limitations for the workplace, researchers suggest treating diversity as a double-edged sword (Sawyer, Houlette, & Yeagley, 2006). The meta-analyses and literature reviews of empirical studies that aimed to analyze the effect of diversity on team outcomes conclude that findings are highly inconsistent (Jackson et al., 2003; Thatcher & Patel, 2012; Webber & Donahue, 2001).

Empirical studies are reporting a variety of effects of faultlines. Some studies are reporting the negative effect of faultlines (Homan et al., 2008; Sawyer et al., 2006; Thatcher et al., 2003), whereas others demonstrate beneficial outcomes (Bezrukova, Spell, & Perry, 2010; Lau & Murnighan, 2005).

Webber and Donahue (2001) have proposed an idea that the type of team may be a potential moderator variable between the type of diversity and team outcomes. They assumed that top management teams are likely to be more heterogeneous on job-related characteristics (e.g. functional and educational background), and are less likely to experience surface-level diversity (e.g. race, sex, age) while the production teams are expected to have the opposite situation. However, Webber and Donahue (2001) could not conduct the meta-analytic test as the vast majority of the research was exclusively focused on top management teams and ignored lower-level teams. Rico et al. (2007) have paid attention to the nature of work, in particular to team autonomy, and suggested that task characteristics are moderating the effect of faultlines on team outcomes as more difficult tasks create a climate appropriate for exploiting team members' differences. Similar idea can be found in the study of Molleman (2005) where he concluded that the role of surface-level diversity would be less important if group has more complex tasks and more autonomy, as the increased need for collaboration would force the team members to share their thoughts, analyze the tasks and reconsider the best working methods collectively. Furthermore, if tasks require the knowledge sharing, subgroups would depend on each other in order to achieve the common goal, and, accordingly, it will make the effect of faultlines insignificant (Molleman, 2005).

Joshi and Roh (2009), in their comprehensive meta-analytic review of diversity research, have noted the importance of industry settings. For example, they found that diversity in age, gender and race had positive effects on performance in service industries and slightly negative effects in manufacturing industries. Subsequently, it is indeed important to take the nature of the task and company setting into account while searching for the effect of faultlines on team outcomes. The confirmation of this statement can be found in the review of variables moderating the effect of diversity made by Guillaume, Dawson, Otake-Ebede, Woods, and West (2017) where it was identified that task characteristic matter as diversity can facilitate innovation and enhance performance on complex tasks but only when it is associated with task-relevant knowledge.

If tasks of the team require innovative thinking and knowledge sharing, the surface-level diversity may be less critical, as team members would pay more attention to their information diversity (e.g. educational and functional background). In order to see the surface-level diversity faultlines' effect, the research should be done in the settings of low-skilled teams, as they are less likely to experience information-based faultlines or investigate those differences during their work. The lack of research on low-skilled teams mentioned by Webber and Donahue (2001) was still found to be relevant eleven years later by Thatcher and Patel (2012) who claim that most of the studies that are made in the field settings are still focusing on R&D or management teams. Hence, we propose to conduct the research

in low-skilled labour setting to understand if demographic faultlines do actually have a negative effect on team outcomes.

HYPOTHESES

Proposed Effect of Dormant Faultlines' Strength on Team Outcomes

The contradictory results of the diversity studies can be explained by the nature of the faultlines concept as it combines two theoretical perspectives. On the one hand, the information/decision-making perspective contends that diversity enhances the elaboration of information, rises discussions and redounds to the integration of ideas (Van Knippenberg et al., 2004). Therefore, more diversity should facilitate better problem solving and have a positive effect on group performance. On the other hand, social-categorization perspective proposes that group may break into “we-they” subgroups that should bring tensions and conflicts on the work floor (Homan et al., 2008). From the categorization-elaboration model (Van Knippenberg et al., 2004) it is predicted that homogeneous teams tend to be better for the more simple tasks as thorough and creative information processing does not take place in routine settings. In this case, disruptive social categorization processes may arise and negatively affect team performance (Turner & Reynolds, 2011; van Knippenberg et al., 2011).

Lau and Murnighan (1998), in their faultline theory, made an assertion that the effect of diversity and faultlines on group outcomes tend to differ, as diversity is more likely to be a potential source of creativity and task conflict, while group faultlines may generate relationship conflict. This raises the question for teams having demographic diversity and working on management or innovative tasks of whether the effect of diversity that facilitates creativity or the effect of faultlines that bring relationship conflict would eventually have more impact on team outcomes. For the surface-level faultlines perspective, innovativeness and knowledge exchange within a team are less pertinent than for functional diversity perspective. We suppose that setting our sights on the low-skilled employees would help to avoid these conflicting effects of diversity faultlines on team outcomes.

From the categorization-elaboration model, differences that are more cognitively accessible (i.e., such as gender, age, race/ethnicity) are more likely to be used for stronger categorization (Van Knippenberg et al., 2004). The principle of normative fit suggests that social categorization is more likely to happen in jobs where stereotypes and prejudices are common. Comparative fit principle proposes that high within-category similarity and high between-category differences result in stronger faultline that undermines team cooperation (Van Knippenberg & Van Ginkel, 2010). As we conduct this research in cleaning teams performing job stereotyped to be predominantly female (Glenn, 1990), we expect all three components to interact and in case of strong dormant faultlines lead to disruptive inter-group biases. Therefore, strong dormant faultlines would likely negatively affect groups' functioning:

Hypothesis 1a (H1a): Strong surface-level Dormant Faultlines are negatively related to Team Performance.

Furthermore, we propose that dormant surface-level faultlines would also negatively affect the social integration of the team. Social integration is “the degree to which team members feel satisfied with and committed to team interactions and their willingness to stay in the team” (Rico et al., 2007, p. 115). While some studies suggest that strong team members’ bonds with their subgroups can result in high levels of subgroup satisfaction, that in turn would lead to positive effects on overall group satisfaction, the vast majority of studies have documented the negative relationship between the faultline strength and the levels of group satisfaction (Thatcher & Patel, 2012). As aforesaid, it is likely to witness the conflicts and disputes between subgroups in the teams with strong dormant faultlines. This, in turn, may lead to reduced levels of social integration in the team as members would constantly feel tension and disunity.

In contrast, groups with weak dormant faultlines are less likely to divide into subgroups, that makes team members more attached to each other and allows to develop a strong feeling of group belonging (Rico et al., 2007). Cronin, Bezrukova, Weingart, and Tinsley (2011) found that stronger subgrouping in a team led to lower satisfaction as group members were less effectively integrated. Thatcher et al. (2003) reported that groups that had two relatively homogeneous subgroups had low levels of satisfaction and commitment. Additionally, faultlines can decrease the team members’ motivation to contribute to the task that may reduce satisfaction and intention to remain of new employees who find themselves in an unmotivated group (Meyer, Schermuly, & Kauffeld, 2016). Scholmerich, Schermuly, and Deller (2016) postulated that a negative relationship exists between socio-demographic faultlines’ strength and perceived cohesion that was also supported by the results. Importantly, this relationship was not found for experience-based faultlines, thus, the negative impact of socio-demographic faultlines exceed that of experience-based faultlines. As we mentioned before, the low-skilled labour context helps to neutralize the influence of possible experience-based faultlines, consequently, we propose the following hypothesis:

Hypothesis 1b (H1b): Strong surface-level Dormant Faultlines are negatively related to Social Integration.

The Dormant Faultlines’ Strength and Faultlines’ Activation

Most of past studies on faultlines were focused on dormant faultlines (Lau & Murnighan, 2005; Li & Hambrick, 2005; Molleman, 2005; Rico et al., 2007) and did not examine if the potential faultlines were active. Zanutto, Bezrukova, and Jehn (2011) criticise this approach because considering whether members actually perceive their group being split into subgroups may help to gain an insight into the group dynamics. In other words, if individuals identify themselves with demographic subgroup the

influence of surface-level faultlines on group process and outcomes would be more explicit. For this reason, it is critical to examine if the activated demographic faultline exists so that employees behave as if they are members of the different subgroups (Zanutto et al., 2011).

Jehn and Bezrukova (2010) have empirically supported the fallacy of the assumption made by the majority of papers that dormant demographic faultlines lead to activated group faultlines. Similarly, the results of the study made by Spoelma and Ellis (2017) demonstrate the importance of considering whether the faultlines are activated, as the team-level threat manipulation check was only beneficial for teams with activated surface-level faultlines. Besides that, Schölmerich et al. (2016) highlight that surface-level diversity attributes are associated with widely shared stereotypic beliefs and faultlines based on these attributes are more likely to create intergroup bias compared to experience-based faultlines.

In the aforementioned trigger section, the mechanisms of faultlines' activation were outlined and five types of triggers were mentioned: simple contact, differential treatment, different values, assimilation, insult or humiliating action. We presume that analogous triggers can occur in the low-skilled teams. For example, the team is composed of members who represent the dominant culture and of those who do not. It could happen that the first ones would automatically surmise that the employees who have another background would endeavour to act just like them (i.e. assimilation trigger). Chrobot-Mason et al. (2009) point out: "Members of the dominant group tend not to like it when co-workers of other cultures express their distinctiveness. They feel threatened by expressions of a different identity and worry about losing their dominance".

In low-skilled teams, the dominant subgroup is unlikely to hold an interest in new knowledge or ideas from the non-dominant group but would rather prefer that everyone just accepts the common mode of doing work. As a result, the non-dominant subgroup would feel affronted by such inequity and the inability to have a say in the workplace matters. In this way, strong dormant faultline is likely to be activated and break a group into subgroups because of the assimilation trigger that occurred due to members' demographic differences. This reasoning leads to the following hypothesis:

Hypothesis 2 (H2): Strong surface-level Dormant Faultlines are positively related to Activated Faultlines in the team.

Conflicts as a mediator of the relationship between Activated Faultlines and Team Outcomes

The linkage between group composition and conflicts is one of the most examined in the faultlines and group research (Choi & Sy, 2010; Jehn, 1995; Pelled et al., 1999). Conflicts between subgroups have been widely recognized as an intermediate process that harms teams' effectiveness (Minichilli, Corbetta, & MacMillan, 2010; Thatcher & Patel, 2012). The attraction paradigm (Byrne, 1971) states that people tend to trust and like others similar to themselves and avoid and dislike dissimilar ones.

Moreover, surface-level differences like gender and age have been particularly found to rise mistrust and interpersonal incompatibilities in teams (Li & Hambrick, 2005; Pelled et al., 1999). Therefore, we propose that relationship and task conflicts play an important role in understanding how activated faultlines influence team outcomes.

Relationship conflicts can be defined as “disagreements and incompatibilities among group members about personal issues that are not task related, such as social events, gossip, and word news” (Thatcher et al., 2003, p. 222). The surface-level differences may be the source of the relationship conflict as members could be biased, have prejudices and negative stereotypes about certain demographic characteristics. The study on thirty baseball teams and the effect of group faultlines on team outcomes confirmed that the conflict among team members intensified and made more salient the splits within a group that had a negative effect on team outcomes (Bezrukova, Spell, Caldwell, & Burger, 2016).

As stated before, people tend to like and trust those who are similar to themselves (Byrne, 1971). For this reason, organizations tend to look for and hire those job applicants that are likely to “fit” their team, thereby making the workforce more homogeneous (Jackson & Joshi, 2011). Similarities shared by members can make the work process more smooth and result in fewer conflicts. However, as little knowledge or experience are needed for low-skilled employment and there is little variation in workers’ productivity, a very limited range of selection procedures is typically used (Carless et al., 2007). Psychological and other forms of assessments can be difficult for low-skilled workers as the minimum reading level may be too high that results in higher personality and cultural heterogeneity (Carless et al., 2007).

According to attraction paradigm (Byrne, 1971), if a new employee starts working in a diverse team, he or she will likely try to search for the similar colleagues, that is actually the beginning of the subgroup formation. At the same time, if a task is creative and group members have to exchange information, the new employee will eventually socialize with all the team members. In contrast, if a task can be done without the extensive deliberations, the chances are high that the group members would communicate only within their subgroups (Lau & Murnighan, 2005). Thus, the relationship conflicts in terms of gossips, doubts and misinterpretation are likely to appear.

Task conflict is “disagreements among group members’ ideas and opinions about the task being performed” (Thatcher et al., 2003, p. 222). If the task is routine, it can be expected that this type of conflict is not likely to occur. However, Pelled et al. (1999) found that people performing the routine tasks experienced task conflict as group members seek task debates with demographically dissimilar others to make their work more interesting. Even though researchers mostly link the task conflict to different educational or experience backgrounds (Molleman, 2005; Thatcher et al., 2003), support for demographic faultlines being the cause of task conflicts can be found in the literature (Thatcher & Patel, 2012). Ultimately, Pelled et al. (1999) conclude that both emotional and task conflicts tend to

accompany each other. In this manner, we expect that if the low-skilled team with activated faultlines has discussions about the task they should perform, conflicts between members of different subgroups who have different opinions would occur. Accordingly, we propose the following hypothesis:

Hypothesis 3 (H3): The stronger employees perceive Activated Faultlines, the more Conflicts in the team will be indicated.

In the previous section it was stated that strong dormant faultlines in the low-skilled team would lead to more conflicts, nevertheless, the effect on subsequent team performance can vary. A number of studies corroborated that conflicts mediate the relationship between team faultlines and performance (Alipour, Mohammed, & Raghuram, 2018; Choi & Sy, 2010; Pelled et al., 1999). Groups with strong faultlines would have problems in communication between subgroups that often leads to emotional conflicts and reduced performance (Lau & Murnighan, 2005). Bezrukova et al. (2016) defined three main task-irrelevant issues in the groups with active faultlines that distract the team from their main work: competition between subgroups, distrust and restricted communication between groups. Thus, the presence of subgroups harms groups performance by consuming time and resources that are directed towards those three issues (Bezrukova et al., 2016). Jehn (1995) found the empirical support that task conflict was consistently negatively related to performance in routine-task groups but had a positive effect on nonroutine-task groups.

Pelled et al. (1999), in their integrative study of the relationship among diversity, conflict and performance, discerned that diversity shapes conflict and that conflict, in turn, shapes performance. Li and Hambrick (2005) have empirically confirmed that group processes (e.g. conflicts) fully mediate the negative association between faultlines and later team performance. Results of the study of Jehn and Bezrukova (2010) demonstrate that conflict mediates the relationship between activated faultlines and group outcomes for both objective and self-rated team performance. Thus, if diversity leads to the formation of multiple “we-they” subgroups, communication may be disrupted and lead to misunderstanding and inability to converge different perspectives. In contrast, if subgroup’s members do not engage in distancing members of other categories, the team would be able to develop joint solutions and share information without having many conflicts and losses in performance (Gibson & Vermeulen, 2003). Therefore, we formalize this logic in the following hypotheses:

Hypothesis 4a (H4a): Conflicts will fully mediate the negative relationship between Activated Faultlines and Team Performance.

Hypothesis 4b (H4b): Conflicts will fully mediate the negative relationship between Activated Faultlines and Social Integration.

Figure 1 visualizes the relationship described in the section above.

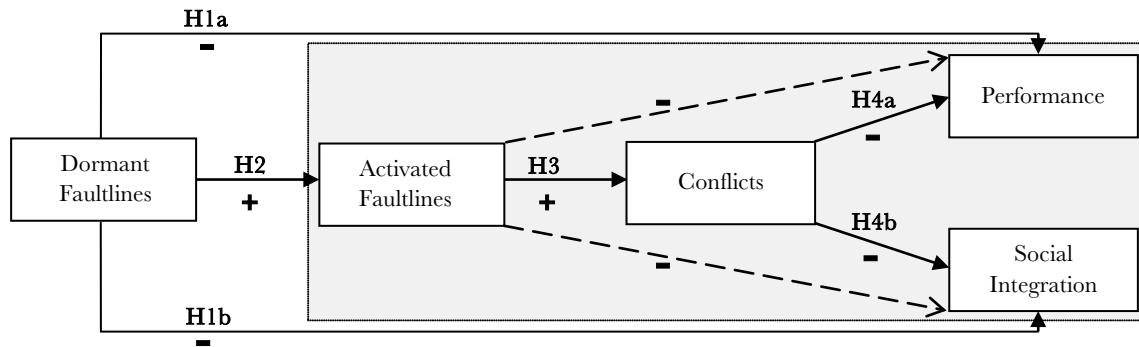


Figure 1. Research Model

Notes: Grey background indicates hypotheses that are applicable to both team and individual levels.

DATA AND METHODS

Sampling and Respondents

During the summer of 2018, we conducted this research in a Dutch cleaning company. The company has approximately ten thousand employees with more than a hundred different nationalities. Our sample consisted of 32 randomly chosen cleaning teams that are working in five Dutch provinces in different sectors such as offices (8 teams), educational institutions (13), hospitals (4), hotels (4) and shopping centres (3). By drawing data from diverse sectors and geographic locations, we aimed to improve the external validity of the research.

The team was defined as “a group of agents with a common goal which can only be achieved by appropriate combinations of individual activities” (Bacharach, 2005, p. xxi). Therefore, only teams that met the criteria of members’ interdependence were considered for participation. Employees are responsible for the cleaning of facades, sanitary, carpets, furniture, office equipment and window coverings. To be able to perform those tasks, teams have to divide functions and plan their activities. This requires a mutual understanding of the activities and willingness to cooperate with the colleagues in order to deal with situational issues.

Teams averaged 12 members and in total 300 respondents were included in the final sample. The average age was 47 years, 72 per cent were female and the number of different nationalities was 32. The demographic characteristics of all respondents are presented in Table II.

Procedure

We employed a survey method to collect data. First, we used the translation-back-translation procedure to translate the original items, specified in the following section, from English to Dutch and then back to English. Due to the fact that our sample consists of low-skilled employees and Dutch was not the mother tongue for most of them, we had to simplify some of the original items. Both versions were cross-checked by representatives of the company and an independent researcher.

Table II. Sample demographics

Variable name	Group	Number		%	
		All respondents	Dutch nationality	All respondents	Dutch nationality
Gender	Male	83	46	27.7%	25.7%
	Female	217	133	72.3%	74.3%
Age	16-25	23	14	7.7%	7.8%
	26-45	120	61	40.0%	34.1%
	46-66	157	104	52.3%	58.1%
Organizational Tenure	< 1	62	29	21.0%	16.7%
	2-3	78	53	26.9%	30.5%
	4-10	90	55	31.1%	31.6%
	> 11	61	37	21.1%	21.2%
Team Tenure	< 1	61	31	21.9%	18.6%
	2-3	74	54	26.5%	32.3%
	4-10	84	45	30.1%	27.0%
	> 11	60	37	21.6%	22.2%
Nationality	The Netherlands	179	179	59.7%	100.0%
	Morocco	26		8.7%	
	Turkey	25		8.3%	
	Ghana	14		4.7%	
	Other*	56		18.7%	
Country of Origin	The Netherlands	103	102	34.3%	57.0%
	Morocco	40	12	13.3%	6.7%
	Turkey	32	8	10.7%	4.5%
	Ghana	25	11	8.3%	6.1%
	Suriname	24	22	8.0%	12.3%
	Other*	76	24	25.3%	13.4%
Language proficiency	Low	56	11	19.1%	6.1%
	Middle	89	38	30.4%	21.6%
	High	148	127	50.5%	72.2%

*In total sample contains respondents with 32 different nationalities from 41 different countries of origin

Second, the questionnaire was pretested with one cleaning team. We asked cleaners to identify any ambiguities in the survey. Based on the responses from the pilot survey and comments of the respondents, we did not need to make any changes as the layout and wording were clear to the employees. Third, the cover letters that described the purpose of the study were mailed to team managers. In total, 9 teams declined to take part in this research. Fourth, the schedule of data collection was created as all of the 32 teams were visited in person. The main reason for that was to have a higher response rate and to increase the quality of the data. The survey included an opening statement explaining to the respondents what would happen to their data and emphasizing the confidential nature of their answers. The survey for team members consisted of 24 self-report, Likert scale questions that were randomly ordered and the survey for the team managers included 13 statements.

Every team was asked to gather together in order to get instruction on filling in the questionnaire and information about the purpose of the research. Respondents could choose the version of the questionnaire in Dutch or English. The researcher helped to those employees who had poor reading

skills or had questions concerning the questionnaire and controlled that all members filled in the questionnaire independently of each other. Team managers filled in the questionnaires at the same time. The researcher has asked team managers to indicate if some team members were absent, and has noted the ages, genders, nationalities and countries of origin of those employees. Thus, the full lists of demographic information have been created for every team. We have controlled that more than 70 per cent of employees in every team were able to fill in the questionnaire, thus, we could get a more objective opinion about every team.

Lastly, all team members returned the questionnaires directly to the researcher. If some items have not been answered, the researcher asked respondents if they would like to fill it in that improved the quality of the data. Visiting teams in person has also helped to create a better understanding of every team by talking to employees and team managers. After the end of the research, all team managers have received a summary of the main conclusions along with practical implications.

Measures

The survey instrument was administered in Dutch and English. Unless otherwise described, all variables were assessed using a 5-point Likert-type scale that ranged from 1 ('strongly disagree') to 5 ('strongly agree'). Complete lists of survey items for team members and managers are provided in Appendix A and Appendix B.

Independent variables. Dormant Faultlines' Strength. In the literature, several ways of calculating the strength of the faultlines (e.g. possibility of splitting the group into subgroups) can be found. The most frequently used index is Fau measure of faultline strength developed by Thatcher et al. (2003). However, Fau assumes that the team is able to crack only in two subgroups. While this may be reasonable for teams up to seven members, we think that teams larger than ten employees may have more than two subgroups. Therefore, we use the average silhouette width (ASW) algorithm developed by Meyer and Glenz (2013). ASW detects the number of subgroups and their homogeneity with a cluster-analytics procedure that involves several steps (for detailed example see Meyer and Glenz, 2013). Meyer, Shemla, Li, and Wegge (2015, p. 365) explain that "silhouette widths quantify how well a team member fits into his or her subgroup. Therefore, the average silhouette width (ASW) quantifies the average fit of all team members to their subgroup". The ASW values may range between -1 and 1, where the closer value is to 1, the better the cluster solution for the whole group (e.g. the stronger faultline the group has) (Meyer & Glenz, 2013). Example of ASW output for one of the teams from our sample and explanations are given in Appendix C.

Dormant Faultlines' Strength was calculated using the `asw.cluster` package in the supplemental material of Meyer and Glenz (2013) in R. To determine the *Strength* of surface-level *Dormant Faultlines*, we calculated ASW across such demographic characteristics as gender, age, nationality and origin of employees. Moreover, we decided to calculate ASW based on different combinations of those

characteristics (like 3: gender, age, nationality, or 2: gender, origin or 1: nationality only), as we were not able to find any guideline in the faultline literature on which attributes would be the most appropriate for our study. An average ASW strength varied from 0.57 (SD = 0.17) if calculated based on gender, age and origin to 0.82 (SD = 0.11) if calculated based on nationality only.

Dependent variables. Team Performance. First, *Perceived Team Performance* was measured by means of two 5-point items adapted from Jehn et al. (1999), including following items: ‘Our team performs well’ and ‘Our team is effective at getting things done quickly’ ($\alpha = 0.73$). Second, items for the *Team Performance* based on managers’ opinions were based on the scale developed by Vermeeren (2014) and Bernardin (2002). We had five adapted sample items that included statements like ‘This team accomplishes its cleaning tasks within the timeframe’ and ‘The quality of cleaning of this team is high’ ($\alpha = 0.81$).

Social Integration. To assess the extent of *Social Integration* in the team, employees were asked questions about their satisfaction, cohesion and intention to remain (Jehn et al., 1999; Rico et al., 2007). Three items assessed the extent to which the employees are satisfied with their colleagues were adapted from the Team Diagnostic Survey (Wageman, Hackman, & Lehman, 2005): ‘My relations with other team members are strained (R)’, ‘I enjoy working with my teammates’, ‘Generally speaking, I am satisfied with this team’. Cohesion was measured with three items adapted from Krogel (2008): ‘I feel accepted by this team’, ‘I feel bad about being a part of this team (R)’, ‘Even though others may disagree with me sometimes, I feel accepted in this team’. The intent to remain was measured by using measure of tenure intention (Kraut, 1975) that was used by Jehn et al. (1999) and includes following items: ‘I would like to stay long in this team’, ‘If I have my own way, I will be working in this same team three years from now’ and ‘I would like to change this teams and work in another one (R)’. Three measures were averaged to create a composite measure of teams’ *Social Integration* ($\alpha = 0.78$).

Mediating variables. Activated surface-level diversity faultlines. Activated Faultlines were measured based on one item adapted from Jehn and Bezrukova (2010): ‘My team is cracked into smaller cliques’. We could not use the full scale as the wording of other items would be not clear to low-skilled employees. Therefore, we included four additional items for several sub-types of faultlines such as ‘In my team people mostly communicate with colleagues of the same nationality’ and ‘In my team young people have less contact with older people’ ($\alpha = 0.71$). These items reflect if group members perceived a division of the group into multiple subgroups based on surface-level characteristics. In order to assess the *Activated Faultlines* based on a manager’s opinion we slightly changed the wording of the statements mentioned above and used the same items ($\alpha = 0.70$).

Conflicts. *Conflicts* in teams were measured by a total of 7 items: three items on task conflict adapted from Jehn (1995) and four items on relationship conflict adapted from Jehn et al. (1999). Examples of included items are: ‘In my team we often disagree about opinions regarding the work being done’, ‘There are conflicts in my team about the order of the tasks’ and ‘There is much friction

(disagreement) among members of my team' ($\alpha = 0.78$). Managers were also asked to indicate the level of *Conflicts* in their teams and four sample items included: 'There is tension among members of this team' and 'This team has conflicts about the way they should perform their tasks' ($\alpha = 0.87$).

Control variables. Several other variables may influence the relationship postulated above. For team-level control variables, we used *Team Size* that was measured by the number of tenured employees. *Team Size* is considered to have an effect on the team processes and outcomes (Li & Hambrick, 2005; Spell, Bezrukova, Haar, & Spell, 2011). We calculated the team-level *Diversity* of age, gender, nationality and origin in order to isolate the unique effect of faultlines (Rico et al., 2007). Blau (1977) heterogeneity index was used for calculation of *Diversity* based on all four demographic variables. For example of calculations see Appendix D. For individual-level analysis we assessed *Perceived-Diversity* by item 'My team consist of many diverse people in terms of nationality, gender and age'. In addition, for individual-level analysis, we used employees' *Team Tenure* measured as the number of years in the team because experience may have an effect on *Team Performance* and *Social Integration* (Bezrukova et al., 2016). Lastly, we controlled for the *Language Proficiency* of the employee as, similarly to tenure, it may have an effect on the team process and outcomes.

Aggregation

The items measuring *Activated Faultlines*, *Conflicts*, *Social Integration* and *Perceived Team Performance* were phrased focusing on the team as a unit of analysis. Likert scale questions investigated an individual viewpoint, thus the data has to be aggregated to the team-level. James, Demaree, and Wolf (1993) developed the index appropriate for within-group agreement of a set of judges rating a single target with a multiple-item scale: $r_{WG(j)}$ coefficients. The application of $r_{WG(j)}$ in this research is based on assumption that each team has a true score on the assessed construct (e.g. level of *Activated Faultlines*) and any variance among responses of team members is assumed to be error variance. As stated in Biemann, Cole, and Voelpel (2012, p. 68): "Interrater agreement can be estimated by comparing the observed variance to the variance expected when there is complete lack of agreement among raters (i.e., random responding)". In order to address assumptions, 5-point response ramp was used and each of multiple items sets are assumed to measures one-dimensional constructs. Biemann et al. (2012, p. 73) provide following threshold for $r_{WG(j)}$: "lack of agreement"= 0.00 to 0.30; "weak agreement"= 0.31 to 0.50; "moderate agreement"= 0.51 to 0.70; "strong agreement"= 0.71 to 0.90, and; "very strong agreement"= 0.91 to 1.00.

Results of the suitability of data for aggregation by using $r_{WG(j)}$ can be seen in Table III. Median interrater agreement scores are 0.65 for *Activated Faultlines*, 0.78 for *Conflicts*, 0.73 for *Perceived Performance* and 0.89 for *Social Integration* that means that all values are on the accepted level. As LeBreton and Senter (2008) do not advice to use uniform null distribution, statistics for moderate skew distribution is reported as well. It can be seen that the mean of $r_{WG(j)}$ falls to the level of weak

agreement. Nonetheless, Biemann et al. (2012) analyzed studies where researchers needed to aggregate the data and in the vast majority of the cases uniform distribution was used and $r_{WG(j)}$ greater than 0.70 was considered sufficient to warrant aggregation.

Additionally, $r_{WG(j)}$ value is the measure of within-group agreement and does not measure the reliability. Bliese (2000) claims that reliability can be high while the agreement is low. Therefore, we conducted the second test that determines whether between-group differences were stronger than within-group differences. We followed the procedure suggested by Bliese (2000). First, ANOVA-based significance test of between-group differences is significant for all variables ($p < 0.01$ for *Activated Faultlines* and *Conflicts* and $p < 0.02$ for *Social Integration*) except *Perceived Performance* ($p = 0.243$). Next, we calculated the interclass correlation coefficient (ICC(1)). This index was interpreted by James (1982) as “the degree of reliability associated with a single assessment of the group mean” and was recommended to be used as a criterion of aggregating (Bliese, 2000, p. 355). Having the value of ICC(1) higher than 0.05, researcher can consider the agreement between variables to be acceptable (Bliese, 2000). For example $ICC(1) = 0.12$ for *Activated Faultlines* suggests that group membership explains 12 per cent of the variance in individual group-members’ ratings. ICC(2) indicates how reliably the aggregate group-level means distinguishes between groups (Bliese, 2000). Based on results, we conclude that some group-level effect takes place for *Activated Faultlines*, *Conflicts* and *Social Integration* and our data can be aggregated to the team-level for subsequent analysis. However, *Perceived Performance* cannot be aggregated to the team-level based on ICC(1) as the between-group variance is low. In this case the team-level analysis will be conducted based on managers’ scores for *Performance*.

Table III. Aggregation results of team member individual ratings to a team-level

Measure	$r_{WG(j)}$, uniform		$r_{WG(j)}$, measure-specific						
	Mean	SD	Shape	σ_E^2	Mean	SD	F ratio	ICC(1)	ICC(2)
Activated Faultlines	0.65	0.28	Moderate Skew	0.90	0.04	0.14	2.32**	0.12	0.57
Conflicts	0.78	0.20	Moderate Skew	0.90	0.16	0.32	2.68**	0.15	0.63
Perceived Performance	0.73	0.23	Moderate Skew	0.90	0.39	0.35	1.18	0.02	0.15
Social Integration	0.89	0.09	Moderate Skew	0.90	0.34	0.41	1.69*	0.07	0.41

SD = standard deviation of $r_{WG(j)}$ values; Shape = the shape of an alternative null distribution; σ_E^2 = variance of an alternative null distribution. Variance estimations for moderate null distributions were taken from LeBreton and Senter (2008, p. 832) for 5-Point Scale.

* $p < 0.02$, ** $p < 0.01$.

RESULTS

Group-level analysis

Means, standard deviations, and correlations for dependent, independent, and control variables at the team-level are summarized in Table IV.

Table IV. Means, standard deviations, and correlations of dependent, independent, and control variables on the group-level

	<i>Mean</i>	<i>SD</i>	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>	<i>6</i>	<i>7</i>	<i>8</i>	<i>9</i>	<i>10</i>	<i>11</i>	<i>12</i>	<i>13</i>	<i>14</i>	<i>15</i>	<i>16</i>	<i>17</i>
Dormant Faultlines' Strength (ASW) based on:																			
1 Gender, Age, Nationality	0.64	0.15	1																
2 Gender, Nationality	0.76	0.13	0.76**	1															
3 Nationality	0.82	0.11	0.70**	0.73**	1														
4 Gender, Age, Origin	0.57	0.17	0.52**	0.31	0.34	1													
5 Gender, Origin	0.68	0.15	0.43*	0.52**	0.35	0.84**	1												
6 Origin	0.77	0.13	0.38*	0.42*	0.50**	0.68**	0.78**	1											
Avg. of individual grades:																			
7 Activated Faultlines	1.21	0.41	-0.13	-0.16	0.07	0.04	0.11	0.23	1										
8 Conflicts	1.16	0.43	-0.17	-0.21	0.02	-0.10	-0.07	0.07	0.74**	1									
9 Social Integration	3.21	0.28	0.10	0.28	-0.02	0.22	0.21	-0.02	-0.71**	-0.76**	1								
Team managers' grades:																			
10 Activated Faultlines	0.84	0.65	0.12	0.12	0.24	0.14	0.08	0.25	0.54**	0.49**	-0.35	1							
11 Conflicts	0.76	0.84	-0.16	-0.17	-0.01	-0.11	-0.05	0.14	0.35	0.51**	-0.48**	0.60**	1						
12 Performance	3.14	0.56	0.05	0.07	-0.03	0.13	0.16	0.02	-0.27	-0.37*	0.26	-0.45*	-0.54**	1					
Control variables:																			
13 Team Size	12.5	3.85	0.01	-0.16	0.07	0.13	0.02	0.14	0.53**	0.52**	-0.35	0.31	0.09	-0.24	1				
14 Origin Diversity	0.64	0.15	-0.12	-0.16	0.01	-0.44*	-0.53**	-0.44*	0.35	0.24	-0.36	0.11	0.03	-0.13	0.30	1			
15 Nationality Diversity	0.55	0.17	-0.43*	-0.48**	-0.29	0.14	-0.07	-0.02	0.43*	0.19	-0.10	0.16	0.03	-0.16	0.34	0.38*	1		
16 Age Diversity	0.30	0.16	-0.30	-0.26	-0.07	-0.34	-0.46**	-0.16	-0.05	0.02	-0.16	-0.04	-0.12	-0.14	0.08	0.21	0.15	1	
17 Gender Diversity	0.48	0.12	-0.25	-0.21	-0.29	-0.29	-0.23	-0.39*	0.11	0.11	-0.07	-0.16	-0.16	0.11	0.28	0.05	0.13	-0.01	1

N = 32 teams; Bold values indicate hypothesized relationships; **, * Correlation is significant at the 0.01 level, 0.05 level (2 - tailed)

Activated Faultlines and *Conflicts* in Table IV are reported as averaged scores given by team members and scores given by managers of the same teams. We see that those scores are significantly positively correlated. Team outcomes variables are given as follows: *Social Integration* score is given as average for every team and *Team Performance* is given based on grades of manager to the team. Teams in our sample have on average 12 employees. *Dormant Faultlines' Strength* variables are positively correlated, however, are not related to the team functioning variables.

Dormant Faultlines' Strength and Activated Faultlines

Prior to testing mediation models for averaged and managers' scores (H3, H4a and H4b), we test Hypotheses 1a and 1b that state that strong *Dormant Faultlines* should have negative influence on team outcomes (*Performance* and *Social Integration*) and the H2 that states that the stronger *Dormant Faultlines* in team, the higher *Activated Faultlines'* score. Results of individual linear regressions are given in Table V. Regressions were performed separately for every value of *Dormant Faultlines' Strength* and for every outcome variable. Thus, 24 linear regressions were performed in order to find whether one of *Dormant Faultlines' Strength* values has some effect on one of the dependent variables. It can be seen that none of the *Dormant Faultlines' Strength* values had a significant effect on *Activated Faultlines* and team outcomes. Thus, H1a, H1b and H2 were not supported. We do not to test the effect of the *Dormant Faultlines' Strength* on *Conflicts* as from the Table IV it can be seen that those variables are not correlated. In order to illustrate some possible reasons why no relationship was found, we provide information about four teams from our sample (see Table VI). Full team demographic information can be seen: age, gender, nationality and country of origin of every member. Next to it, the values of *Dormant Faultlines' Strength*, *Diversity* and *Activated Faultlines* are provided.

Table V. Simple Linear Regression results of the effect of Dormant Faultlines' Strength on Faultlines' Activation ^a and team outcomes ^b

Predicting Variable	Activated Faultline (Avg.)		Activated Faultline (Manager)		Social Integration (Avg.)		Team Performance (Manager)	
	Beta	t	Beta	t	Beta	t	Beta	t
DFGAN	-0.122	-0.676	0.116	0.631	0.073	0.403	0.046	0.248
DFGN	-0.161	-0.895	0.121	0.657	0.285	1.631	0.074	0.397
DFN	0.069	0.380	0.237	1.316	-0.008	-0.046	-0.034	-0.184
DFGAO	0.041	0.225	0.136	0.741	0.219	1.232	0.134	0.726
DFGO	0.107	0.592	0.082	0.441	0.219	1.229	0.162	0.885
DFO	0.226	1.269	0.252	1.404	-0.008	-0.045	0.019	0.102

N = 32 at the team-level; Standardized Regression coefficients are reported;

DF – Dormant Faultline Strength; G – Gender, A – Age, N- Nationality, O – Country of Origin;

^a based on averages of team members' grades and managers' grades;

^b based on averages of team members' grades for Social Integration and managers' grades for Performance;

* p < 0.1, ** p < 0.05

Table VI. Values of Dormant Faultlines' Strength, Diversity and Activated Faultlines for four teams from the sample

Team Demographics																
N	Team 1				Team 2				Team 3				Team 4			
	Age ^a	Gender ^b	Natl. ^c	Origin	Age	Gender	Natl.	Origin	Age	Gender	Natl.	Origin	Age	Gender	Natl.	Origin
1	2	m	NL	GH	2	m	NL	NL	3	f	TK	TK	3	f	FR	SM
2	2	m	NL	GH	3	m	NL	TK	3	f	TK	TK	2	f	TK	TK
3	2	m	SR	SR	3	m	M	M	3	f	NL	NL	2	f	TK	TK
4	2	m	GH	GH	3	f	CH	CH	3	f	NL	M	3	f	TK	TK
5	1	m	NL	SM	3	m	NL	M	3	f	NL	TK	2	f	TK	TK
6	3	m	NL	NL	2	m	CO	CO	2	f	NL	TK	3	f	NL	NL
7	3	m	NL	SO	3	m	M	M	2	f	NL	TK	3	m	NL	SL
8	3	m	NL	SR	2	f	CO	CO	3	f	NL	NL	3	f	NL	NL
9	3	f	CZ	CZ	3	m	NL	VT	3	f	NL	NL				
10	2	f	NL	SR	3	m	GH	GH	3	f	TK	TK				
11	2	m	NI	NI	3	m	M	M								
12	3	m	NL	IN	2	f	NL	NL								
13					3	f	M	M								
14					3	m	M	M								

Indexes and results					
		Team 1	Team 2	Team 3	Team 4
Dormant Faultlines' Strength ^d	DFGAN	0.5	0.56	1	0.63
	DFGN	0.61	0.6	1	0.75
	DFGAO	0.36	0.43	0.9	0.63
	DFGO	0.44	0.41	0.9	0.75
	DFN	0.67	0.86	1	0.88
	DFO	0.5	0.71	0.9	0.75
Diversity	Gender	0.28	0.41	0	0.22
	Age	0.57	0.41	0.32	0.47
	Natl.	0.53	0.71	0.42	0.59
	Origin	0.83	0.76	0.54	0.66
Activated Faultlines	Manager	1.2 (Yes) ^e	1.8 (Yes)	0.0 (No)	0.0 (No)
	Avg. of team	1.0 (Yes)	1.8 (Yes)	0.7 (No)	0.9 (No)

^a 1 = 16-25; 2= 26-45; 3 = 46-66;

^b m = male; f = female;

^c Natl. = Nationality; NL = the Netherlands, SR= Suriname, GH= Ghana, CZ = Czech Republic, SO = Somalia, NI = Nigeria, M = Morocco, CH = China, CO = Colombia, TK = Turkey, FR = France, SM = Saint Martin;

^d DF – Dormant Faultline Strength; G – Gender, A – Age, N- Nationality, O – Country of Origin; For both Faultline Strength and Diversity: 0 = none; 1 = very high;

^e Evaluation is given based on comparison to results of other teams in the sample.

First of all, we would like to draw your attention to the significant change in *Dormant Faultlines' Strength* based on variables used for calculation. For example, the second team in Table VI has weak faultline (0.43) if we take into account all information on ages, genders and origins of its members, but if we only consider their nationalities, the faultline becomes extremely strong (0.86). We could not find any guideline in current diversity literature on which variables it is reasonable to choose, therefore, we had to calculate all possible indexes and conduct analysis with all of them.

Secondly, we show that the first two teams have low *Dormant Faultlines' Strength* in comparison to teams three and four. Yet, we see that in the first two teams faultlines are activated based on the opinions of team members and managers while in the last two teams they stay dormant. Thus, we assume that some non-demographic variables may play a role in subgroups formation. During our data collection, we observed different managers' attitudes towards diversity, and those managers that initiate joint team celebrations (e.g. celebrate both Easter and Ramadan) and organize special diversity-oriented events (diversity dinner, where all members bring self-cooked national dishes), have also reported that their team has no subgroups. However, we have not conducted structured interviews with those managers, hence, this observation needs to be validated.

One more reason for such unexpected subgroup formation may be deep-level diversity. Here we suppose that team members do not categorize each other based on their surface-level differences but based on their job-attitude, functional- background or many other possible underlying attributes. To give an example, we have seen that some employees prefer to come to their work earlier in order to drink coffee. This habit clearly does not depend on demographic differences but on the different personalities of employees. However, it may split the team into two subgroups: those who drink coffee before their work versus those who do not.

In addition, we also would like to mention that if an employee has Dutch nationality it does not necessarily mean that he or she speaks Dutch on a high level or feels like a part of Dutch society. For example, we have seen teams where two foreign employees have worked for more than 20 years in the team but always stick together and do not seek out contact with Dutch colleagues. In contrast, we have also met people who have lived less than five years in the Netherlands and do not have Dutch nationality but speak Dutch fluently. Therefore, we conclude that all those differences cannot be captured in one number that explains why we do not observe any relationship between *Dormant Faultlines' Strength* and *Activated Faultlines* or team outcomes. *Dormant Faultlines' Strength* variables will not be included in the following analyses.

Activated Faultlines, Conflicts and Team Outcomes

We continue to analyze the data in order to support or reject our following hypotheses. H3 states that if *Faultlines* are *Activated*, then this team would have more *Conflicts*. H4a and H4b state that *Conflicts* in the team would fully mediate the relationship between *Activated Faultlines* and team outcomes. To test it, we used PROCESS macro for SPSS and the bootstrapping procedure from Hayers (2013, Model 4). We have performed two analyses based on averages of team members' grades and on managers' grades. The results can be seen in Table VII and Table VIII. For every mediator and outcome variable we also performed analyses with only controls to ensure that our results are not driven by teams' size or diversity on demographic attributes.

Table VII shows that *Activated Faultlines* (Avg.) have a significant positive effect on *Conflicts* (Avg.) ($\beta = 0.756, p < 0.01$). *Conflicts* (Avg.) have a significant negative effect on *Social Integration* (Avg.) ($\beta = -0.293, p < 0.05$). The indirect effect of *Activated Faultlines* (Avg.) on *Social Integration* (Avg.) through *Conflicts* (Avg.) is confirmed by the bootstrap results as the bootstrapped 95 per cent confidence interval (CI) around the indirect effect does not contain zero (-0.4553, -0.0126). However, the direct effect of *Activated Faultlines* (Avg.) on *Social Integration* (Avg.) is -0.341 that is stronger than the strength of indirect effect via *Conflicts* (Avg.) that is -0.221. Thus, H4b is only partially supported as while the indirect effect was found, the direct effect is slightly stronger. That means that partial instead of full mediation was observed. One of the control variables has a significant positive effect on *Social Integration: Team Diversity* based on nationality. So the more different nationalities work in one team, the more satisfied and cohesive the team seems to be.

The results of the analysis based on grades given by managers can be seen in Table VIII. We followed the same procedure as described before. We see similar results to those obtained from analysis of averaged team members' grades. *Activated Faultlines* have a significant positive effect on *Conflicts* ($\beta = 0.804, p < 0.01$). This finding again provides support for H3. The *Conflicts* have a significant negative effect on *Team Performance* ($\beta = -0.313, p < 0.05$). The indirect effect of *Activated Faultlines* on *Team Performance* through *Conflicts* is confirmed as the bootstrapped 95 per cent CI around the indirect effect excludes zero (-0.660, -0.050). However, *Activated Faultlines* (Manager) do not have a direct effect on *Team Performance* as the bootstrapped CI around the direct effect includes zero (-0.494, 0.300). Thus H4a is supported as full mediation is found. None of the control variables had a significant effect on either *Conflicts* or *Team Performance*.

Table VII. Regression results for simple mediation model of Activated Faultlines (Avg.) on Team Social Integration (Avg.) through Conflicts (Avg.)

Model	b coeff	SE	t		
Control variables (DV = Conflicts)					
Constant	3.614	0.425	8.510**		
Team size	-0.056	0.020	-2.734*		
Origin Diversity	-0.299	0.521	-0.574		
Nationality Diversity	0.046	0.463	0.098		
Age Diversity	0.133	0.626	0.212		
Gender Diversity	0.088	0.443	0.199		
R ² = 0.280, F= 2.026, p= 0.108					
Mediator variable model (DV = Conflicts)					
Constant	0.219	0.328	0.669		
Activated Faultlines	0.756	0.169	4.467**		
Team size	0.022	0.017	1.254		
Origin Diversity	-0.041	0.404	-0.101		
Nationality Diversity	-0.502	0.373	-1.346		
Age Diversity	-0.008	0.474	-0.161		
Gender Diversity	0.205	0.343	0.598		
R ² = 0.600, F= 6.2436, p= 0.0004					
Control variables (DV = Social Integration)					
Constant	3.753	0.287	13.078**		
Team size	-0.022	0.014	-1.557		
Origin Diversity	-0.543	0.352	-1.542		
Nationality Diversity	0.221	0.313	0.706		
Age Diversity	0.005	0.423	0.013		
Gender Diversity	-0.159	0.299	-0.530		
R ² = 0.215, F= 1.421, p= 0.250					
Dependent variable model (DV = Social Integration)					
Constant	3.930	0.182	21.563**		
Activated Faultlines	-0.341	0.125	-2.727*		
Conflicts	-0.293	0.110	-2.658*		
Team size	0.010	0.010	1.018		
Origin Diversity	-0.307	0.223	-1.378		
Nationality Diversity	0.430	0.213	2.019*		
Age Diversity	-0.074	0.261	-0.284		
Gender Diversity	-0.318	0.191	-1.670		
R ² = 0.721, F= 8.84, p= 0.0000					
Total, direct and indirect effects					
Total effect of Activated Faultlines on Social Integration	Effect	SE	t	LLCI	ULCI
	-0.563	0.104	-5.412**	-0.7772	-0.3487
Direct effect of Activated Faultlines on Social Integration	Effect	SE	t	LLCI	ULCI
	-0.341	0.125	-2.727*	-0.5996	-0.0830
Indirect effect of Activated Faultlines on Social Integration through Conflicts	Effect	Boot SE	BootLLCI	BootULCI	
	-0.221	0.116	-0.4553	-0.0126	

N = 32 teams. Unstandardized regression coefficients are reported. Bootstrap sample size is 5000.

LL=lower limit, UL=upper limit, CI=confidence interval. The bootstrap CI were estimated using the SPSS PROCESS macro (version 3.00; Hayes, 2013); * p < .05; ** p < .01, two-tailed.

Table VIII. Regression results for simple mediation model of Activated Faultlines (Manager grade) on Team Performance (Manager grade) through Conflicts (Manager grade)

Model	b coeff	SE	t		
Control variables (DV = Conflicts)					
Constant	1.118	0.973	1.149**		
Team size	0.029	0.047	0.620		
Origin Diversity	0.085	1.198	0.071		
Nationality Diversity	0.231	1.076	0.215		
Age Diversity	-1.436	1.434	-1.001		
Gender Diversity	-0.711	1.019	-0.698		
R ² = 0.062, F= 0.329, p= 0.890					
Mediator variable model (DV = Conflicts)					
Constant	0.566	0.831	0.681		
Activated Faultlines	0.804	0.230	3.494**		
Team size	-0.019	0.042	-0.464		
Origin Diversity	0.144	0.998	0.144		
Nationality Diversity	-0.164	0.912	-0.180		
Age Diversity	-0.211	1.233	-0.171		
Gender Diversity	-0.452	0.850	-0.532		
R ² = 0.377, F= 2.423, p= 0.057					
Control variables (DV = Team Performance)					
Constant	3.430	0.633	5.423**		
Team size	-0.037	0.031	-0.040		
Origin Diversity	-0.031	0.779	-0.358		
Nationality Diversity	-0.251	0.700	-0.563		
Age Diversity	0.937	0.932	-1.218		
Gender Diversity	-0.373	0.662	1.005		
R ² = 0.111, F= 0.624, p= 0.683					
Dependent variable model (DV = Team Performance)					
Constant	3.867	0.570	6.781**		
Activated Faultlines	-0.097	0.192	-0.505		
Conflicts	-0.313	0.139	-2.253*		
Team size	-0.021	0.029	-0.747		
Origin Diversity	0.019	0.678	0.028		
Nationality Diversity	-0.225	0.620	-0.363		
Age Diversity	0.333	0.839	0.398		
Gender Diversity	-0.619	0.581	-1.066		
R ² = 0.382, F= 2.029, p= 0.095					
Total, direct and indirect effects					
Total effect of Activated Faultlines on Team Performance	Effect	SE	t	LLCI	ULCI
	0.620	0.169	-2.059	-0.697	0.001
Direct effect of Activated Faultlines on Team Performance	Effect	SE	t	LLCI	ULCI
	0.173	0.192	-0.505	-0.494	0.300
Indirect effect of Activated Faultlines on Team Performance through Conflicts	Effect	Boot SE	BootLLCI	BootULCI	
	-0.251	0.153	-0.660	-0.050	

N = 31 teams. Unstandardized regression coefficients are reported. Bootstrap sample size is 5000.

LL=lower limit, UL=upper limit, CI=confidence interval. The bootstrap CI were estimated using the SPSS PROCESS macro (version 3.00; Hayes, 2013); * p < .05; ** p < .01, two-tailed.

Individual-level analysis

In the first analysis we again computed means and standard deviations for individual scores of *Activated Faultlines*, *Conflicts*, *Social Integration*, *Perceived Performance* and control variables as well as the correlation between these variables (see Table IX). Table IX shows that four main variables are significantly related to each other. To further test our hypotheses, we used a simple mediation model (see Table X and Table XI). In a similar manner, H3 states that the stronger employee perceives *Activated Faultlines*, the more *Conflicts* in the team will be indicated, and *Conflicts* will fully mediate the relationship between *Activated Faultlines* and *Team Performance* (H4a) and *Social Integration* (H4b).

Table IX. Descriptive statistics and pairwise correlations

	Mean	SD	1	2	3	4	5	6	7
Individual scores:									
1 Activated Faultlines	1.20	0.81	1						
2 Conflicts	1.18	0.76	0.55**	1					
3 Social Integration	3.23	0.61	-0.49**	-0.59**	1				
4 Perceived Performance	3.14	0.78	-0.34**	-0.44**	0.52**	1			
Control variables:									
5 Team Tenure ^a	2.08	1.8	0.01	0.06	0.10	-0.00	1		
6 Language Proficiency ^b	2.30	0.85	0.01	0.02	-0.01	-0.12	-0.06	1	
7 Perceived Team Diversity ^c	3.36	0.81	-0.12	-0.12*	0.17**	0.12	-0.09	-0.11	1

N = 300; **, * Correlation is significant at the 0.01 level, 0.05 level (2 – tailed);

^a 0 = less than one year; 3 = 7 to 10 years; 6 = more than 26 years;

^b 0 = very low; 3 = high;

^c 0 = no diversity; 4 = high diversity.

First of all, results obtained from a model with *Perceived Performance* as the outcome variable will be described (see Table X). We tested models with control variables to test whether *Team Tenure*, *Language Proficiency* or *Perceived Team Diversity* explain why individual reports more *Conflicts* or better *Team Performance*. Control variables models explained less than 3% of the variance in *Conflicts* and *Perceived Performance* and none of the coefficients was statistically significant.

From the mediation variable model in Table X we see that if an employee perceives subgroups in the team then also more *Conflicts* are likely to be indicated ($\beta = 0.507$, $p < 0.01$). This provides support for H3. Furthermore, the stronger team members perceive *Conflicts* the lower they grade *Team Performance* ($\beta = -0.375$, $p < 0.01$). Finally, *Activated Faultlines* have a direct effect on *Perceived Performance* as the bootstrapped 95 per cent CI around the direct effect excludes zero (-0.266; -0.018). However, the indirect effect through *Conflicts* is stronger than the direct one: -0.190 and -0.142 respectively. This provides support to H4a that is in line with our previous results obtained from grades of team managers. In addition, one control variable *Language Proficiency* has a significant negative effect on *Perceived Performance*. Thus, the better employee speaks Dutch the lower grades to *Team Performance* are given.

Table X. Regression results for simple mediation model of Activated Faultlines (Indv.) on Perceived Performance (Indv.) through Conflicts (Indv.)

Model	b coeff	SE	t		
Control variables (DV = Conflicts)					
Constant	1.479	0.259	5.709**		
Team Tenure	0.017	0.026	0.643		
Language Proficiency	0.005	0.055	0.089		
Perceived Team Diversity	-0.100	0.057	-1.743		
R ² = 0.014, F= 1.261, p= 0.288					
Mediator variable model (DV = Conflicts)					
Constant	0.665	0.231	2.881**		
Activated Faultlines	0.507	0.048	10.568**		
Team Tenure	0.019	0.022	0.860		
Language Proficiency	0.010	0.046	0.227		
Perceived Team Diversity	-0.045	0.049	-0.934		
R ² = 0.306, F= 29.258, p= 0.000					
Control variables (DV = Perceived Performance)					
Constant	3.053	0.269	11.358**		
Team Tenure	0.007	0.027	0.251		
Language Proficiency	-0.099	0.057	-1.743		
Perceived Team Diversity	0.085	0.060	1.420		
R ² = 0.021, F= 1.918, p= 0.127					
Dependent variable model (DV = Perceived Performance)					
Constant	3.835	0.258	14.851**		
Activated Faultlines	-0.142	0.063	-2.255*		
Conflicts	-0.375	0.068	-5.559**		
Team Tenure	0.013	0.025	0.518		
Language Proficiency	-0.099	0.051	-1.948*		
Perceived Team Diversity	0.032	0.054	0.592		
R ² = 0.227, F= 15.536, p= 0.0000					
Total, direct and indirect effects					
Total effect of Activated Faultlines on Perceived Performance	Effect	SE	t	LLCI	ULCI
	-0.332	0.056	-5.962**	-0.442	-0.223
Direct effect of Activated Faultlines on Perceived Performance					
	-0.142	0.063	-2.255*	-0.266	-0.018
Indirect effect of Activated Faultlines on Perceived Performance through Conflicts	Effect	Boot SE	BootLLCI	BootULCI	
	-0.190	0.047	-0.292	-0.105	

N = 271 individuals. Unstandardized regression coefficients are reported. Bootstrap sample size is 5000.

LL=lower limit, UL=upper limit, CI=confidence interval. The bootstrap CI were estimated using the SPSS PROCESS macro (version 3.00; Hayes, 2013); * p < .05; ** p < .01, two-tailed.

Table XI shows how *Social Integration* is affected by the perception of *Activated Faultlines* and *Conflicts*. Firstly, we test models with only controls in order to see if *Team Tenure*, *Language Proficiency* or *Perceived Team Diversity* explain why individuals are satisfied with the team. It is found that *Team Tenure* and *Perceived Team Diversity* are positively associated with *Social Integration*. However, the coefficient of determination is very low ($R^2 = 0.021$), thus, the very small proportion of variance in *Social Integration* is explained by the model with control variables. In addition, we again see that if employee perceives *Activated Faultlines* then also more *Conflicts* are likely to be indicated ($\beta = 0.502$, $p < 0.01$) that supports

H2b (slightly different numbers come from the fact that due to missing values in the previous model 271 individuals were analysed while now 266 were included).

Table XI. Regression results for simple mediation model of Activated Faultlines (Indv.) on Social Integration (Indv.) through Conflicts (Indv.)

Model	b coeff	SE	t		
Control variables (DV = Conflicts)					
Constant	1.479	0.259	5.709**		
Team Tenure	0.017	0.026	0.643		
Language Proficiency	0.005	0.055	0.089		
Perceived Team Diversity	-0.100	0.057	-1.743		
R ² = 0.014, F= 1.261, p= 0.288					
Mediator variable model (DV = Conflicts)					
Constant	0.667	0.233	2.904**		
Activated Faultlines	0.502	0.049	10.334**		
Team Tenure	0.021	0.023	0.911		
Language Proficiency	0.014	0.047	0.302		
Perceived Team Diversity	-0.051	0.049	-1.046		
R ² = 0.303, F= 28.314, p= 0.000					
Control variables (DV = Social Integration)					
Constant	2.657	0.205	12.957**		
Team Tenure	0.041	0.021	1.968*		
Language Proficiency	0.014	0.044	0.313		
Perceived Team Diversity	0.135	0.046	2.966**		
R ² = 0.043, F= 3.881, p= 0.010					
Dependent variable model (DV = Social Integration)					
Constant	3.477	0.174	19.971**		
Activated Faultlines	-0.171	0.042	-4.025**		
Conflicts	-0.366	0.046	-8.043**		
Team Tenure	0.048	0.017	2.897**		
Language Proficiency	0.016	0.034	0.478		
Perceived Team Diversity	0.075	0.036	2.067*		
R ² = 0.412, F= 36.389, p= 0.0000					
Total, direct and indirect effects					
Total effect of Activated Faultlines on Social Integration	Effect	SE	t	LLCI	ULCI
	-0.355	0.040	-8.896**	-0.433	-0.276
Direct effect of Activated Faultlines on Social Integration	-0.171	0.042	-4.025**	-0.254	-0.087
Indirect effect of Activated Faultlines on Perceived Performance through Conflicts	Effect	Boot SE	BootLLCI	BootULCI	
	-0.184	0.034	-0.256	-0.122	

N = 266 individuals. Unstandardized regression coefficients are reported. Bootstrap sample size is 5000.

LL=lower limit, UL=upper limit, CI=confidence interval. The bootstrap CI were estimated using the SPSS PROCESS macro (version 3.00; Hayes, 2013); * p < .05; ** p < .01, two-tailed.

From the Table XI we see that the more *Conflicts* team member perceives the lower is the grade for *Social Integration* ($\beta = -0.366$, $p < 0.01$) that means that employee is not satisfied and is not willing to stay long in the team. The direct effect of *Activated Faultlines* on *Social Integration* is found as the bootstrapped CI does not contain zero (-0.0254, -0.087). But the indirect effect through *Conflicts* is

stronger (-0.184) than the direct one (-0.171). Thus *Conflicts* mediate the relationship between *Activated Faultlines* and *Social Integration* that supports H4b. This result is in line with previous findings obtained from averaged team-level scores. Two control variables are found to have a significant positive effect on *Social Integration*: *Team Tenure* and *Perceived Diversity*. Thus, those employees who work longer in their team are not willing to leave their team and are satisfied with it. Also, those who report that their team is diverse are more likely to have higher *Social Integration* score.

Taken all together, our results indicate that if the team has *Activated Faultlines* then more *Conflicts* are likely to happen. This leads to lower team outcomes: worse *Team Performance* (graded externally by the manager and self-perceived) and *Social Integration* (satisfaction, cohesion and intention to remain). We also found that *Activated Faultlines* have a direct influence on team outcomes, so that if the group is divided into smaller subgroups its performance and cohesiveness are likely to diminish. Interestingly, the strength of *Dormant Faultlines* was not found to have any effect on *Activated Faultlines*, so those teams where we expected to see subgroups based on their demographic composition actually did not have them, and those with weak chance to break into subgroups did. *Dormant Faultlines* did not influence team outcomes, thus, by blindly calculating the chance of particular team to be broken in smaller subgroups based on age, gender, nationality/origins of its members we cannot predict how this team will perform or how satisfied its employees will be. Figures 2 and 3 summarize the findings of the mediating relationships explored in the study.

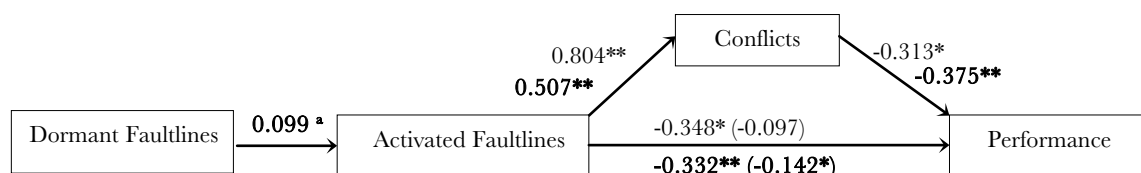


Figure 2. Effect of Dormant Faultlines on Activated Faultlines and significant direct and indirect effects of Activated Faultlines on Performance

Notes: Unstandardized regression coefficients are reported;

* $p < 0.05$, ** $p < 0.01$, two-tailed;

First number – score at the group-level (Manager), the second number (in bold) - score at the individual-level (Team members); Coefficients in parentheses represent the effect after the introduction of Conflicts into the regression equations.

^a Unstandardized regression coefficient for relationship between Dormant Faultline strength calculated based on Gender, Age and Country of Origin and averaged Activated Faultlines is provided as an example; direct effect of Dormant Faultlines on Activated Faultlines was not found (see Table V).

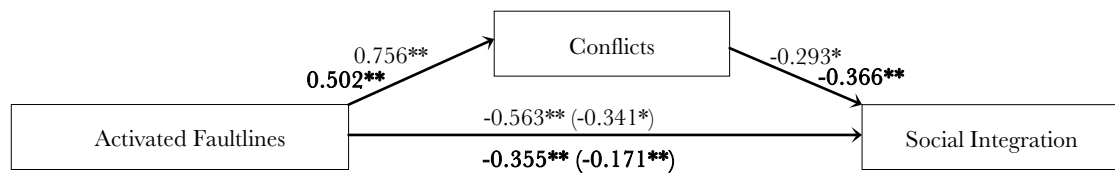


Figure 3. Significant direct and indirect effects of Activated Faultlines on Social Integration

Notes: Unstandardized regression coefficients are reported;

* $p < 0.05$, ** $p < 0.01$, two-tailed;

First number – score at the group-level (Average), the second number (in bold) - score at the individual-level (Team members); Coefficients in parentheses represent the effect after the introduction of Conflicts into the regression equations.

DISCUSSION

Previous research on organizational diversity has focused on quantifying it by inventing a diversity (e.g. faultlines) indexes. We followed the theory of faultlines and measured probability of the team to break into subgroups based on genders, ages, nationalities and origins of its members. Drawing on previous research, we argued that teams that have stronger dormant faultlines would more likely to have subgroups based on demographic differences. If subgroups would be evident, we expected that more conflicts will happen between team members due to gossips and misunderstanding between the subgroups. Eventually, that would lead to lower team performance and social integration (satisfaction, intention to remain and cohesion). While we found strong support for the fact that if the group is broken into multiple subgroups it leads to conflicts and worsens team outcomes, we did not find any proof for the fact that dormant faultlines lead to activated ones. Thus, our results showed that preventing team members from being divided into multiple subgroups would be most relevant for improvement of low-skilled team outcomes. However, subgrouping process cannot be predicted based solely on information about a team's demographic composition.

Theoretical Contributions

Our study offers several implications for the faultlines and diversity literature. Up until now, diversity researchers have primarily focused on team performance and innovativeness in R&D and management teams. However, as employees continue to interact with each other, deep-level diversity (e.g. attitudinal or educational dissimilarities) may over time become more important than surface-level differences (Harrison et al., 1998). This claim specifically applies to teams comprised of the higher-educated employees as they would more likely try to uncover deeper dissimilarities in each other. Therefore, conducting research in low-skilled teams helps to neutralize the effect of deep-level diversity attributes.

Firstly, in a sample composed of cleaning teams, we found that dormant faultlines based on surface-level differences were not able to predict if the group would actually have faultlines activated. Therefore, in low-skilled teams, its members learn deeper-level information about each other and demographic similarity does not create a sufficient base for forming subgroups. In addition, we expected that dormant faultlines would negatively affect team performance and social integration but these hypotheses were not confirmed. So if only team's demographic composition is known, we cannot make any predictions about group outcomes. That means that the relationship between group composition and team outcomes is less straight-forward than researchers originally thought.

Secondly, even though dormant demographic faultlines do not necessarily cause the team members to perceive subgroups, results were largely supportive of the rest of our hypotheses. We found that activated surface-level faultlines negatively affected team outcomes that is consistent with findings of research on group process (Carton & Cummings, 2013; Cronin et al., 2011; Jehn & Bezrukova, 2010). It seems like if team members recognize that their group is divided into subgroups, they feel less satisfied and committed to their team. One explanation may be that if team members perceive having subgroups, it means that they have less contact with one part of their colleagues. Therefore, our research provides strong evidence that multiple subgroups within a team signal about problems in relationships between colleagues that negatively affect team performance.

Thirdly, our study contributes to the conflict literature and provides an understanding of how the influence of evident subgroups is mediated by conflicts in the team. Specifically, our theory and findings show that activated faultlines increase the level of conflicts in the team, which then worsens team performance and social integration. This is in line with the social-categorization theory (Jackson & Joshi, 2011; Van Knippenberg et al., 2004), which states that subgroups create negative bias towards out-group members that influences subsequent performance (Homan et al., 2008). In our study, this mediation effect was found based on the perception of team members and based on the managers' opinions. We think that when the team is clearly divided into subgroups, the quality of communication is likely to reduce, thus conflicts in the team appear more often. The mediating role of conflicts was more evident for team performance than for social integration, so if faultlines are activated in the team, performance will drop due to an increased level of conflicts.

In addition, our considerable contribution to the faultlines literature is to respond to calls for the conducting research in real work teams (Thatcher & Patel, 2012). Most of the previous research was conducted in artificially made student teams or in experiment settings. This research was conducted in a company with the highly diverse workforce. Very few studies were conducted in the field settings as controlling for the strength of the faultlines becomes difficult. Furthermore, in lab settings researchers tend to establish a stronger negative relationship between faultlines and team performance (Thatcher & Patel, 2012). We were able to obtain a sample of teams with different levels of dormant faultlines. Based on our observations of 32 teams, we state that while many researches tried to find the level of

faultline that is optimal for the team functioning, in practice it still would be not possible to shape teams based on some faultlines' strength threshold. Similar conclusions can be found in the research of Cronin et al. (2011) and Harrison and Klein (2007), where it is stated that if people work longer together it is difficult to find a relationship between demographic predictors and perceived subgroups or performance. In our sample, around 80 per cent of employees work two years or longer in the same team. Thus, even in low-skilled teams, demographic differences are likely to become not important by lapse of time.

Our findings clearly indicate that even demographic differences are not that easy to capture in one index (leaving aside deep-level differences). In our case, we calculated faultlines based on both nationality and country of origin. This already may give two different values for the same team: while all employees may have Dutch nationality (dormant faultline is 0), half of them may be born in Portugal (dormant faultline becomes 1). We have met some employees that were born in the Netherlands and had Dutch nationality but speak another language at home and choose friends from another culture. In contrast, we also have met employees without Dutch nationality who have lived less than five years in the Netherlands but at the same time have spoken Dutch like native speakers and have not explicitly searched contact with people from their home country. From this reasoning we draw the inference that not only current literature does not provide guidelines on which variables to choose in order to assess the level of diversity in the team, but also taking into account all the imaginable differences of employees (like political preferences, sexual orientation, religion, work mentality etc.) in order to build a team with particular level of innovativeness or productivity, seems unrealizable.

In sum, our main contribution is in demonstrating that it is impossible to capture diversity or faultline strength in the index and based on that form work teams or try to predict team outcomes. However, obviously, diversity exists as otherwise, subgroup formation would not take a place. Eventually, the most important is to explore whether the team is divided into subgroups as that provides valuable information about team process (conflicts) and outcomes (performance and social integration).

Practical Implications

The results of our study have implications for diversity management. Nationality, age and other forms of team diversity are growing rife in many companies. It especially applies to organizations using low-skilled labour where foreigners often get their first jobs in a new country. Therefore, exploring both beneficial and detrimental consequences of diversity on team process and performance becomes a valuable part for the effective management (Van Knippenberg et al., 2004).

For HR/Diversity Departments. While earlier studies conclude that during the team design stage managers should make their choices based on arrangement of demographic attributes such as

nationality, age and gender (Bezrukova, Jehn, Zanutto, & Thatcher, 2009; Carton & Cummings, 2013; Lau & Murnighan, 1998), we did not find any support for this claim. On the contrary, we have found that in low-skilled teams with high chance to form subgroups based on demographic differences, it did not necessarily happen. Thus, we do not find it realistic to state that managers can prevent subgroups' formation based on certain demographic attributes while establishing the team. If some predisposition for subgroup formation is evident, then attention should be given to prevent that. For instance, four Turkish old women and four young Dutch women work in the same team. It is likely that the two cliques will be formed, but it would be rash to assert that this scenario will definitely happen. Their manager needs to know the language proficiency of the Turkish women, background information for all employees (e.g. some of the Dutch women may have Turkish relatives), and deep-level diversity attributes (e.g. religion) as all of those characteristics can play a role in activation of dormant faultlines. Therefore, because so many details should be controlled, we suggest not trying to prevent subgroup formation based solely on demographic information of applicants during the team design stage.

Another argument for not using indexes for diversity comes from the observation that it is almost impossible to quantify it. For example, calculation of the faultline strength for the same team based on members' nationalities and later based on their origins may give two completely different values. Consequently, we suggest that managers individually work with every team and try to prevent subgrouping based on any possible difference by promoting cohesion, helping and friendliness in their team. In case faultlines seem already be activated, the manager should immediately start figuring out the cause and then make future decisions about team composition and needed actions.

After that, HR department of a company with a diverse workforce has to ensure that line managers are educated in areas of diversity and conflict management as only they are able to monitor group dynamics on the daily basis. It is important to recognize the beginning of the subgrouping process in order to take an action. Moreover, it is critical to take a right action because for some cultures open group discussions would help, while for others individual discussions would be more appropriate. Thus, the team manager has to be aware that avoiding subgroups is his or her responsibility and that using all available tools in time would help to achieve group cohesion and understanding.

Moreover, our results show that if subgroups were observable in the team, it was indicated by both employees and the manager. We advice to annually monitor whether employees or managers indicate having subgroups at their workplace. This can be done by including additional items in the employees' satisfaction survey that is usually used by most of the companies. Visiting company's units may also provide valuable information as observation of communication in the team during lunchtime may already provide a rough idea of group dynamics. For example, if diversity manager sees that Dutch employees are having their lunch at one side of the table while Turkish employees are sitting on the other side, it will be a clear sign that action has to be taken by the line manager.

Lastly, it is important to foster a diversity-friendly atmosphere in the company. Every team may have an event when employees cook their national dishes and wear their traditional clothes. Benefits of diversity may be constantly reported via organizational newspaper or emails with examples of success stories of colleagues from the different countries.

For Line Management. In addition to the recommendations given above, we developed guidelines for team managers who directly see their teams on the daily basis. First of all, the line manager should know that having subgroups in the teams brings conflicts and worsens performance. For instance, Dutch employees meet at 9 a.m. while Polish employees come half an hour later. Obviously, those two subgroups barely see each other, which in turn diminishes cohesion. If on one occasion Dutch employees need extra help, it is very likely that Polish employees will not be ready to work extra for their colleagues without having additional compensation. Thus, it is highly important to create team feeling and emphasise that the job is done by the joint effort of all team members.

Second, line managers should avoid triggers indicated by Chrobot-Mason et al. (2009). For example, giving to younger employees more difficult work than to older colleagues may activate age faultline, or allowing Turkish employees stop working 10 minutes earlier while never allowing it to Spanish ones may activate nationality faultline as different treatments would be immediately perceived by both subgroups. The team manager should treat all the employees equally and in case that some distinction has to be made, the reason for that action should be clearly reported and explained. In addition, the main language should be spoken and line manager should try to prevent conversations in different languages if all team is gathered together.

Thirdly, team manager should take care that employees have enough time for conversations with each other, for example, during the lunch. Once a week longer lunch may be organized when all team news and ideas can be discussed. It would show that the line manager is open to new propositions and is interested that everyone in the team feels well. We have seen that some successful line managers organize bus trips to another city so that employees were able to communicate informally and see their colleagues from a new perspective. This stimulates employees to avoid stereotypical ideas and be more open to all the colleagues.

Lastly, if the faultlines are activated and the line manager is convinced that it is not possible to break existing subgroups, this should be reported to the HR department. The solution may be to relocate employees or to propose additional language courses if some colleagues have difficulties in communication with others. We believe that all conflicts and bad group dynamics may be improved with correct management actions taken in time.

During our data collection, we have noticed that those managers, who emphasize the value of diversity in the team, prevent team members from making stereotypes about each other and preclude conversations in another language than Dutch while the team is together, reported fewer subgroups and conflicts in their teams. The same observation was found by Spoelma and Ellis (2017) who write

that if leaders encourage their team members to enhance their diversity beliefs and establish shared norms, it will increase team cohesion. Hence, managers of diverse teams should pay enough attention to honest and respectful interactions with employees in order to recognize whether subgroup formation is possible and on time start to prevent it. If subgroups are already formed, it has a strong negative effect on team functioning as more conflicts arise, that leads to lower performance and social integration (satisfaction, cohesion and intention to remain) of the team.

Limitations and Future Research

Despite these contributions, our study has limitations that may provide direction for future research. We used data collected from 300 individuals from 32 teams. Despite the strength that we gathered full demographic information for every team and from each more than 75 per cent of members have participated in this research, the number of groups in the sample is still rather small. Although having a bigger sample by surveying more teams would have been desirable, we think that our findings already provide useful information about the applicability of faultline strength and diversity indexes. Future research may replicate the study using a larger sample.

As with all studies, unmeasured variables may affect the findings. We think that more data about team managers' diversity attitudes or involvement in stimulating positive team dynamics would provide valuable information about subgrouping process. Yet, we have no reason to think that this data would affect our main findings. We also did not include deep-level diversity variables like members' attitudes towards the organization, heterogeneity in educational backgrounds, job values and other not easily observable diversity attributes. Even though we state that it is not possible to quantify those differences in one diversity value, we advise to explore deep-level diversity in the future research as demographic attributes do not seem to significantly predict team success. In addition, we found that conflict only partially mediates the relationship between activated faultlines and team outcomes, thus future research should focus on finding new possible mediators that would provide more explanation on how subgroups influence team performance and social integration.

Our data collection had several limitations. First of all, all the data was collected from teams that were willing to participate. Secondly, our sample contains teams that are not diverse in their work tenure as more than 78 per cent of employees work longer than two years in their teams. What is more, it would be interesting to focus on the new teams where the effect of demographic differences may be more observable and then collect the same data two years later. Measuring these variables over time would help to assess the dynamic processes of faultline activation and to determine their impact on performance. Moreover, it would be useful to conduct more case research in order to reveal information about subgroup formation in teams with different faultlines' strengths.

Thirdly, all teams in our sample perform cleaning work and even though employees have to make joint decisions and solve rising problems together, their task interdependence could be higher as

cleaning is still mostly done individually. Future research may try to obtain the sample of low-skilled teams where more collective decision-making and cooperation take place.

Like most survey-based studies, our study has some method-based limitations. As our research took part in a cleaning organization, we had to adapt some of the original measurement items to make them more understandable for low-skilled workers. But at the same time, the employees' levels of understanding can still vary, so it could be that native speakers could provide more thoughtful answers than non-native respondents. For the future research, we may advise translating the questionnaire into the language of respondents if they constitute more than 10 per cent of the sample. Like in our case we could translate the questionnaire in Turkish and Arabic. In addition, all the variables were measured with self-report instruments. The use of other sources and other types of measurements would have strengthened the reliability and validity of the study.

Another interesting research avenue would be to explore if members of different subgroups perceive team performance and social integration differently. While members of one subgroup perceive high performance, another subgroup may have a contrary opinion. As we found that activated faultlines are a strong predictor for the team outcomes, we encourage more future research on studying dynamics at subgroup and team-levels at the same moment.

Our analysis also has several limitations. Our dependent variables regarding social integration captured employees' perceptions and did not capture the objective cohesiveness of the team. While we were able to obtain managers' opinions about conflicts and performance of the teams, future research should employ more objective performance measures. One more limitation is the value of ICC(2) as they are at the middle level and Buengeler and den Hartog (2015, p. 848) note that "results obtained based on variables with modest ICC(2) values represent rather conservative tests of the hypothesized effects". Future research may try to sample teams of a larger size in order to increase the value of ICC(2).

CONCLUSION

In conclusion, the present study suggests that diversity in teams is a very complex phenomenon. Our findings indicate that when members and managers of the team perceive the group be divided into smaller subgroups, it has a strong negative effect on team processes as more conflicts arise and subsequently deteriorate team outcomes. At the same time, dormant faultlines' strength does not have any effect on faultlines' activation or team outcomes. Furthermore, we propose that future research begins to refocus from trying to find a specific measure that captures demographic or overall team diversity to exploring the process of faultline activation in more detail. In sum, the findings of this study underline the importance of the effect of subgroups on teams' performance and social integration.

REFERENCES

- Alexander, J., Nuchols, B., Bloom, J., & Lee, S.-Y. (1995). Organizational demography and turnover: An examination of multiform and nonlinear heterogeneity. *Human Relations, 48*(12), 1455-1480. doi:10.1177/001872679504801204
- Alipour, K. K., Mohammed, S., & Raghuram, S. (2018). Differences in the Valuing of Power Among Team Members: a Contingency Approach Toward Examining the Effects of Power Values Diversity and Relationship Conflict. *Journal of Business and Psychology, 33*(2), 231-247.
- Bacharach, M. (2005). Foreword: teamwork. In *Gold, Teamwork—Multi-Disciplinary Perspectives* (pp. xxi-xxv). New York: Palgrave Macmillan.
- Barkema, H. G., & Shvyrkov, O. (2007). Does Top Management Team Diversity Promote or Hamper Foreign Expansion? *Strategic Management Journal, 28*(7), 663-680.
- Bernardin, H. J. (2002). *Human resource management: An experiential approach*: Irwin Professional Pub.
- Bezrukova, K., Jehn, K. A., Zanutto, E. L., & Thatcher, S. M. (2009). Do workgroup faultlines help or hurt? A moderated model of faultlines, team identification, and group performance. *Organization Science, 20*(1), 35-50. doi:10.1287/orsc.1080.0379
- Bezrukova, K., Spell, C. S., Caldwell, D., & Burger, J. M. (2016). A multilevel perspective on faultlines: Differentiating the effects between group- and organizational-level faultlines. *The Journal of applied psychology, 101*(1), 86-107. doi:10.1037/apl0000039
- Bezrukova, K., Spell, C. S., & Perry, J. L. (2010). Violent splits or healthy divides? Coping with injustice through faultlines. *Personnel Psychology, 63*(3), 719-751. doi:10.1111/j.1744-6570.2010.01185.x
- Bezrukova, K., Thatcher, S. M., Jehn, K. A., & Spell, C. S. (2012). The effects of alignments: examining group faultlines, organizational cultures, and performance. *The Journal of applied psychology, 97*(1), 77-92. doi: 10.1037/a0023684
- Biemann, T., Cole, M. S., & Voelpel, S. (2012). Within-group agreement: On the use (and misuse) of rWG and rWG (J) in leadership research and some best practice guidelines. *The Leadership Quarterly, 23*(1), 66-80. doi:10.1016/j.leaqua.2011.11.006
- Blau, P. M. (1977). *Inequality and heterogeneity : a primitive theory of social structure*. New York :: Free Press.
- Bliese, P. D. (2000). Within-group agreement, non-independence, and reliability: Implications for data aggregation and analysis. In *Multilevel theory, research, and methods in organizations: Foundations, extensions, and new directions*. (pp. 349-381). San Francisco, CA, US: Jossey-Bass.
- Buengeler, C., & den Hartog, D. N. (2015). National diversity and team performance: the moderating role of interactional justice climate. *International Journal of Human Resource Management, 26*(6), 831-855. doi:10.1080/09585192.2014.991345
- Byrne, D. E. (1971). *The attraction paradigm* (Vol. 11): Academic Pr.

- Carless, S. A., Fewings-Hall, S., Hall, M., Hay, M., Hemsworth, P. H., & Coleman, G. J. (2007). Selecting Unskilled and Semi-Skilled Blue-Collar Workers: The criterion-related validity of the PDI-employment inventory 1. *International Journal of Selection and Assessment*, *15*(3), 335-340.
- Carton, A. M., & Cummings, J. N. (2013). The impact of subgroup type and subgroup configurational properties on work team performance. *The Journal of applied psychology*, *98*(5), 732-758. doi:10.1037/a0033593
- Choi, J. N., & Sy, T. (2010). Group-level organizational citizenship behavior: Effects of demographic faultlines and conflict in small work groups. *Journal of Organizational Behavior*, *31*(7), 1032-1054.
- Chrobot-Mason, D., Ruderman, M. N., Weber, T. J., & Ernst, C. (2009). The challenge of leading on unstable ground: Triggers that activate social identity faultlines. *Human Relations*, *62*(11), 1763-1794. doi:10.1177/0018726709346376
- Cohen, S. G., & Bailey, D. E. (1997). What makes teams work: Group effectiveness research from the shop floor to the executive suite. *Journal of Management*, *23*(3), 239-290. doi:https://doi.org/10.1016/S0149-2063(97)90034-9
- Cronin, M. A., Bezrukova, K., Weingart, L. R., & Tinsley, C. H. (2011). Subgroups within a team: The role of cognitive and affective integration. *Journal of Organizational Behavior*, *32*(6), 831-849. doi:10.1002/job.707
- Cunningham, G. B. (2009). The Moderating Effect of Diversity Strategy on the Relationship Between Racial Diversity and Organizational Performance. *Journal of Applied Social Psychology*, *39*(6), 1445-1460.
- Geological fault. (2003-2008). *WordNet 3.0, Farlex clipart collection*. Retrieved from <https://www.thefreedictionary.com/geological+fault>
- Gibson, C., & Vermeulen, F. (2003). A Healthy Divide: Subgroups as a Stimulus for Team Learning Behavior. *Administrative science quarterly*, *48*(2), 202-239. doi:10.2307/3556657
- Glenn, E. N. (1990). Cleaning up/kept down: A historical perspective on racial inequality in women's work. *Stan. L. Rev.*, *43*, 1333.
- Guillaume, Y. R. F., Dawson, J. F., Otaye-Ebede, L., Woods, S. A., & West, M. A. (2017). Harnessing demographic differences in organizations: What moderates the effects of workplace diversity? *Journal of Organizational Behavior*, *38*(2), 276-303. doi:10.1002/job.2040
- Harrison, D. A., & Klein, K. J. (2007). What's the Difference? Diversity Constructs as Separation, Variety, or Disparity in Organizations. *The Academy of Management Review*, *32*(4), 1199-1228. doi:10.2307/20159363
- Harrison, D. A., Price, K. H., & Bell, M. P. (1998). Beyond Relational Demography: Time and the Effects of Surface- and Deep-Level Diversity on Work Group Cohesion. *Academy of Management Journal*, *41*(1), 96-107. doi:10.5465/256901

- Homan, A. C., Hollenbeck, J. R., Humphrey, S. E., Knippenberg, v. D., Ilgen, D. R., & Kleef, v. G. A. (2008). Facing differences with an open mind: Openness to experience, salience of intra-group differences, and performance of diverse work groups. *Academy of Management Journal*, 51(6), 1204-1222. doi:10.5465/AMJ.2008.35732995
- Jackson, S. E., & Joshi, A. (2011). Work team diversity. *APA handbook of industrial and organizational psychology, 1*, 651-686. doi:10.1037/12169-020
- Jackson, S. E., Joshi, A., & Erhardt, N. L. (2003). Recent Research on Team and Organizational Diversity: SWOT Analysis and Implications. *Journal of Management*, 29(6), 801-830. doi:10.1016/S0149-2063(03)00080-1
- James, L. R., Demaree, R. G., & Wolf, G. (1993). rwg: An assessment of within-group interrater agreement. *Journal of applied psychology*, 78(2), 306-309.
- Jehn, K. A. (1995). A Multimethod Examination of the Benefits and Detriments of Intragroup Conflict. *Administrative science quarterly*, 40(2), 256. doi:10.2307/2393638
- Jehn, K. A., & Bezrukova, K. (2010). The faultline activation process and the effects of activated faultlines on coalition formation, conflict, and group outcomes. *Organizational Behavior and Human Decision Processes*, 112(1), 24-42. doi:10.1016/j.obhdp.2009.11.008
- Jehn, K. A., Northcraft, G. B., & Neale, M. A. (1999). Why differences make a difference: A field study of diversity, conflict and performance in workgroups. *Administrative science quarterly*, 44(4), 741-763.
- Joshi, A., & Roh, H. (2009). The role of context in work team diversity research: a meta-analytic review. *The Academy of Management Journal*, 52(3), 599-627. doi:10.5465/AMJ.2009.41331491
- Kraut, A. I. (1975). Predicting turnover of employees from measured job attitudes. *Organizational Behavior and Human Performance*, 13(2), 233-243. doi:https://doi.org/10.1016/0030-5073(75)90047-1
- Krogel, J. (2008). *The Group Questionnaire: A new measure of the group relationship*: Brigham Young University.
- Lau, D. C., & Murnighan, J. K. (1998). Demographic diversity and faultlines: The compositional dynamics of organizational groups. *Academy of Management Review*, 23(2), 325-340.
- Lau, D. C., & Murnighan, J. K. (2005). Interactions within groups and subgroups: the effects of demographic faultlines. *The Academy of Management Journal*, 48(4), 645-659. doi:10.5465/AMJ.2005.17843943
- LeBreton, J. M., & Senter, J. L. (2008). Answers to 20 Questions About Interrater Reliability and Interrater Agreement. *Organizational Research Methods*, 11(4), 815-852. doi:10.1177/1094428106296642

- Li, J., & Hambrick, D. C. (2005). Factional Groups: A New Vantage on Demographic Faultlines, Conflict, and Disintegration in Work Teams. *The Academy of Management Journal*, *48*(5), 794-813. doi:10.5465/AMJ.2005.18803923
- Maxwell, N. L. (2006). Low-skilled jobs: The reality behind the popular perceptions. doi:https://doi.org/10.17848/9781429454902
- Meyer, B., & Glenz, A. (2013). Team faultline measures: A computational comparison and a new approach to multiple subgroups. *Organizational Research Methods*, *16*(3), 393-424. doi:10.1177/1094428113484970
- Meyer, B., Schermuly, C. C., & Kauffeld, S. (2016). That's not my place: The interacting effects of faultlines, subgroup size, and social competence on social loafing behaviour in work groups. *European Journal of Work and Organizational Psychology*, *25*(1), 31-49. doi:10.1080/1359432X.2014.996554
- Meyer, B., Shemla, M., Li, J., & Wegge, J. r. (2015). On the Same Side of the Faultline: Inclusion in the Leader's Subgroup and Employee Performance. *Journal of Management Studies*, *52*(3), 354-380. doi:10.1111/joms.12118
- Minichilli, A., Corbetta, G., & MacMillan, I. C. (2010). Top management teams in family-controlled companies: 'familiness', 'faultlines', and their impact on financial performance. *Journal of Management Studies*, *47*(2), 205-222.
- Molleman, E. (2005). Diversity in Demographic Characteristics, Abilities and Personality Traits: Do Faultlines Affect Team Functioning? *Group Decision and Negotiation*, *14*(3), 173-193. doi:10.1007/s10726-005-6490-7
- Mortensen, M., & Haas, M. R. (2018). Perspective—Rethinking Teams: From Bounded Membership to Dynamic Participation. *Organization Science*, *29*(2), 341-355.
- Pelled, L. H., Eisenhardt, K. M., & Xin, K. R. (1999). ARTICLES - Exploring the Black Box: An Analysis of Work Group Diversity, Conflict and Performance. *Administrative science quarterly*, *44*(1), 1. doi:10.2307/2667029
- Pesch, R., & Bouncken, R. B. (2018). How to achieve benefits from diversity in international alliances: Mechanisms and cultural intelligence. *Global Strategy Journal*, *8*(2), 275-300.
- Pfeffer, J. (1985). Organizational demography: Implications for management. *California Management Review*, *28*(1), 67-81.
- Rico, R., Molleman, E., Sánchez-Manzanares, M., & Van der Vegt, G. S. (2007). The effects of diversity faultlines and team task autonomy on decision quality and social integration. *Journal of Management*, *33*(1), 111-132. doi:10.1177/0149206306295307
- Sawyer, J. E., Houlette, M. A., & Yeagley, E. L. (2006). Decision performance and diversity structure: Comparing faultlines in convergent, crosscut, and racially homogeneous groups. *Organizational Behavior and Human Decision Processes*, *99*(1), 1-15. doi:10.1016/j.obhdp.2005.08.006

- Schneider, B. (1987). The people make the place. *Personnel Psychology*, *40*(3), 437-453.
- Schölmerich, F., Schermuly, C. C., & Deller, J. r. (2016). How Leaders' Diversity Beliefs Alter the Impact of Faultlines on Team Functioning. *Small Group Research*, *47*(2), 177-206. doi:10.1177/1046496416630960
- Spell, C. S., Bezrukova, K., Haar, J., & Spell, C. (2011). Faultlines, fairness, and fighting: A justice perspective on conflict in diverse groups. *Small Group Research*, *42*(3), 309-340. doi:10.1177/1046496411402359
- Spoelma, T. M., & Ellis, A. P. J. (2017). Fuse or fracture? Threat as a moderator of the effects of diversity faultlines in teams. *The Journal of applied psychology*, *102*(9), 1344-1359. doi:10.1037/apl0000231
- Tajfel, H., & Turner, J. (1986). The social identity theory of inter group behavior in S Worchel & WG Austin (Eds) *Psychology of intergroup relations*. Chicago: Nelson.
- Thatcher, S. M., Jehn, K. A., & Zanutto, E. (2003). Cracks in diversity research: The effects of diversity faultlines on conflict and performance. *Group Decision and Negotiation*, *12*(3), 217-241.
- Thatcher, S. M., & Patel, P. C. (2012). Group faultlines: A review, integration, and guide to future research. *Journal of Management*, *38*(4), 969-1009. doi:10.1177/0149206311426187
- Turner, J. C. (1982). Towards a cognitive redefinition of the social group. *Social identity and intergroup relations*, 15-40.
- Turner, J. C., & Reynolds, K. J. (2011). Self-categorization theory. *Handbook of theories in social psychology*, *2*(1), 399-417.
- van Knippenberg, D., Dawson, J. F., West, M. A., & Homan, A. C. (2011). Diversity faultlines, shared objectives, and top management team performance. *Human Relations*, *64*(3), 307-336. doi:10.1177/0018726710378384
- Van Knippenberg, D., De Dreu, C. K., & Homan, A. C. (2004). Work group diversity and group performance: an integrative model and research agenda. *The Journal of applied psychology*, *89*(6), 1008-1022. doi:10.1037/0021-9010.89.6.1008
- Van Knippenberg, D., & Schippers, M. C. (2007). Work group diversity. *Annual Review of Psychology*, *58*, 515-541. doi:10.1146/annurev.psych.58.110405.085546
- Van Knippenberg, D., & Van Ginkel, W. (2010). The Categorization-Elaboration Model of Work Group Diversity: Wielding the Double-Edged Sword. In.
- Vermeeren, B. (2014). Variability in HRM implementation among line managers and its effect on performance: a 2-1-2 mediational multilevel approach. *The International Journal of Human Resource Management*, *25*(22), 3039-3059. doi:10.1080/09585192.2014.934891
- Wageman, R., Hackman, J. R., & Lehman, E. (2005). Team diagnostic survey: Development of an instrument. *The Journal of Applied Behavioral Science*, *41*(4), 373-398. doi:10.1177/0021886305281984

- Webber, S. S., & Donahue, L. M. (2001). Impact of highly and less job-related diversity on work group cohesion and performance: a meta-analysis. *Journal of Management*, 27(2), 141-162. doi:10.1016/S0149-2063(00)00093-3
- Williams, K. Y., & O'Reilly, C. A. (1998). Demography and diversity in organizations: a review of 40 years of research. *Research in organizational behavior*, 20, 77-140.
- Zanutto, E. L., Bezrukova, K., & Jehn, K. A. (2011). Revisiting faultline conceptualization: Measuring faultline strength and distance. *Quality & Quantity*, 45(3), 701-714. doi:10.1007/s11135-009-9299-7

Appendix A

Survey Items for team members

	Construct	Question in Dutch	Question in English	Note	Reference
1	Social Integration: Satisfaction	De omgang met andere teamleden is ongemakkelijk	My relations with other team members are strained (uneasy)	<i>Reversed</i>	Wageman, R., Hackman, J. R., & Lehman, E. (2005). Team diagnostic survey: Development of an instrument. <i>The Journal of Applied Behavioral Science</i> , 41(4), 373-398. doi:10.1177/0021886305281984
2		Ik vind het werken met mijn teamgenoten plezierig	I enjoy working with my teammates		
3		Over het algemeen ben ik tevreden over dit team	Generally speaking, I am satisfied with this team		
4	Social Integration: Cohesion	Ik voel dat ik bij het team hoor	I feel accepted by this team		Kroegel, J. (2008). <i>The Group Questionnaire: A new measure of the group relationship</i> : Brigham Young University.
5		Ik vind het vervelend dat ik deel uitmaak van dit team	I feel bad about being a part of this team	<i>Reversed</i>	
6		Ook al zijn anderen het soms met mij oneens, ik voel me geaccepteerd in mijn team	Even though others may disagree with me sometimes, I feel accepted in this team		
7	Social Integration: Intention to remain	Ik wil graag lang in dit team blijven	I would like to stay long in this team		Kraut, A. I. (1975). Predicting turnover of employees from measured job attitudes. <i>Organizational Behavior and Human Performance</i> , 13(2), 233-243. doi:https://doi.org/10.1016/0030-5073(75)90047-1 (
8		Als ik zelf zou kunnen kiezen, zou ik over drie jaar nog steeds in hetzelfde team werken	If I have my own way, I will be working in this same team three years from now		
9		Ik wil liever bij een ander team gaan werken	I would like to change this team and work in another one	<i>Reversed</i>	
10	Activation of surface-level diversity faultlines	In mijn team hebben vrouwen minder contact met mannen	In my team women have less contact with men		Jehn, K. A., & Bezrukova, K. (2010). The faultline activation process and the effects of activated faultlines on coalition formation, conflict, and group outcomes. <i>Organizational Behavior and Human Decision Processes</i> , 112(1), 24-42. doi:10.1016/j.obhdp.2009.11.008 Self – made questions
11		In mijn team hebben jongere mensen minder contact met oudere mensen	In my team young people have less contact with older people		
12		In mijn team communiceren mensen meestal met collega's van dezelfde nationaliteit	In my team people mostly communicate with colleagues of the same nationality		
13		In mijn team hebben degenen die langer werken minder contact met nieuwe werknemers	In my team those who work longer have less contact with new employees		
14		Mijn team is verdeeld in kliekjes (kleine groepjes)	My team is divided into smaller cliques (small groups)		
15	Conflicts: Task conflict	In mijn team zijn we het vaak oneens met elkaar over het werk	In my team we often disagree about opinions regarding the work being done		Jehn, K. A. (1995). A Multimethod Examination of the Benefits and Detriments of Intragroup Conflict. <i>Administrative science quarterly</i> , 40(2), 256. doi:10.2307/2393638
16		Er zijn conflicten over de volgorde van de taken in mijn team	There are conflicts in my team about the order of the tasks		

17		Mijn team heeft geen conflicten over de manier waarop we onze taken moeten uitvoeren	My team does not have conflicts about the way we should perform our tasks	<i>Reversed</i>	
18	Conflicts: Relationship conflict	Er is onenigheid tussen collega's in mijn team	There is much friction (disagreement) among members of my team		Jehn, K. A., Northcraft, G. B., & Neale, M. A. (1999). Why differences make a difference: A field study of diversity, conflict and performance in workgroups. <i>Administrative science quarterly</i> , 44(4), 741-763.
19		Leden in mijn team zijn onredelijk tegen elkaar	Members of my team are unfair to each other		
20		Er zijn geen spanningen tussen de leden van mijn team	There is no tension among members of my team	<i>Reversed</i>	
21		Er wordt geroddeld in mijn team	There are gossips in my team		
22	Team Performance	Mijn team presteert goed	My team performs well		Jehn, K. A., Northcraft, G. B., & Neale, M. A. (1999). Why differences make a difference: A field study of diversity, conflict and performance in workgroups. <i>Administrative science quarterly</i> , 44(4), 741-763.
23		Mijn team krijgt dingen snel en op de juiste manier gedaan	My team is effective at getting things done quickly		
All questions below are self-compiled					
24	Self-perceived diversity	Mijn team bestaat uit veel verschillende mensen (in termen van nationaliteit, geslacht en leeftijd)	My team has many diverse people (in terms of nationality, gender and age)		
25	Demographic information	Hoe oud bent u?	How old are you?		
26		Wat is uw geslacht?	What is your gender?		
27		Wat is uw nationaliteit?	What is your nationality?		
28		Wat is uw geboorteland?	What is your country of birth?		
29	Control for the language proficiency	Wat is uw moedertaal?	What is your first language?		
30		Welke taal gebruikt uw meest op uw werk?	Which language do you mostly speak at your work?		
31		Op welke niveau spreekt u Nederlands? <input type="checkbox"/> Ik spreek en begrijp geen Nederlands <input type="checkbox"/> Beginner - Ik ken basiswoorden en kan eenvoudige zinnen begrijpen <input type="checkbox"/> Gemiddeld - Ik heb geen problemen om over alledaagse onderwerpen te praten <input type="checkbox"/> Geavanceerd / Native - Ik kan alles gemakkelijk begrijpen en zeggen	Which level of Dutch (Nederlands) do you have: <input type="checkbox"/> I do not speak and understand Dutch <input type="checkbox"/> Beginner – I know basic words and am able to understand simple sentences <input type="checkbox"/> Intermediate – I have no difficulties in speaking to any everyday topic <input type="checkbox"/> Advanced/Native – I can easily understand and say everything		
32	Control question for the team size	Hoeveel mensen zitten er in uw team?	How many people are there in your team?		
33	Work and Team Tenure	Sinds wanneer werkt u al bij Asito?	When did you start to work at Asito?		
34		Sinds wanneer bent u onderdeel van dit team?	When did you start to work in this team?		
35	Control question for tenure in the team	Hoeveel uren per week werkt u gemiddeld in dit team?	How many hours per week on average do you work in this team?		

Appendix B

Survey items for team managers

	Construct	Question ^a	Reference
1	Team Performance	Dit team presteert (relatief) goed op productiviteit	Vermeeren, B. (2014). Variability in HRM implementation among line managers and its effect on performance: a 2-1-2 mediational multilevel approach. <i>The International Journal of Human Resource Management</i> , 25(22), 3039-3059. doi:10.1080/09585192.2014.934891 Bernardin, H. J. (2002). <i>Human resource management: An experiential approach</i> . Irwin Professional Pub.
2		De kwaliteit van schoonmaken van dit team is hoog	
3		Het aantal schoonmaak taken dat dit team verleent is hoog	
4		Dit team behaalt zijn schoonmaak taken binnen de daarvoor gestelde tijd	
5		Dit team behaalt de gestelde doelstellingen	
6	Conflicts	Er zijn emotionele spanningen binnen het team	Jehn, K. A. (1995). A Multimethod Examination of the Benefits and Detriments of Intragroup Conflict. <i>Administrative science quarterly</i> , 40(2), 256. doi:10.2307/2393638 Jehn, K. A., Northcraft, G. B., & Neale, M. A. (1999). Why differences make a difference: A field study of diversity, conflict and performance in workgroups. <i>Administrative science quarterly</i> , 44(4), 741-763.
7		Er is onenigheid tussen collega's in het team	
8		Collega's in het team zijn onredelijk tegen elkaar	
9		Het team heeft conflicten over de manier waarop ze hun taken moeten uitvoeren	
10	Activated Faultlines	Het team is verdeeld in klikjes (kleine groepjes)	Jehn, K. A., & Bezrukova, K. (2010). The faultline activation process and the effects of activated faultlines on coalition formation, conflict, and group outcomes. <i>Organizational Behavior and Human Decision Processes</i> , 112(1), 24-42. doi:10.1016/j.obhdp.2009.11.008 Self – made questions
11		Er zijn conflicten door de culturele verschillen tussen collega's in het team	
12		Er zijn conflicten door de leeftijdsverschil tussen collega's in het team	
13		Er zijn conflicten tussen mannen en vrouwen in het team	

^a All items were asked only in Dutch as all team managers were proficient Dutch speakers

Appendix C

Explanation of calculation of Faultlines Indexes

Table C1. Team Demographic Information

N	Age	Gender	Nationality	Origin
1	young	m	ER	ER
2	old	f	NL	NL
3	old	m	NL	AN
4	old	m	NL	FR
5	middle	f	NL	NL
6	young	f	ER	ER
7	young	f	NL	NL
8	middle	f	NL	NL
9	middle	f	NL	NL
10	middle	m	NL	NL
11	middle	f	NL	AF
12	middle	m	NL	NL
13	old	f	NL	ID
14	young	f	ER	ER

Table C2. Results of ASW analysis conducted using free R package “asw.cluster”

Faultlines' Strength is based on:	ASW Value	N of subgroups	Size of subgroups	Members to subgroups
Age+ Gender+ Nationality	0.815	6	2 2 2 4 2 2	1 2 3 3 4 5 1 4 4 6 4 6 2 5
Gender+ Nationality	0.929	4	1 7 4 2	1 2 3 3 2 4 2 2 2 3 2 3 2 4
Age+ Gender + Origin	0.505	6	3 3 2 3 2 1	1 2 3 3 4 1 2 4 4 5 6 5 2 1
Gender+ Origin	0.643	5	3 5 2 2 2	1 2 1 1 2 3 2 2 2 4 5 4 5 3
Nationality	1	2	3 11	1 2 2 2 2 1 2 2 2 2 2 2 1
Origin	0.714	3	3 7 4	1 2 3 3 2 1 2 2 2 2 3 2 3 1

Notes. Colours in both tables indicate how ASW algorithm in R distributed team members to 6 different subgroups if age, gender and nationality are taken into account

Appendix D

Explanation of calculation of Blau's Diversity Indexes

Blau's (1977) Diversity Index formula: $1 - \sum_{i=1}^k p_i^2$

Table D1. Team Demographic Information and number of different Demographic Categories

N	Age	Gender	Nationality	Origin	N of different demographic categories			
					Age	Gender	Natl.	Origin
1	middle	f	NL	NL	Middle:6	f: 9	NL: 11	NL: 7
2	middle	f	NL	NL				
3	middle	f	NL	NL				
4	middle	m	NL	NL		m: 5		
5	middle	f	NL	AF				AF: 1
6	middle	m	NL	NL				
7	old	f	NL	NL	Old: 4			
8	old	m	NL	AN				AN: 1
9	old	m	NL	FR				FR: 1
10	old	f	NL	ID				ID: 1
11	young	m	ER	ER	Young: 4		ER: 3	ER: 3
12	young	f	ER	ER				
13	young	f	NL	NL				
14	young	f	ER	ER				
Blau's Diversity Index Value:					0.653	0.456	0.337	0.684

$$\text{Blau's Diversity Index (Age)} = 1 - \left(\left(\frac{6}{14} \right)^2 + \left(\frac{4}{14} \right)^2 + \left(\frac{4}{14} \right)^2 \right) = 0.653$$

$$\text{Blau's Diversity Index (Gender)} = 1 - \left(\left(\frac{9}{14} \right)^2 + \left(\frac{5}{14} \right)^2 \right) = 0.459$$

$$\text{Blau's Diversity Index (Nationality)} = 1 - \left(\left(\frac{11}{14} \right)^2 + \left(\frac{3}{14} \right)^2 \right) = 0.337$$

$$\text{Blau's Diversity Index (Origin)} = 1 - \left(\left(\frac{7}{14} \right)^2 + \left(\frac{1}{14} \right)^2 + \left(\frac{1}{14} \right)^2 + \left(\frac{1}{14} \right)^2 + \left(\frac{1}{14} \right)^2 + \left(\frac{3}{14} \right)^2 \right) = 0.684$$