



OPEN ACCESS

EDITED BY

Carlos María Alcover,
Rey Juan Carlos University, Spain

REVIEWED BY

Jennie Weiner,
University of Connecticut,
United States
Sheila Keener,
Old Dominion University,
United States

*CORRESPONDENCE

Rebecca Gerlach
Rebecca.gerlach@
psychologie.tu-chemnitz.de

SPECIALTY SECTION

This article was submitted to
Organizational Psychology,
a section of the journal
Frontiers in Psychology

RECEIVED 27 August 2021

ACCEPTED 10 August 2022

PUBLISHED 29 September 2022

CITATION

Gerlach R and Gockel C (2022) A question
of time: How demographic faultlines and
deep-level diversity impact the
development of psychological safety in
teams.

Front. Psychol. 13:765793.

doi: 10.3389/fpsyg.2022.765793

COPYRIGHT

© 2022 Gerlach and Gockel. This is an
open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that
the original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution or
reproduction is permitted which does not
comply with these terms.

A question of time: How demographic faultlines and deep-level diversity impact the development of psychological safety in teams

Rebecca Gerlach^{1*} and Christine Gockel²

¹Department of Psychology, Chemnitz University of Technology, Chemnitz, Germany, ²Berlin School of Management, SRH University Berlin, Berlin, Germany

Psychological safety (PS) is a shared belief among team members that it is safe to take interpersonal risks. It can enhance team learning, experimentation with new ideas, and team performance. Considerable research has examined the positive effects of PS in diverse organizational contexts and is now shifting its focus toward exploring the nature of PS itself. This study aims to enhance our understanding of PS antecedents and development over time. Based on the model of team faultlines and research on team diversity, we examined the effects of demographic faultlines, team member personality, and member competencies on the development of PS. Over 5 months, 61 self-managed teams ($N=236$) assessed their PS at the beginning, midpoint, and end of a research project. Results of a multilevel growth curve model show that PS decreased from project beginning to end. Initial levels of PS were especially low when teams had strong demographic faultlines and when team members differed in neuroticism. PS decreased more strongly over time when team members were diverse in agreeableness and assessed their task-related competencies to be relatively high. Our study identifies time and team composition attributes as meaningful predictors for the development of PS. We present ideas for future research and offer suggestions for how and when to intervene to help teams strengthen PS throughout their collaboration.

KEYWORDS

psychological safety, development, team member personality, competency, growth curve modeling, demographic faultlines

Introduction

Psychological safety (PS) is a key factor for successful collaboration (Bergmann and Schaeppli, 2016). It is defined as team members' shared belief that their environment is conducive to interpersonally risky behaviors such as speaking up, asking for help, or owning up to mistakes (Edmondson and Lei, 2014). In teams with a high level of PS, members send and receive signals to each other, prompting them to share ideas or admit mistakes,

compared to teams with a low level of PS. They report more errors because they are not afraid of being seen as incompetent or overstrained by other team members or their team leader (Edmondson, 1999). When team members experience PS, they regard the team to be a protected, safe context in which they can take risks without the potential for negative consequences (Edmondson, 1999; Garvin et al., 2008). Thus, PS forms the basis for a supportive learning environment for individuals, teams, and organizations (Edmondson, 1999; Garvin et al., 2008; Frazier et al., 2017; Harvey et al., 2019).

Although considerable theory and research link PS with numerous organizational outcomes, for the most part, extant studies are based on state descriptions at one specific point in time (Frazier et al., 2017). Recently, researchers' attention has shifted to PS as an emerging phenomenon with limits in its dynamics and mode of action over time (e.g., Gerlach and Gockel, 2018; Deng et al., 2019; Harvey et al., 2019; Higgins et al., 2020). For example, researchers have focused on differences within team members' perception of PS and showed that belonging to a leader's in-group (compared to an out-group) protected team members from the negative effect of task conflict on PS (Gerlach and Gockel, 2018). Also, PS had the most positive effect on organizational performance over time when being relatively low (as compared to high) combined with high levels of employees feeling accountable for their actions or decisions (in a 3-year study by Higgins et al., 2020). Moreover, a recent interview study explored how leaders experienced and created PS under extreme situations such as the COVID-19 pandemic allowing PS to unfold its positive impact on learning opportunities together with environmental and organizational factors (Weiner et al., 2021). Altogether, these findings have highlighted the importance of subgroup dynamics and boundary conditions for the creation of PS by using more complex study designs.

Regarding the development of PS, theory suggests that PS—as an emergent state—is bound to time and changes as a function of what happens in the team (see a dynamic model of team learning climate, Harvey et al., 2019). However, empirical evidence is mixed so far. In a study of innovative project teams, no changes in PS were found (Edmondson and Mogelof, 2005). Another study indicates that PS decreases, as it co-evolves with network ties over time (Schulte et al., 2012). Thus, we hope to shed light on these contradictory findings by exploring the following questions: How does PS form and develop in teams? Which antecedents related to team member characteristics facilitate changes in PS?

This study is designed to address these questions using an approach that is grounded in theories on team learning and team diversity: More specifically, we build on the theoretical model of group faultlines (Lau and Murnighan, 2005) that links subgroup dynamics from group diversity (what members bring into the group) to the formation and development of PS. We tested the effects of demographic faultlines, team member personality characteristics, and team member competencies as predictors for the formation and development of PS in teams. In our study, we examined self-managed project teams from their formation

until task completion. Employing a multilevel growth curve model, we investigate initial levels and changes in PS over time. This study advances our understanding of how member diversity and time affect the development of PS, with an emphasis on team learning opportunities.

Time, team development, and psychological safety

The perception of PS by single members, subgroups, and across the entire team forms early after team formation and continues to develop over time (Harvey et al., 2019). Teams might grow together. Yet, critical incidences or transitions might impair a desirable positive development (Marks et al., 2001; Harvey et al., 2019). For example, whereas leaders can facilitate the emergence of PS by inviting members to discuss opposing views (Chen et al., 2011), conflicts or disagreements between team members can reduce the chance for a learning climate within a team (Bunderson and Boumgarden, 2010).

Teams are bound to time and task progress in their work: The beginning, the midpoint, and the end of a project have been identified to be important for team development and the design of team interventions (e.g., Gersick, 1988; Hackman and Wageman, 2005; Bush et al., 2018, for an overview on team development theories, see Delice et al., 2019). As described in the punctuated equilibrium model, teams pay attention to time and organize their work in response to the time that is available to complete their task (Gersick, 1988). Specifically, team members are sensitive to the passage of time and the corresponding increase in time pressure and may address this by changing their strategies or approaches to the task. Time is found to be understood as part of the task that needs to be managed by teams stressing the midpoint as one important turning point for team processes and outcomes (Okhuysen and Waller, 2002; Ford and Sullivan, 2004). As such, teams show a particular tendency to make radical changes around the midpoint, for example in their informal leadership structures, marking a transition between two rather balanced phases of collaboration (Gersick, 1988). With the passage of the midpoint, teams can change their motivational frame from being task-focused to performance-focused, adjusting their exploratory search for information during these phases accordingly (Knight, 2015). Team mood is the driver of this change of focus over time. Teams might ask, around the temporal midpoint, "Did we do a good job so far, and are we satisfied with where we stand? Did we put enough resources into searching for information and ideas on how to solve the task, or do we need more information for a solid decision?," as compared to "Can we move on to the performance phase?" These aspects might be related to team members' specific learning behavior *search for information* that is bound to the ongoing performance of the task. In this study, we focus on PS, which is the supportive environment for learning throughout a team project.

We examine the kind of change that PS shows during a project and the antecedents for initial levels and changes in PS.

To date, only three studies have investigated the temporal development of PS in project groups: First, innovative teams were examined from the beginning of a project to the end (Edmondson and Mogelof, 2005). In this ground-breaking study, PS was found to be stable across three measurements ($M_{T1} = 3.06$; $M_{T2} = 3.07$; $M_{T3} = 3.01$). Second, the relation between PS and friendship ties in social networks was examined in consulting teams working together for the time of a project (Schulte et al., 2012), with PS decreasing from beginning to the end ($M_{T1} = 3.81$; $M_{T2} = 3.59$; $M_{T3} = 3.37$; no tests of significance reported). A third study found only a moderate correlation ($r = 0.27$) between two PS assessments taken at the individual level 6 weeks apart (Liang et al., 2012), which is an indication that PS assessments have changed.

We believe that PS should diminish in teams from the beginning to the end of a project. Research on the need for closure (Pierro et al., 2003) implies that when teams need to make decisions under time pressure, they tend to cut off discussions and strive for task completion. Therefore, we predict that, as time passes, time pressure will contribute to an atmosphere of closed-mindedness in the group. As a deadline approaches, team members should be less likely to pose questions, speak up, or share controversial opinions and, if they continue to do so, are more likely to be sanctioned by their fellow team members. Supporting this prediction, research also shows that novel contributions are more positively valued when being made early during collaboration compared to past the midpoint when this might disrupt performance and lead to frustration (Ford and Sullivan, 2004). Sharing novel contributions is a learning behavior closely linked to PS in teams. Thus, PS might develop similarly in project teams. Therefore, we propose the following.

Hypothesis 1: As time passes in a team project, psychological safety decreases.

Demographic faultlines predict the formation of psychological safety

Group members are sensitive to signals of PS from the very beginning of their collaboration (Gersick, 1988; Hackman and Wageman, 2005). They assess each other instantly depending on overt demographic attributes such as gender, ethnicity, age, tenure, and—if available—functional background (Lau and Murnighan, 2005). These overt demographic characteristics are summarized under the term surface-level diversity or demographic diversity, which can be easily assessed after a brief contact (Bell, 2007). Specifically, team members use demographic differences and overt skills to classify themselves and others into homogeneous (sub-) groups with which they can identify. Research shows that demographic diversity may result in improved decision-making and problem-solving (Watson et al., 1993), but it is also associated with reduced interpersonal liking, intergroup communication,

and team cohesion (Tsui and O'Reilly, 1989; Tsui et al., 1992; Smith et al., 1994), as well as an increase in team conflict (Jehn, 1995).

For a long time, different conceptualizations and measurements have led to an enormous body of inconsistent findings integrated for the first time by the theory and unified measurement of team faultlines (Thatcher and Patel, 2012). The term “faultline” comes from the field of geography and refers to the boundary zones between tectonic plates of continents. Transferred to the context of group research, it describes a hypothetical line that potentially splits a team into homogeneous subgroups according to team members' similarity along with a set of multiple attributes (Lau and Murnighan, 1998). The term “team faultlines” refers only to the joint effect of surface-level attributes such as gender, age, ethnicity, or functional background that serve as the basis for the categorization processes of team members into possible homogeneous subgroups (Lau and Murnighan, 1998, 2005). To be more specific about the type of faultlines, Lau and Murnighan (1998) introduced the term demographic faultlines which can be weak or strong. As such, demographic faultline strength was found to have a consistently negative effect on team processes and outcomes (Thatcher and Patel, 2012).

The model of demographic faultlines (Lau and Murnighan, 2005) links team diversity with perceptions of PS on a theoretical level. The model proposes that demographic faultlines negatively affect PS as well as team learning, anticipated team performance, and team satisfaction (Lau and Murnighan, 2005). In teams with a strong demographic faultline, team members identify, communicate, and share information more often with members of the same subgroup than with members from other subgroups (Lau and Murnighan, 1998, 2005), which should undermine the overall team perception of PS. The model of group faultlines suggests further that demographic faultlines affect PS at the start of a team project, because it is essential for members to find productive subgroups right away, and they often do so based on a preliminary assessment of overt surface-level attributes (Lau and Murnighan, 1998, 2005). Thus, we propose the following.

Hypothesis 2: Demographic faultline strength is negatively associated with initial levels of psychological safety.

Team member personalities and competencies predict the formation and change of psychological safety

Deep-level diversity refers to underlying psychological characteristics, such as personality factors, abilities, and values (Bell, 2007). Team member personality characteristics and competencies have been identified as the most important predictors for effective teamwork, which is an important outcome associated with PS (Kozlowski and Bell, 2003; Newman et al., 2017). Team members' personality characteristics translate to the team level, especially in self-managed teams where team members'

roles emerge and transform over time (Barry and Stewart, 1997). Compared to surface-level diversity, deep-level diversity regarding team member personality and competencies is found to be more important for team functioning over time (Harrison et al., 2002). The authors explain the finding in such a way that team members have to get to know each other first to learn about each other's characteristics. On the other hand, the joint task may also reveal the requirements of certain characteristics at different points in time. In the current study, we refer to the five-factor model of personality, which is a widely accepted model of personality and is often called the Big Five (Goldberg, 1993; Costa and McCrae, 1994). It organizes broad individual differences in social and emotional experiences into five different categories—based on factor analyses applied to personality survey data (McAdams and Pals, 2006). These categories are usually labeled extraversion (vs. introversion), neuroticism (vs. emotional stability), openness to experience, conscientiousness, and agreeableness. The Big Five are relatively consistent and stable over time, have psychobiological underpinnings, and can predict behavioral tendencies and life outcomes (e.g., Matthews et al., 2009). They have also been used effectively in several past studies focused on individual differences between team members (e.g., Bonner et al., 2007).

Previous research links three of the Big Five personality characteristics, namely extraversion, openness to experiences, and neuroticism, to PS on a theoretical level (Edmondson and Mogelof, 2005). This research revealed that, on an individual level, neuroticism and openness to experience predict individual PS perceptions at the midpoint of a team project and that neuroticism predicts PS perceptions at the end of the project. Contrary to the authors' expectations, extraversion did not predict PS perceptions at any measurement point. Edmondson and Mogelof (2005) explain the lack of an effect of extraversion by arguing that extraversion encompasses a wide variety of behaviors and thus, might not be close enough to PS-relevant behaviors such as speaking up—although extraversion is understood to be an important antecedent of this specific communication behavior. They proposed that the investigated personality characteristics extraversion, openness, and neuroticism might translate to the team level as well. Results from a team study on composition and compilation effects identified conscientiousness as an important predictor of PS if its distribution is skewed in the team (Ostermeier et al., 2020). Taken together, the theoretical link between personality characteristics and PS is promising and needs more clarity.

Despite the broader assumption that time catalyzes the effect of team personality (Harrison et al., 2002), we have reason to believe that extraversion and neuroticism are exceptions that affect the formation of PS. As such, extraversion is important for PS formation because individuals with high levels of extraversion enter new ambiguous social situations self-confidently (Anthony et al., 2007; Gebauer et al., 2015) and promote the experience of a positive mood in the team (Barrick et al., 1998). Neuroticism is important for PS formation, as it is closely linked to the concept of negative affectivity (Stokes and Levin, 1990), a dispositional

tendency to experience aversive feelings, which has been shown to predict initial levels of PS (Ostermeier et al., 2020). Individuals with high levels of neuroticism and likewise teams with neurotic team members experience higher feelings of distress, anger, frustration, or anxiety in social interactions (Levin and Stokes, 1989) impairing the emergence of PS at the beginning and its development. Furthermore, deep-level diversity in terms of the personality characteristics of openness to experience, conscientiousness, and agreeableness should become more important during the later stages of group work when PS has developed further (Harrison et al., 2002).

We base our reasoning on the model proposed by LePine et al. (2011) on how team member personality characteristics can translate to the team level either through individual personality traits that directly affect team composition or corresponding behaviors that affect team processes indirectly. Following the principle of parsimony, we build our team research on two functional relations between team composition and PS proposing that team composition might have additive (when the combination of group members' scores is equal to the sum of the effects) and/or compensatory effects (when group members' scores on an attribute balance one another) regarding team characteristics. Thus, we relate team composition attributes with the respective team mean values and/or within-standard deviations as an index of variation to PS. We explain our reasoning in more detail below and postulate our assumptions on the team level linking team members' personality characteristics to the formation and changes in PS.

In the current study, we focus on extraversion as a predictor for the formation of PS. Extraverts are sociable, enthusiastic, energetic, and assertive (Goldberg, 1993; Barrick et al., 1998). Behaviors relating to extraversion can be more easily observed than behaviors relating to the other Big Five (John and Robins, 1993). Therefore, extraversion should be particularly important at the beginning—as compared to the other traits. In teams with highly extroverted members, PS should form quickly when members signal to each other that the team has a friendly atmosphere where open communication is encouraged. This assumption is supported by a positive correlation between team extraversion (mean values) and communication on the team level (Barrick et al., 1998). Further, extraversion comprises the experience of positive emotional states (Barrick et al., 1998) that promote well-being and prepare the ground for a supportive learning atmosphere right from the start. This positive effect on PS is likely to be additive: the more extraverts belong to a team, the more signals of sociability and positive affectivity they might send to each other within the first meetings. If the team is high in extraversion (i.e., has a high mean in team member extraversion), the positive effect might wear out at a later stage of the collaboration when PS has been established. However, in the beginning, team members' extraversion should facilitate communication behaviors and a positive sentiment toward each other that form the basis for PS. Thus, we propose:

Hypothesis 3a_(E, Mean): The higher a team's mean level in extraversion, the higher the initial levels of psychological safety will be in the team.

The distribution of extraversion in the team should also be relevant to the formation of PS. Extraverts are assertive and tend to show leadership behaviors (Judge et al., 2002). Thus, if team members vary in their level of extraversion, some members might dominate other members from the start with a rather disconcerting, negative effect on other team members' perception of the team environment for inter-individual risk-taking. Thus, we propose

Hypothesis 3b_(E, SD): The more diverse team members are in extraversion, the lower initial levels of psychological safety will be in the team.

As a second deep-level predictor for the formation and changes in PS, we chose neuroticism, which is closely linked with negative affectivity (Goldberg, 1993; Barrick et al., 1998), and should obstruct trust and communication behaviors that form the basis for PS. Team members who are high in neuroticism are less emotionally stable and tend to be anxious, depressive, and angry (George, 1990; Barrick et al., 1998). In contrast to their emotionally stable counterparts, highly neurotic team members should impede a continued process of positive interactions. More specifically, they should reduce the team's responsiveness to feedback, because mistakes that are mentioned might be taken personally and not as constructive feedback. Team neuroticism should have an additive and a compensatory effect. A team composed of highly neurotic members is likely to create a tense atmosphere inhibiting cooperation, which translates into the joint perception that it is not safe to take risks from the beginning. A recent study about the relationship between negative affectivity and early PS supports this assumption: Teams high in negative affectivity, operationalized as a trait that reflects individuals' dispositional tendency to experience aversive emotions, started teamwork with lower levels of PS (Ostermeier et al., 2020). Thus, we propose

Hypothesis 3c_(N, Mean): The higher a team's mean level in neuroticism is, the lower initial levels of psychological safety will be.

In addition, diversity in neuroticism might become more important when mistakes occur, and task conflicts arise during project work. When team members are diverse in neuroticism, then there are a few members high in neuroticism; these members might destroy PS and impede team learning for everyone in the team by expressing their discomfort with feedback and open ideas. This effect can be explained by the positive–negative asymmetry, a phenomenon in impression formation that states that bad information carries more weight and influences the impression more than good information (e.g., Peeters and Czapinski, 1990). The underlying reason is that “bad” behaviors (in this case, the

neurotics who express discomfort with feedback) are more diagnostic than “good” behaviors (in this case, the emotionally stable who express comfort with open feedback) because the consistency requirements for PS are more stringent than for the absence of PS. In other words, to experience PS, all team members must be comfortable with feedback. For a low level of PS, a few comments from neurotic team members are sufficient. Thus, we propose

Hypothesis 3d_(N, SD): The more diverse team members are in neuroticism, the more psychological safety will decrease over time.

As a third deep-level predictor for changes in PS, we chose openness to experience, because it is closely linked with learning behaviors. Open individuals are naturally curious and open-minded regarding unconventional ideas and experimentation (Goldberg, 1993; Barrick et al., 1998). As such, teams with high levels of openness are likely to foster the discussion of problems by having members offer unconventional ideas or inviting each other to think outside the box. This kind of communication is more important during collaboration—and especially when teams encounter difficulties—than at the beginning. Translating the relationship between openness to experience and PS found by Edmondson and Mogelof (2005) to the team level, we postulate an additive effect of team member openness on the development of PS:

Hypothesis 3e_(O, Mean): The higher a team's mean level in openness to experience, the more psychological safety within the team will increase over time.

As the fourth deep-level predictor for PS, we chose conscientiousness, an important characteristic for thorough work, organization, and the detection of errors (Goldberg, 1993; Barrick et al., 1998) that can be linked to specific learning behaviors. A team composed of highly conscientious members works thoroughly, is achievement-oriented, and engages in task-focused roles (Bell, 2007). Especially the aspect of duty fulfillment is an important part of the concept of conscientiousness that should foster PS in teams because members have a high interest in doing the right thing for themselves and others. They understand that sharing relevant information is part of their duty that might create a norm on the team level (Ostermeier et al., 2020). Further, conscientiousness is related to helping behavior in a consistent way (Hurtz and Donovan, 2000). In a recent study, team mean conscientiousness was hypothesized to have a positive effect on PS (Ostermeier et al., 2020). Yet, results from a path model did not support this assumption. Instead, team conscientiousness was found to have a negative effect on PS, if its distribution was skewed in the team (Ostermeier et al., 2020). More specifically, if conscientiousness was positively skewed in a team, that is, when there were more highly conscientious members compared to members low in conscientiousness, PS decreased. Based on these findings, we propose the following

Hypothesis 3_{f(C, SD)}: The more diverse team members are in conscientiousness, the more psychological safety will decrease over time.

Finally, as the fifth deep-level predictor for PS, we chose agreeableness. It describes characteristics such as “helpful, friendly, warm, trusting, and tolerant. In fact, the very essence of agreeableness is cooperation” (Barrick et al., 1998, p. 381). We find a theoretical overlap between the concepts of agreeableness and PS, as it encompasses attributes of a comfortable learning atmosphere such as trustful, warm, tolerant, as well as specific learning behaviors such as being helpful, considerate, or caring. A team composed of highly agreeable members might encompass a group norm of being prosocial and inclined toward others (Graziano et al., 2007). Such teams report having fewer conflicts (Barrick et al., 1998). A recent study investigated the relation between team agreeableness and initial levels and changes in cohesion—the forces that help members belong to the team (Festinger, 1950). Acton et al. (2020) found that teams composed of highly agreeable members reported an increase in cohesion over time. PS and cohesion are both emergent states, only PS fosters the exchange of divergent opinions (Edmondson, 2002), whereas cohesion is partly related to speaking up less often and exchanging opposing opinions less strongly (Rose et al., 2011). Based on these findings, we propose

Hypothesis 3_{g(A, Mean)}: The higher a team’s mean level of agreeableness, the more psychological safety within the team will increase over time.

Still, a single highly disagreeable member can be enough to undermine the team’s capability to work together (Barrick et al., 1998). When members of a team strongly vary in their agreeableness, this can damage the relationships between team members. For instance, a disagreeable team member might start conflict stressing agreeable members in their striving for harmony and a task-focused working atmosphere at a later point of collaboration. The effect of variability in team member agreeableness can also be explained by the positive–negative asymmetry (e.g., Peeters and Czapinski, 1990). The “bad” behaviors of disagreeable team members who start conflicts have a stronger impact on PS as compared to the “good” behaviors of agreeable team members striving for harmonious interactions. In previous research, variance in team agreeableness was negatively related to social cohesion, communication, and workload sharing, while being positively related to team conflict (Barrick et al., 1998). We therefore propose.

Hypothesis 3_{h(A, SD)}: The more diverse team members are in agreeableness, the more psychological safety will decrease over time.

Besides team member personality characteristics, team member competencies should also predict the development of

PS. Member competencies are conceptualized as personal characteristics or specific behaviors that are related to performance in a certain context (Krumm et al., 2012). We refer to task-relevant competencies as an important resource for team learning because they help members identify and correct mistakes and encourage them to consult each other for advice (Doblinger, 2022). They influence what and how team members communicate and how the team operates and performs (Hackman et al., 1976). Teams need to generate performance strategies that are appropriate to solve the task, and they do so based on their knowledge and competencies (Hackman et al., 1976). However, a high level of competency does not necessarily need to translate into positive results. On the contrary, members generate private hypotheses about how a task needs to be approached or solved and rarely discuss it explicitly with each other (Hackman et al., 1976). Task-oriented teams tend to meet the implicit norm of NOT addressing process issues (Argyris, 1969), which means that team members share less or no information about their individual strategies. In an experiment with student teams (Hackman et al., 1976), teams with unequal information about the task performed better, if an explicit discussion of performance strategies was a part of their team task (as compared to teams in the control group without intervention or team in another intervention group that prohibited the discussion of performance strategies as inefficient and waste of time). Teams with equal information only performed well when they did not discuss performance strategies at the beginning of the team task. Previous research supports the current relevance of this phenomenon, showing that teams differ in their ability to utilize their resources and successfully transform them into good team decisions, team performance, or creativity (Innami, 1994; Taggar, 2002; Mayo and Woolley, 2021). The reason for this unexpected negative link lies in the teams’ communication behavior. Teams with ample resources (such as high task-relevant knowledge) tend to have discussions with members sticking to their positions and exchanging fewer facts and reasons to explain their positions as compared to teams with little resources (Innami, 1994). This finding is in line with former observations as described above (e.g., Argyris, 1969; Hackman et al., 1976), thereby relating team members’ competencies to specific learning behaviors, namely exploring new knowledge and information and engaging in a collaborative conversation to refine, build or modify collective knowledge (Wiese and Burke, 2019). Based on these previous findings, we believe that team member competencies can also impact PS perceptions. More specifically, they should impact how PS changes over time. Team member competencies should have an additive effect, that is, teams with high task-relevant competencies might be more likely to detect problems or provide solutions, yet, they might have difficulties elaborating on adequate solutions due to their reduced communication behavior (Innami, 1994). Thus, we propose

Hypothesis 4_(Competency, Mean): Team member task-relevant competencies predict changes in psychological safety: The

greater a team's competency, the more psychological safety will decrease over time.

Materials and methods

Participants

Our study was part of a bigger research project about time and changes in teams at a university in Germany. Sixty-one student teams of two to five members each participated in this study ($M = 3.87$, $SD = 0.53$). They worked on a university research project, which was part of their curriculum, over the course of 5 months. In total, 236 students participated in this study. Of these students, 68% were psychology students and 32% were students from a bachelor's program combining psychology, physics, and cognitive sciences. Data were collected over a period of three semesters. The majority of the participants were female (77.12%) and in their second year of study. Participants were between 19 and 44 years old ($M = 22.34$, $SD = 3.43$).

Team task and participant recruitment

In their second year of study, psychology students needed to complete a course in experimental research. Their task was to design and conduct a research project: They developed hypotheses, recruited participants, collected data, ran analyses, and critically discussed their results. At the first course meeting, we invited students to participate in our study. At this time, some teams had already formed on a voluntary basis. We briefly described the goal of our study and explained how data would be collected. In exchange for every completed questionnaire, students received research credits. In addition, teams could win a movie voucher in a lottery to celebrate their project accomplishment. All 62 teams who were invited to the study participated. We excluded one team consisting of four members from further analyses because one member quit the class after the midpoint of the project leaving team assessments of PS incomplete.

Design and procedure

The longitudinal design makes it possible to study temporal dynamics in teams. We asked team members to complete questionnaires at three points in time that are important for team development, namely at the beginning, at the midpoint, and at the end of the project (Gersick, 1988; Chang et al., 2003). We further related our measurement points to the following tasks that teams had to fulfill during their course: at the first point of measurement, approximately 4 weeks following the start of the class, teams presented the theoretical framework for their projects. We assumed that

team members required this time to meet and get acquainted with each other before PS could be assessed adequately. At the second point of measurement, when half of the project time had passed, teams presented their methods. At the third point of measurement, which was at the end of the project, teams presented and discussed their results. Data were collected anonymously. Team diversity variables were assessed at the first measurement point. Team members assessed their PS at all three measurement points.

Measures

All scales had the same response format: On a Likert-type scale, team members rated the extent to which they agreed with the statements, ranging from 1 (*not at all*) to 7 (*absolutely*).

Team psychological safety

We assessed PS with the Team Psychological Safety Scale (Edmondson, 1999). An example item is "If you make a mistake on this team, it is often held against you" (reverse scored). We translated the scale into German, remaining as close to the original expressions as possible. After the first point of measurement, several team members reported having difficulties in responding to the fourth item "It is safe to take a risk on this team." We excluded this item so that six items remained in our analyses (We report omega total as suggested by McNeish, 2018 and Hayes and Coutts, 2020: $\omega_{total1} = 0.59$, $\omega_{total2} = 0.77$, and $\omega_{total3} = 0.70$ for six items each).

Demographic faultline strength

We calculated demographic faultline strength using the average silhouette width (ASW) cluster algorithm in R (Meyer and Glenz, 2013). The surface-level faultline included the demographic characteristics of gender, age, course of study (psychology or cognitive science), and semester term (third, fourth, or fifth semester).

The ASW measure lies between -1 and 1 . Values near 1 represent the emergence of mostly homogeneous subgroups, whereas values near 0 represent mostly diverse subgroups as no homogeneous subgroups exist. ASW categorizes members into subgroups in a stepwise approach. First, each team member has his or her own subgroup, consisting of him or herself only. Next, members with the most similarities are merged into a subgroup of two, and so on, until all existing subgroup possibilities are calculated. Then, the mean ASW value is computed for each team member representing how well the member fits into the subgroup. The output in R provides faultline values per team and further information on the team level such as the number of subgroups, the subgroup size, and an explicit allocation of members to subgroups (see Meyer et al., 2015). In most teams ($n = 52$), only two homogeneous subgroups emerged. The subgroup size was, on average, 2.2 members per subgroup.

Team member personality

We used the Big Five Inventory-25 to measure the personality dimensions of extraversion, neuroticism, openness to experience, conscientiousness, and agreeableness with five items each. The inventory is a shorter version of the 44-item Big Five Inventory by John et al. (1991), which was translated into German and validated by Rammstedt et al. (2004). Example items are “I see myself as someone who is “...talkative” for extraversion ($\omega_{\text{total}}=0.90$), “...worries a lot” for neuroticism ($\omega_{\text{total}}=0.84$), “...inventive” for openness ($\omega_{\text{total}}=0.86$), “... tends to be disorganized” (inverted item) for conscientiousness ($\omega_{\text{total}}=0.84$), and “... generally trusting” for agreeableness ($\omega_{\text{total}}=0.69$).

Team member competency

We assessed team members' competencies in statistics and methodology because group members needed them in all steps of their empirical research project. For example, when reading the introductory articles, they needed to understand what had been done before and how. When planning the experiment, basic methodological knowledge helped to select a design and appropriate measures. When analyzing data, statistical knowledge is essential for testing hypotheses. We adapted a German version of the Academic Self-Description Questionnaire (ASDQ; Marsh et al., 2005b) to the context of these groups. In general, the ASDQ measures self-assessed competencies in school subjects with six items each (e.g., in math, language, or sports). We replaced the school subjects with statistics and methodology. An example item is “In comparison to my fellow students, I am very good at statistics and methodology” ($\omega_{\text{total}}=0.90$). Marsh et al. (2005a) reported moderate correlations between ASDQ assessments and the respective grades. Because students had taken courses in statistics before participating in this group research project, we assumed that they were able to properly answer the questionnaire.

Control variables

When testing the effects of demographic faultlines, we included demographic attributes that may affect the team outcome above and beyond faultline strength on the team level (see Meyer et al., 2015). We operationalized demographic diversity effects on a team level by including the Blau index (Blau, 1977) for gender diversity and the within-team standard deviation in tenure (number of semesters).

Results

We used multilevel-based linear growth curve modeling (LGCM) to test our hypotheses (Gajek and Burzykowski, 2013; Bliese, 2016). Linear growth curve analysis is used in several psychological disciplines to investigate developmental trajectories, e.g., those of personality development or the impact of time on the development of social behavior (Jung and Wickrama, 2008). This method allows us to study the development of PS in the form of

overall increase or decrease by modeling initial levels (the intercept) and changes within teams (interaction with time). With this statistical method, we can examine differences within and between teams over time and we can test our hypotheses about how team faultlines and diversity in member personality and competencies predict initial levels and changes in PS.

Agreement between team members and data aggregation

In the first step, we calculated agreements between team members to test for the nested structure of our data (Bliese, 2016). Thus, we calculated intraclass correlation coefficients $ICC(1)$ and $ICC(2)$. The $ICC(1)$ value indicates the amount of variance that can be explained by team membership. The $ICC(2)$ value indicates the extent to which the team's mean rating is reliable, thereby taking average team size into account (Bliese, 2000; Grawitch and Munz, 2004). As expected, we found high and significant ICC values for PS, with $ICC(1)$ ranging between 0.14 and 0.26, and $ICC(2)$ between 0.37 and 0.58 across measurement points. The findings indicate high agreements within teams and differences between teams. In comparison, we found non-significant ICC values for four of the five personality characteristics and competencies indicating differences between team members, with $ICC(1)$ ranging between 0 and 0.08, and $ICC(2)$ between 0 and 0.22; except for conscientiousness with $ICC(1)=0.12$, $p < 0.05$ and $ICC(2)=0.35$. Because we were interested in how team diversity variables, and not individual characteristics, affect initial assessments and changes in PS, we aggregated all predictor variables to the team level for hypothesis testing. We used the standard deviation within teams as an index for team diversity regarding team member personality (see Harrison et al., 2002). We also tested for measurement invariance (MI) of the PS Scale before conducting growth analysis as recommended by Van de Schoot et al. (2012). Results confirmed a strict invariance of the PS scale across the three measurement points because the models had a very good fit and did not differ significantly from one another (configural MI, metric MI, adjusted scalar MI, and strict MI). Correlations between demographic faultline strength, team diversity variables, and team PS are displayed in Table 1.

The reasoning of control variables

Due to the complexity of our tested model, we reduced control variables to the least possible number according to their theoretical contribution to our model (Becker, 2005). First, we controlled for diversity regarding gender, because women differ in their perceptions of PS as compared to men (Carmeli and Gittell, 2009). Second, we controlled for tenure diversity. Tenure has a distinct meaning for personal and organizational identity building as compared to age, especially if members enter the organization at the same time (e.g., Meyer et al., 2015). Faultline attributes often

TABLE 1 Means, SD, and correlations among variables at team level at all measurement points (N=61 teams).

Variable	M	SD	1.1	1.2	1.3	1.4	1.5 _E	1.5 _N	1.5 _O	1.5 _C	1.5 _A	1.6	1.7 _E	1.7 _N	1.7 _O	1.7 _C	1.7 _A	1.8	1.9	2	3
Beginning																					
1.1 Team Size	3.87	0.53	-																		
Surface-Level Diversity																					
1.2 Gender, Blau's Index	0.21	0.27	-0.13	-																	
1.3 Tenure, S.D.	0.28	0.65	-0.20	0.26	-																
1.4 Demographic	0.51	0.18	0.41	-0.22	0.14	-															
Faultline Strength																					
Deep-Level Diversity																					
1.5 _E Extraversion, Mean	4.57	0.68	0.25	0.01	-0.32	0.09	-														
1.5 _N Neuroticism, Mean	4.29	0.61	-0.03	-0.24	-0.13	0.02	-0.22	-													
1.5 _O Openness to	5.30	0.50	0.05	0.07	0.08	-0.05	0.11	-0.39	-												
Experience, Mean																					
1.5 _C Conscientiousness, Mean	5.31	0.66	0.08	-0.49	-0.09	0.17	0.12	0.18	-0.10	-											
1.5 _A Agreeableness, Mean	5.70	0.39	0.00	-0.18	0.16	0.04	0.08	0.05	0.15	0.34	-										
1.6 Competency, Mean	4.43	0.74	-0.03	0.27	-0.11	-0.19	0.16	-0.02	0.01	-0.15	-0.14	-									
1.7 _E Extraversion, S.D.	1.23	0.52	0.21	-0.03	-0.07	-0.00	-0.03	0.03	0.02	0.08	-0.13	-0.02	-								
1.7 _N Neuroticism, S.D.	1.08	0.50	0.10	0.23	-0.22	0.12	0.13	0.04	0.02	-0.06	0.00	0.08	0.07	-							
1.7 _O Openness to Experience, S.D.	0.93	0.46	0.11	-0.07	-0.14	0.05	0.25	0.06	-0.45	0.06	-0.15	0.02	0.14	0.04	-						
1.7 _C Conscientiousness, S.D.	0.91	0.42	0.33	0.04	0.23	0.21	-0.01	0.21	0.02	0.05	0.09	-0.04	0.15	-0.10	0.05	-					
1.7 _A Agreeableness, S.D.	0.77	0.30	-0.00	0.05	0.05	0.14	-0.08	0.13	-0.04	-0.07	-0.30	0.08	-0.13	0.12	0.05	0.27	-				
1.8 Competency, S.D.	1.13	0.47	0.25	-0.17	0.08	0.30	-0.06	-0.08	0.21	0.19	0.02	-0.14	0.14	0.02	0.11	0.34	0.30	-			
Team Emergent State																					
1.9 Team Psychological Safety	6.01	0.41	-0.14	0.01	0.08	-0.26	0.10	-0.17	0.20	0.13	0.19	-0.03	0.02	-0.32	-0.09	-0.20	0.06	0.05	-		
Midpoint																					
2 Team Psychological Safety	5.85	0.52	-0.16	-0.03	-0.06	-0.22	0.18	-0.06	0.06	0.09	0.06	-0.05	-0.02	-0.24	-0.01	-0.13	0.01	0.01	0.58	-	
End																					
3 Team Psychological Safety	5.81	0.56	-0.09	-0.14	-0.12	-0.22	0.09	-0.04	0.13	0.29	0.23	-0.26	0.08	-0.00	-0.15	-0.18	-0.19	0.06	0.52	0.66	-

M and SD are used to represent mean and standard deviation, respectively. Correlations with $r > 0.25$ are significant at $p < 0.05$ and are bold printed.

correlate with each other, such as age and tenure. However, as they share the same variance, they might shadow possible effects due to multicollinearity. Because the difference in team members' theoretical knowledge of psychology in terms of their tenure was more important for the successful completion of the task than their age differences, we included tenure (semester term) in our final model calculations.

Finally, our teams were all relatively small (with 3–5 members each and one dyad), and team size did not correlate with PS. Because team faultlines and team size correlated significantly, we excluded team size as a control variable when testing our hypotheses referring to initial levels of PS. When testing for changes in PS, however, we included team size again, because decisions and conflicts might differ in teams with three members compared to four or five members, when majorities and coalitions can be formed.

Analytical procedure

We used Linear Growth Curve Modeling in R (R Development Core Team, 2012), a multilevel approach that allows us to study changes (Bliese and Ployhart, 2002; Bliese, 2016). This way, we can predict initial levels (the intercepts) and changes (the slopes) of PS over time. We had two levels of analysis with measurement points on the lowest level (level 1), nested in teams (level 2). We restructured the dataset according to a new variable, namely Measurement Time with the attributes 1 (initial levels), 2 (midpoint), and 3 (end). In our model-building approach (Bliese and Ployhart, 2002), we started by calculating a null model and increased the complexity of the models in a stepwise way.

In the first step, we examined the base growth curve and fitted the model's structure. In the second step, we predicted initial levels of PS by adding demographic faultline strength as a focal predictor to the model, when controlling for the components of the faultlines, namely diversity in gender and tenure. In the third step, we predicted initial levels of PS by adding deep-level diversity in terms of personality (mean values and SDs) for extraversion and neuroticism step by step to the model, again controlling for diversity in gender and tenure. In the fourth step, we predicted changes in PS by adding deep-level diversity in terms of the hypothesized personality characteristics and competencies (mean values and SDs) with the respective interactions with time to the baseline model, when controlling for team size as well as diversity in gender and tenure. When testing our hypotheses, we added both the mean and the within-team deviation of the deep-level predictors to the model to accurately consider their dependency when interpreting the results (e.g., we combined mean values and controlled for deviation, and vice versa). All predictors on the team level were grand-mean centered, except for demographic faultline strength and team diversity regarding gender, which were *z*-standardized.

Fitting the model's structure

First, we ran an unconstrained (null) model and found confirmation for our multilevel structure. In this model, 25% of the variance in PS could be explained by team properties. Second, we calculated a random intercept, fixed slope model by adding measurement time, a level-1 variable, to the model. A comparison of the models' deviations showed that a random intercept, fixed slopes model had a significantly better fit compared to the null model [Δ -2LL (1) = 11.71, $p < 0.0006$]. In the next step, we tested a random intercept, random slopes model, allowing slopes to vary as well. Again, this model had a significantly better fit compared to the former model [Δ -2LL (5) = 26.30, $p < 0.0001$]. Third, we tested for autocorrelation among measurement points. This model had no significantly better fit compared to the former model [Δ -2LL (1) = 0.01, $p = 0.93$]. Therefore, we allowed intercepts and slopes to vary in further calculations without controlling for autocorrelation in our data. We report our base growth curve model as Model 1 (Table 2).

Hypothesis tests

All tested models with the respective coefficients, variances and fit indices are presented in tables. Table 2 shows the test of predictors for initial levels in PS, and Tables 3–6 show the tests of predictors for changes in PS. Table 7 presents an overview of the tested hypotheses and respective support.

In Model 1 (Table 2), we found that time significantly predicted changes in PS with PS decreasing over time ($b = 0.09$, $p = 0.002$). Thus, we found support for H1. In Model 2, we predicted initial levels (intercept) of PS with demographic faultline strength and the control variables team diversity regarding gender (Blau index) and semester term (within-team SD). Initial levels depended significantly on faultline strength: The stronger the demographic faultline, the lower teams assessed initial levels of PS to be. Thus, we found support for H2.

In Model 3, we predicted initial levels (intercept) of PS by adding deep-level diversity in terms of team member extraversion (H3a regarding mean values and H3b regarding SDs), when controlling for effects of demographic faultline strength, gender diversity, and tenure diversity. We found that neither mean levels of team member extraversion nor diversity in team member extraversion predicted initial levels of PS (no support for H3a and H3b).

We repeated this procedure for team member neuroticism separately in Model 4. Contrary to our assumption, it was not the mean value, but diversity in neuroticism (the within-team SD) that affected initial levels of PS: The more diverse the team was in neuroticism, the lower teams assessed initial levels of PS to be (no support for H3c). In Model 5, the combined model with all tested predictors is depicted.

In Model 6 (Table 3), we present our baseline model for predicting changes (the slope) in PS including the control variables

TABLE 2 Linear growth curve model ($n=699$ Measurements Nested in Three Time Points, in 233 Members in 61 Teams) for initial levels of team psychological safety.

Parameter	Model 1 _{Time}			Model 2 _{Faultlines}			Model 3 _{Extraversion}			Model 4 _{Neuroticism}			Model 5 _{Combined}		
	<i>b</i>	95% CI	<i>t</i>	<i>b</i>	95% CI	<i>t</i>	<i>b</i>	95% CI	<i>t</i>	<i>b</i>	95% CI	<i>t</i>	<i>b</i>	95% CI	<i>t</i>
Fixed Effects Level 1 (Time)															
Intercept	6.09	[5.96, 6.22]	91.30	6.09	[5.96, 6.22]	90.42	6.09	[5.96, 6.23]	91.30	6.09	[5.97, 6.22]	96.35	6.09	[5.97, 6.21]	96.46
Time	-0.09	[-0.15, -0.03]	-3.09	-0.09	[-0.15, -0.03]	-3.11	-0.09	[-0.15, -0.03]	-3.09	-0.09	[-0.15, -0.03]	-3.08	-0.09	[-0.15, -0.03]	-3.10
Fixed Effects Level 2 (Teams)															
Surface-Level Diversity															
Gender, Blau's Index				-0.04	[-0.15, 0.07]	-0.76	-0.01	[-0.12, 0.09]	-0.23	0.02	[-0.09, 0.13]	0.41	-0.02	[-0.13, 0.10]	-0.32
Tenure, <i>S.D.</i>				0.06	[-0.12, 0.23]	0.64	0.06	[-0.12, 0.24]	0.70	-0.05	[-0.22, 0.13]	-0.53	0.05	[-0.14, 0.24]	0.55
Demographic Faultline Strength				-0.11	[-0.21, -0.002]	-2.03							-0.09	[-0.20, 0.01]	-1.72
Team Personality															
Extraversion, <i>Mean</i>							0.12	[-0.04, 0.29]	1.53				0.14	[-0.03, 0.30]	1.69
Extraversion, <i>S.D.</i>							0.06	[-0.14, 0.26]	0.62				0.07	[-0.12, 0.26]	0.70
Neuroticism, <i>Mean</i>										-0.11	[-0.27, 0.06]	-1.29	-0.08	[-0.25, 0.09]	-0.95
Neuroticism, <i>S.D.</i>										-0.25	[-0.46, -0.04]	-2.42	-0.22	[-0.44, -0.01]	-2.10
Variance Components															
Level 2 Intercept (Teams)	0.33	[0.20, 0.52]		0.33			0.33			0.28	[0.16, 0.50]		0.28		
Level 2 Slope (Teams)	0.17	[0.12, 0.24]		0.17			0.17			0.17	[0.12, 0.24]		0.17		
Intercept-Slope Covariance (Teams)	-0.61	[-0.85, -0.16]		-0.61			-0.60			-0.57	[-0.85, -0.04]		-0.59		
Level 1 Intercept (Time)	0.42	[0.31, 0.55]		0.42			0.43			0.43	[0.30, 0.63]		0.43		
Level 1 Slope (Time)	0.05	[0.01, 0.21]		0.06			0.06			0.06	[0.005, 0.71]		0.06		
Intercept-Slope Covariance (Time)	0.74	[-0.96, 0.999]		0.55			0.53			0.47	[-0.99, 0.999]		0.46		
Within-Team Variance	0.44	[0.41, 0.48]		0.43			0.43			0.43	[0.38, 0.50]		0.43		
<i>AIC</i>	1294.68			1305.53			1315.26			1310.33			1316.18		
<i>BIC</i>	1335.37			1364.20			1378.48			1373.56			1392.80		
$-2LL$	1276.67			1279.53			1287.26			1282.33			1282.19		
Overall Pseudo- R^2_m	0.01			0.03			0.02			0.04			0.07		

All models are random intercepts, random slopes, and $-2 \ln = -2 \log$ likelihood values. For Models 2, 3 and 5, only confidence intervals for fixed effects are reported; confidence intervals could not be calculated in R due to a non-positive definite in the variance-covariance matrix. Values significant according to 95% CI are bold.

TABLE 3 Linear Growth Curve Model ($n=699$ measurements nested in three time points, in 233 members in 61 teams) for changes in team psychological safety predicted by team personality and competency (Models 6 and 7).

<i>Parameter</i>	<i>Model 6</i>			<i>Model 7_{Neuroticism}</i>		
	<i>b</i>	<i>95% CI</i>	<i>t</i>	<i>b</i>	<i>95% CI</i>	<i>t</i>
Fixed Effects Level 1 (Time)						
Intercept	6.09	[5.95, 6.23]	89.80	6.09	[5.97, 6.22]	95.69
Time	-0.09	[-0.15, -0.03]	-3.10	-0.09	[-0.15, -0.03]	-3.12
Fixed Effects Level 2 (Teams)						
Team Size	-0.08	[-0.29, 0.13]	-0.77	-0.06	[-0.26, 0.14]	-0.63
Surface-Level Diversity						
Gender, Blau's Index	-0.01	[-0.12, 0.10]	-0.18	0.02	[-0.09, 0.13]	0.30
Tenure, <i>S.D.</i>	-0.003	[-0.17, 0.18]	0.04	-0.05	[-0.23, 0.12]	-0.60
Big Five and Team Competency						
Neuroticism, <i>Mean</i>				-0.13	[-0.35, 0.08]	-1.25
Openness to Experience, <i>Mean</i>						
Conscientiousness, <i>Mean</i>						
Agreeableness, <i>Mean</i>						
Competency, <i>Mean</i>						
Neuroticism, <i>S.D.</i>				-0.39	[-0.66, -0.12]	-2.93
Openness to Experience, <i>S.D.</i>						
Conscientiousness, <i>S.D.</i>						
Agreeableness, <i>S.D.</i>						
Competency, <i>S.D.</i>						
Cross-Level Interactions						
Time×Neuroticism, <i>Mean</i>				0.02	[-0.08, 0.11]	0.35
Time×Openness to Experience, <i>Mean</i>						
Time×Conscientiousness, <i>Mean</i>						
Time×Agreeableness, <i>Mean</i>						
Time×Competency, <i>Mean</i>						
Time×Neuroticism, <i>S.D.</i>				0.11	[-0.01, 0.23]	1.78
Time×Openness to Experience, <i>S.D.</i>						
Time×Conscientiousness, <i>S.D.</i>						
Time×Agreeableness, <i>S.D.</i>						
Time×Competency, <i>S.D.</i>						
Variance Components						
Level 2 Intercept (Teams)	0.34	[0.22, 0.54]		0.29	[0.17, 0.50]	
Level 2 Slope (Teams)	0.17	[0.12, 0.24]		0.17	[0.12, 0.24]	
Intercept-Slope Covariance (Teams)	-0.59	[-0.84, -0.15]		-0.56	[-0.84, -0.05]	
Level 1 Intercept (Time)	0.43	[0.32, 0.59]		0.43	[0.30, 0.62]	
Level 1 Slope (Time)	0.06	[0.01, 0.39]		0.06	[0.01, 0.48]	
Intercept-Slope Covariance (Time)	0.45	[-0.93, 0.99]		0.75	[-0.988, 0.999]	
Within-Team Variance	0.43	[0.38, 0.49]		0.43	[0.37, 0.50]	
<i>AIC</i>	1311.98			1323.27		
<i>BIC</i>	1370.71			1399.97		
<i>-2LL</i>	1285.98			1289.27		
Overall Pseudo- R^2_m	0.01			0.04		

Note. All models are random intercepts, random slopes, and $-2 LL = -2$ log likelihood values. Values significant according to 95% *CI* are bold.

team size, gender diversity, and tenure diversity. In Model 7 (Table 3), we predicted changes in PS with diversity in team member neuroticism (H3d *SDs*). We, therefore, added the

respective interaction (i.e., the product terms) of this predictor with time and found that changes in PS did not depend on diversity in neuroticism (no support for H3d). In the following, we repeated

Table 4 Linear Growth Curve Model ($n=699$ measurements nested in three time points, in 233 members in 61 teams) for changes in team psychological safety predicted by team personality and competency (Models 8 and 9).

<i>Parameter</i>	<i>Model 8</i> _{Openness}			<i>Model 9</i> _{Conscientiousness}		
	<i>b</i>	<i>95% CI</i>	<i>t</i>	<i>b</i>	<i>95% CI</i>	<i>t</i>
Fixed Effects Level 1 (Time)						
Intercept	6.09	[5.96, 6.23]	90.52	6.09	[5.96, 6.22]	89.83
Time	-0.09	[-0.15, -0.04]	-3.12	-0.09	[-0.15, -0.03]	-3.11
Fixed Effects Level 2 (Teams)						
Team Size	-0.09	[-0.30, 0.12]	-0.89	-0.02	[-0.24, 0.20]	-0.18
Surface-Level Diversity						
Gender, Blau's Index	-0.02	[-0.12, 0.09]	-0.30	0.03	[-0.09, 0.15]	0.47
Tenure, <i>S.D.</i>	-0.01	[-0.18, 0.17]	-0.07	0.04	[-0.14, 0.22]	0.41
Big Five and Team Competency						
Neuroticism, <i>Mean</i>						
Openness to Experience, <i>Mean</i>	0.25	[-0.06, 0.55]	1.63			
Conscientiousness, <i>Mean</i>				0.002	[-0.23, 0.23]	0.02
Agreeableness, <i>Mean</i>						
Competency, <i>Mean</i>						
Neuroticism, <i>S.D.</i>						
Openness to Experience, <i>S.D.</i>	0.15	[-0.19, 0.49]	0.86			
Conscientiousness, <i>S.D.</i>				-0.18	[-0.53, 0.17]	-1.04
Agreeableness, <i>S.D.</i>						
Competency, <i>S.D.</i>						
Cross-Level Interactions						
Time × Neuroticism, <i>Mean</i>						
Time × Openness to Experience, <i>Mean</i>	-0.04	[-0.18, 0.09]	-0.65			
Time × Conscientiousness, <i>Mean</i>				0.08	[-0.01, 0.17]	1.65
Time × Agreeableness, <i>Mean</i>						
Time × Competency, <i>Mean</i>						
Time × Neuroticism, <i>S.D.</i>						
Time × Openness to Experience, <i>S.D.</i>	-0.09	[-0.24, 0.06]	-1.18			
Time × Conscientiousness, <i>S.D.</i>				-0.02	[-0.16, 0.13]	-0.22
Time × Agreeableness, <i>S.D.</i>						
Time × Competency, <i>S.D.</i>						
Variance Components						
Level 2 Intercept (Teams)	0.33	[0.21, 0.53]		0.34	[0.21, 0.54]	
Level 2 Slope (Teams)	0.17	[0.12, 0.25]		0.17	[0.12, 0.24]	
Intercept-Slope Covariance (Teams)	-0.60	[-0.85, 0.15]		-0.62	[-0.85, 0.18]	
Level 1 Intercept (Time)	0.43	[0.31, 0.59]		0.43	[0.32, 0.57]	
Level 1 Slope (Time)	0.06	[0.01, 0.42]		0.06	[0.01, 0.26]	
Intercept-Slope Covariance (Time)	0.52	[-0.98, 0.998]		0.52	[-0.86, 0.98]	
Within-Team Variance	0.43	[0.39, 0.49]		0.44	[0.39, 0.48]	
<i>AIC</i>	1327.67			1326.38		
<i>BIC</i>	1404.37			1403.08		
<i>-2LL</i>	1293.67			1292.38		
Overall Pseudo- R^2_m	0.03			0.04		

Note. All models are random intercepts, random slopes, and $-2 LL = -2 \log$ likelihood values. Values significant according to 95% CI are bold.

the procedure separately for our deep-level characteristics according to our hypotheses. In Model 8 (Table 4), we predicted changes in PS with team member openness (H3e mean values) and

found that changes in PS did not depend on team member openness (no support for H3e). In Model 9 (Table 4), we predicted changes in PS with diversity in team member conscientiousness

Table 5 Linear Growth Curve Model ($n=699$ measurements nested in three time points, in 233 members in 61 teams) for changes in team psychological safety predicted by team personality and competency (Models 10 and 11).

<i>Parameter</i>	<i>Model 10</i> _{Agreeableness}			<i>Model 11</i> _{Competency}		
	<i>b</i>	95% CI	<i>t</i>	<i>b</i>	95% CI	<i>t</i>
Fixed Effects Level 1 (Time)						
Intercept	6.09	[5.96, 6.22]	90.54	6.09	[5.96, 6.23]	88.67
Time	-0.09	[-0.15, -0.04]	-3.17	-0.09	[-0.15, -0.04]	-3.18
Fixed Effects Level 2 (Teams)						
Team Size	-0.09	[-0.30, 0.12]	-0.84	-0.09	[-0.31, 0.13]	-0.85
Surface-Level Diversity						
Gender, Blau's Index	0.007	[-0.10, 0.12]	0.13	0.003	[-0.11, 0.12]	0.07
Tenure, <i>S.D.</i>	-0.03	[-0.21, 0.15]	-0.31	-0.02	[-0.20, 0.17]	-0.17
Big Five and Team Competency						
Neuroticism, <i>Mean</i>						
Openness to Experience, <i>Mean</i>						
Conscientiousness, <i>Mean</i>						
Agreeableness, <i>Mean</i>	0.18	[-0.20, 0.56]	0.93			
Competency, <i>Mean</i>				0.07	[-0.13, 0.27]	-0.72
Neuroticism, <i>S.D.</i>						
Openness to Experience, <i>S.D.</i>						
Conscientiousness, <i>S.D.</i>						
Agreeableness, <i>S.D.</i>	0.39	[-0.09, 0.87]	1.64			
Competency, <i>S.D.</i>				0.02	[-0.29, 0.33]	0.12
Cross-Level Interactions						
Time × Neuroticism, <i>Mean</i>						
Time × Openness to Experience, <i>Mean</i>						
Time × Conscientiousness, <i>Mean</i>						
Time × Agreeableness, <i>Mean</i>	0.03	[-0.13, 0.18]	0.31			
Time × Competency, <i>Mean</i>				-0.08	[-0.16, -0.001]	-1.98
Time × Neuroticism, <i>S.D.</i>						
Time × Openness to Experience, <i>S.D.</i>						
Time × Conscientiousness, <i>S.D.</i>						
Time × Agreeableness, <i>S.D.</i>	-0.21	[-0.41, -0.0001]	-1.97			
Time × Competency, <i>S.D.</i>				0.01	[-0.11, 0.14]	0.23
Variance Components						
Level 2 Intercept (Teams)	0.33			0.35		
Level 2 Slope (Teams)	0.16			0.17		
Intercept-Slope Covariance (Teams)	-0.57			-0.59		
Level 1 Intercept (Time)	0.43			0.43		
Level 1 Slope (Time)	0.06			0.06		
Intercept-Slope Covariance (Time)	0.54			0.53		
Within-Team Variance						
<i>AIC</i>	1322.75			1329.18		
<i>BIC</i>	1399.45			1405.88		
<i>-2LL</i>	1288.75			1295.18		
Overall Pseudo- R^2_m	0.03			0.03		

Note. All models are random intercepts, random slopes models, $-2 LL = -2 \log$ likelihood values. For Models 10 and 11, only confidence intervals for fixed effects are reported; confidence intervals for the variance components could not be calculated in R due to a non-positive definite in the variance-covariance matrix. Values significant according to 95% CI are bold.

(H3f SDs) and found that changes did not depend on diversity in team member conscientiousness (no support for H3f).

In Model 10 (Table 5), we predicted changes in PS with team member agreeableness (H3g mean values) and diversity in

agreeableness (H3h SDs) and found that changes in PS depended on diversity in team member agreeableness (support for H3h, but no support for H3g). The more diverse team members were in agreeableness, the more did PS decline over

Table 6 Linear Growth Curve Model ($n=699$ measurements nested in three time points, in 233 members in 61 teams) for changes in team psychological safety predicted by team personality and competency (Model 12).

<i>Parameter</i>	<i>Model 12</i> _{Combined}		
	<i>b</i>	<i>95% CI</i>	<i>t</i>
Fixed Effects Level 1 (Time)			
Intercept	6.09	[5.97, 6.22]	96.24
Time	-0.09	[-0.15, -0.04]	-3.27
Fixed Effects Level 2 (Teams)			
Team Size	-0.04	[-0.26, 0.19]	-0.33
Surface-Level Diversity			
Gender, Blau's Index	0.08	[-0.05, 0.21]	1.27
Tenure, <i>S.D.</i>	-0.07	[-0.27, 0.12]	-0.75
Big Five and Team Competency			
Neuroticism, <i>Mean</i>	-0.07	[-0.32, 0.18]	-0.56
Openness to Experience, <i>Mean</i>	0.22	[-0.11, 0.56]	1.35
Conscientiousness, <i>Mean</i>	0.03	[-0.20, 0.26]	0.26
Agreeableness, <i>Mean</i>	0.25	[-0.13, 0.64]	1.32
Competency, <i>Mean</i>	0.05	[-0.14, 0.23]	0.51
Neuroticism, <i>S.D.</i>	-0.40	[-0.68, -0.13]	-2.94
Openness to Experience, <i>S.D.</i>	0.18	[-0.16, 0.52]	1.08
Conscientiousness, <i>S.D.</i>	-0.15	[-0.52, 0.22]	-0.79
Agreeableness, <i>S.D.</i>	0.46	[-0.04, 0.97]	1.86
Competency, <i>S.D.</i>	-0.07	[-0.41, 0.26]	-0.43
Cross-Level Interactions			
Time × Neuroticism, <i>Mean</i>	0.02	[-0.09, 0.13]	0.30
Time × Openness to Experience, <i>Mean</i>	-0.05	[-0.20, 0.10]	-0.65
Time × Conscientiousness, <i>Mean</i>	0.06	[-0.04, 0.16]	1.15
Time × Agreeableness, <i>Mean</i>	-0.04	[-0.21, 0.13]	-0.45
Time × Competency, <i>Mean</i>	-0.07	[-0.15, 0.01]	-1.66
Time × Neuroticism, <i>S.D.</i>	0.11	[-0.006, 0.23]	1.87
Time × Openness to Experience, <i>S.D.</i>	-0.10	[-0.25, 0.04]	-1.38
Time × Conscientiousness, <i>S.D.</i>	-0.02	[-0.17, 0.14]	-0.20
Time × Agreeableness, <i>S.D.</i>	-0.22	[-0.44, 0.006]	-1.91
Time × Competency, <i>S.D.</i>	0.06	[-0.08, 0.21]	0.83
Variance Components			
Level 2 Intercept (Teams)	0.28		
Level 2 Slope (Teams)	0.16		
Intercept-Slope Covariance (Teams)	-0.52		
Level 1 Intercept (Time)	0.43		
Level 1 Slope (Time)	0.06		
Intercept-Slope Covariance (Time)	0.57		
Within-Team Variance			
<i>AIC</i>	1382.87		
<i>BIC</i>	1530.96		
<i>-2LI</i>	1316.86		
Overall Pseudo- R^2_m	0.09		

Note. All models are random intercept, random slopes models, $-2 LL = -2$ log-likelihood value. For Model 12, only confidence intervals for fixed effects are reported; confidence intervals for the variance components could not be calculated in *R* due to a non-positive definite in the variance-covariance matrix. Values significant according to 95% *CI* are bold.

time. We plotted this interactive effect of the predictor team member diversity in agreeableness with time and used the web tool provided by Preacher et al. (2006) to test for simple slopes. Figure 1 illustrates how PS changes as a function of team

member diversity in agreeableness and time. Psychological safety decreased in teams that were diverse in agreeableness [$\gamma = -0.15$ (0.03), $p < 0.001$] and remained stable in teams with homogeneous agreeableness.

TABLE 7 Hypotheses and support for initial levels and changes in team psychological safety.

	Support?
Hypothesis on the Effect of Time on PS	
H 1: As time passes in a team project, PS decreases.	Supported
Initial Level Hypotheses	
H 2: Demographic faultline strength is negatively associated with initial levels of PS.	Supported
H 3a: The higher a team's mean level in extraversion, the higher the initial levels of PS will be in the team.	Not supported
H 3b: The more diverse team members are in extraversion, the lower initial levels of PS will be in the team.	Not supported
H 3c: The higher a team's mean level in neuroticism is, the lower initial levels of PS will be.	Not supported
Change Level Hypotheses	
H 3d: The more diverse team members are in neuroticism, the more PS will decrease over time.	Not supported
H 3e: The higher a team's mean level in openness to experience, the more PS within the team will increase over time.	Not supported
H 3f: The more diverse team members are in conscientiousness, the more PS will decrease over time.	Not supported
H 3g: The higher a team's mean level of agreeableness, the more PS within the team will increase over time.	Not Supported
H 3h: The more diverse team members are in agreeableness, the more PS will decrease over time.	Supported
H 4: Team member task-relevant competencies predict changes in PS: The greater a team's competency, the more PS will decrease over time.	Supported

In Model 11 (Table 5), we predicted changes in PS with team member competencies (H4 mean values) and found that changes in PS depended on team members' competencies (support for H4). The more competent team members were, the more did PS decline over time. We plotted the interactive effect of the predictor team member competencies with time to test for simple slopes (Preacher et al., 2006). Figure 2 illustrates how PS changes as a function of team member competencies and time. Psychological safety decreased in teams with relatively high member competencies [$\gamma = -0.15$ (0.04), $p = 0.0003$] and remained stable in teams with relatively low competencies. Residuals for all reported models were normally distributed so that the requirements for computing mixed models were met. We estimated the overall variance explanation of the models with the pseudo- R^2 for generalized mixed-effect models (Nakagawa and Schielzeth, 2013).

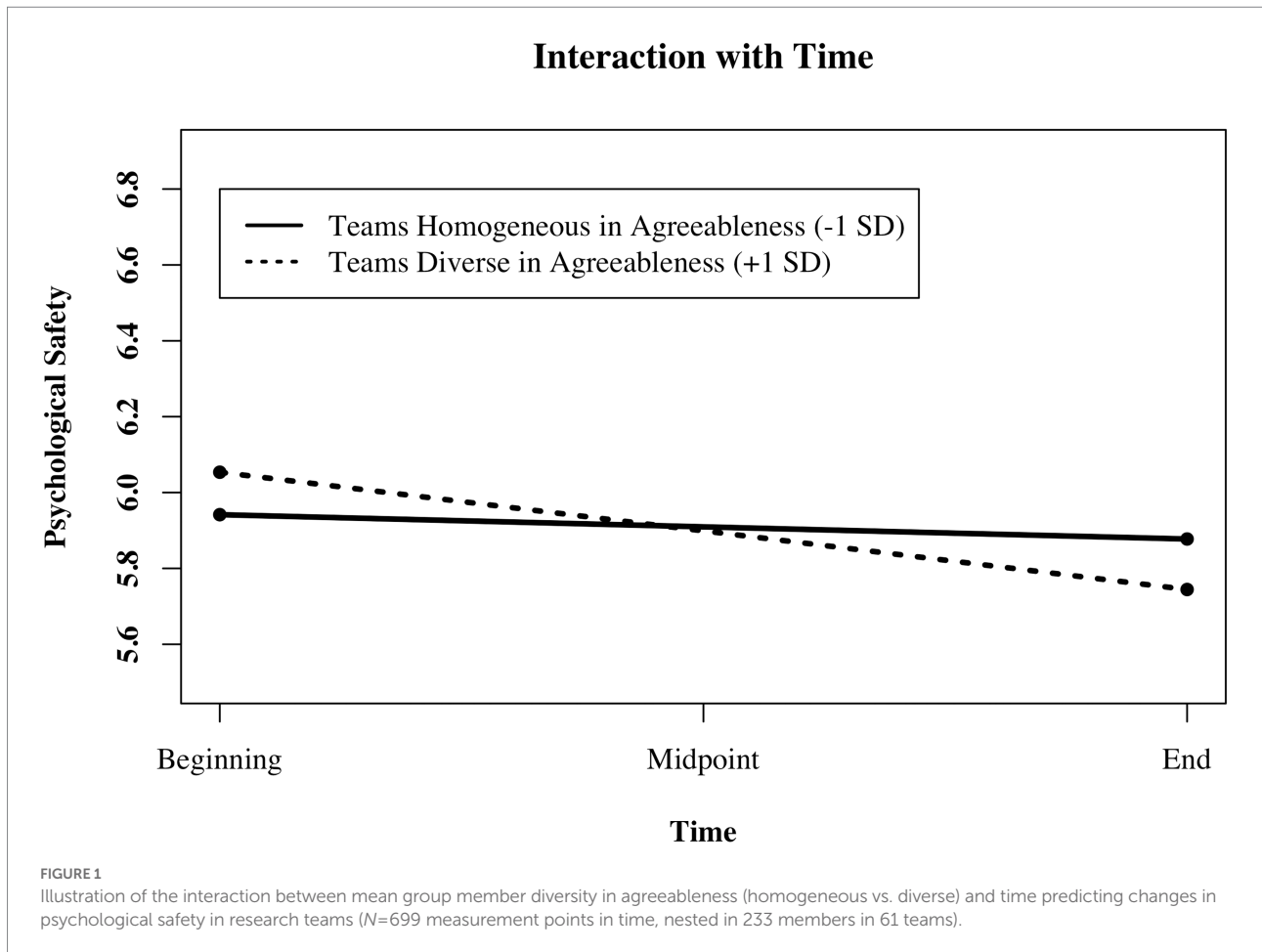
In Model 12 (Table 6), all control variables and predictors are included simultaneously, for the reader's information. In summary, we found that PS significantly decreased over time. Initial levels of PS depended on demographic faultline strength and team members' diversity in neuroticism: The stronger the faultlines were, and the more diverse team members were in their neuroticism, the lower they assessed initial levels of PS to be. Changes in PS depended on team member diversity in agreeableness and on team member competencies. PS decreased more strongly over time when team members were diverse in agreeableness and assessed their task-related competencies to be relatively high.

Discussion

This study serves to deepen our understanding of the dynamics of PS. It sheds a light on how PS forms and develops over time in the context of member diversity—from the very

beginning to the end of project teams. In a sample of self-managed teams, we found that PS decreased over the course of a team project. Initial levels in PS were especially low when teams had strong faultlines and team members were diverse in neuroticism. Changes in PS depended on team member agreeableness and competencies such that PS decreased to a stronger extent when team members varied in their agreeableness and assessed their task-related competencies to be high. We start our discussion by addressing the development of PS found in our study: Most teams started at initially high levels, which significantly decreased over time.

The decrease in PS over time is in line with the descriptive statistics reported by Schulte et al. (2012). There are two explanations for the decline in PS: First, at the beginning of a team project, it is commonly acceptable to allow many questions from team members to form a shared mental model of the task and a transactive memory system about members' areas of expertise (Gockel and Brauner, 2013). Also, at this early stage, not many problems need to be addressed because they have not occurred or are not obvious yet. However, after this initial period, most team members may think that the task is clear and that first actions should be taken. In this phase, additional questions and feedback could be perceived as rather irritating and may slow down the work. Yet, because the first sessions set standards, PS is very high at the beginning, followed by a consecutive decrease over time. At the finish line, the final critical consideration of one's teamwork is important for perceived team performance; however, the team may already be too "closed off" to reopen the discussion or allow feedback loops, and this closing phenomenon is reflected in a decrease in PS. Two findings support this consideration: Teams with a high need for closure, which might be triggered by time pressure, are generally less open to critical discussion (Pierro et al., 2003). In addition, the acceptance of novel contributions decreases as time passes in teams facing a deadline (Ford and Sullivan,



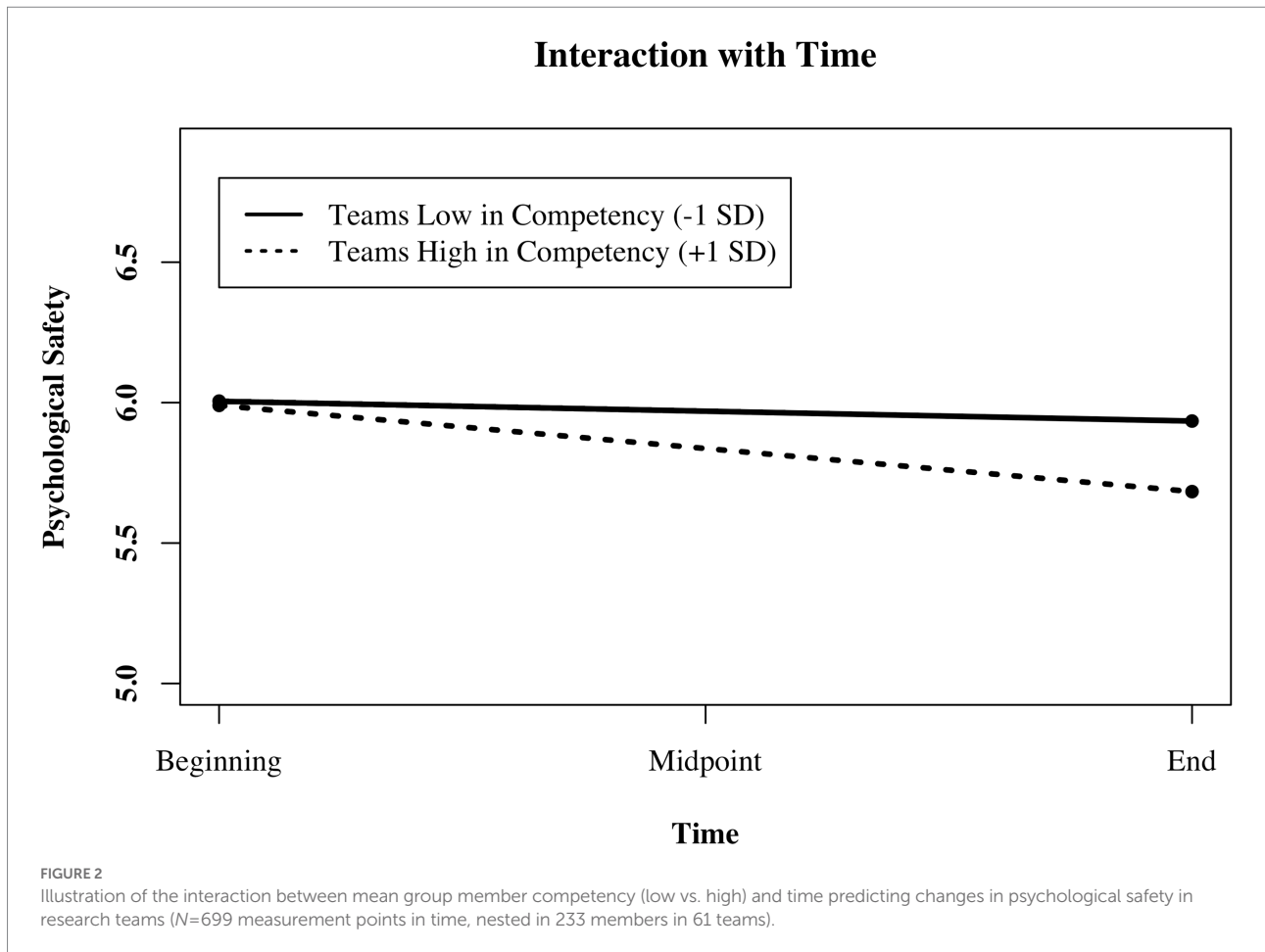
2004). Particularly around the temporal midpoint, teams appear to experience a transition between two work orientations that provides a starting point for interventions to increase or stabilize PS (Ford and Sullivan, 2004; Knight, 2015). Our results support a continuous decrease over a nonlinear decrease. The team's assessment of risk-taking before asking questions or seeking support (and not just task information) decreases from beginning to end across all teams depending on the time passed. Altogether, this suggests that PS as a supportive learning environment has a boundary concerning time or specific temporal milestones—at least in teams with a clear deadline.

The second explanation for the decline in PS refers to a phenomenon found in trust research (Lewicki et al., 2006). It is reasonable that, at their first meeting, team members evaluate their PS based on other criteria than on their experience with risky behavior in the team, because they may not yet have had enough experience with this specific team. For instance, they could have used the teams' structure (self-managed teams without formal leaders) or other status signals such as gender, age, or floor-gaining behavior for first impressions. Thus, members give a leap of faith and PS to the team, which is to be evaluated later during teamwork (reflected in lower team agreements at midpoint).

Perhaps, up until this point, team members shared the anticipated credence of PS. Enthusiasm at the start as a kind of initial ignition and its reduction over time appears not to be unusual for teamwork (see also, e.g., the romance of teams or the honeymoon-hangover effect from research on newcomers' job satisfaction, Allen and Hecht, 2004; Boswell et al., 2005).

In our study, initial levels of PS depended on demographic faultline strength. Our findings complement previous research in this area. For example, teachers reported lower individual levels of PS in the presence of strong team faultlines compared to weak faultlines (Gerlach and Gockel, 2018). The present study is the first to demonstrate the temporal link between team faultlines and initial levels of PS. This is in line with a discussion by Lau and Murnighan (2005) that strong demographic faultlines have a "direct and potentially pervasive effect" (p. 655) on team processes right from the start. As such, team members are more likely to identify with their subgroup and share sensitive information with close members of their subgroup instead the entire team, which leads to more isolated communication and information sharing and an overall reduced PS (Lau and Murnighan, 2005).

Different from our expectation, team members' diversity in neuroticism (and not mean values) predicted the formation of



PS. The more diverse team members were in neuroticism, the lower teams assessed their PS at the beginning to be. This finding adds to former research that stressed the mean value compared to diversity in neuroticism to be important for early levels of PS (Ostermeier et al., 2020). One explanation for our finding could be that teams diverse in neuroticism scores felt initially uncomfortable in the team, as members perceived others to be different in their way of communication. This might hamper a smooth collaboration and all communication behaviors that are associated with PS right from the start (Barrick et al., 1998). Nevertheless, this relation should be interpreted with caution because we did not hypothesize the effect to occur. It needs future research to test its robustness.

If neuroticism is an important concept for communication behavior, especially if individuals have extreme values, future studies might explore the influence of single members on the creation of PS. For future research, the use of other operationalizations such as taking the minimum or maximum or skewness of a trait into consideration might bring new insights into the possible dynamics of PS (see Bell, 2007; Ostermeier et al., 2020). Still, it is important to note that only neuroticism, and not extraversion, was found to predict the formation of PS. Future research might focus on attitudes that are important for team

performance such as team members' orientation toward teamwork (Ulloa and Adams, 2004).

Changes in PS over time depended on only one personality characteristic, namely team members' agreeableness. The more team members varied in their agreeableness, the more PS decreased throughout the project. In other words, compared to teams homogeneous in agreeableness, teams diverse in agreeableness started with relatively high PS but decreased more steeply over time. We assume that in these diverse teams, the more agreeable members create a trustful, warm, and tolerant atmosphere, in which the more disagreeable members feel comfortable taking risks and pushing for learning. This behavior could be beneficial for PS at the beginning of the team interaction but lead to more disagreements, conflicts, and a deterioration in PS over time. In teams homogeneous in agreeableness, PS seems rather stable throughout the project. This might create a holding environment for risk-taking behavior. Our study is the first to test the link between team member agreeableness and PS, to our knowledge. The personality characteristic agreeableness has lately received increased attention as an important predictor of team cohesion and team performance, especially if teams work face-to-face (Bradley et al., 2013). Recent research showed that mean levels

of agreeableness in teams predicted an increase in team cohesion (Acton et al., 2020). Our results add to these findings by stressing team members' variation over mean values of agreeableness for the development of PS. A fruitful avenue for future research is to focus on specific behaviors of agreeable and disagreeable team members and their mutual influence over time to tease apart the more positive and negative effects on PS and subsequent learning behavior. In this context, accountability could be a promising element in explaining why also low PS might be beneficial for producing learning behavior (Weiner et al., 2021).

Changes in PS did not depend on team member neuroticism, openness to experience, or conscientiousness in our study. Thus, our findings only partly support the idea that PS unfolds based on team members' personality characteristics. This adds to Edmondson and Mogelof's (2005) findings in innovative teams, where openness and neuroticism (at an individual level) predicted perceptions of PS at the midpoint and the end of the collaboration. It is plausible that if teams understand the value of PS for innovation as an index of team performance, other personality facets might have a greater effect on PS depending on their relevance for task performance. Further, our findings are in line with recent results that could not support the hypothesis that mean values in conscientiousness predict the development of PS (Ostermeier et al., 2020).

As a second predictor for changes, we identified team members' task-related competencies. The higher team members rated their competencies to be at the beginning of the project, the more PS decreased throughout the project. Members of a group with high competencies are more likely to believe that they are right, stick to their positions, and disregard the opinions of others (Innami, 1994). In a previous study, groups that needed to find a consensus to solve the NASA moon landing task showed different communication patterns dependent on the diversity of their task-relevant skills (Innami, 1994). As such, groups with high task-relevant competencies held on to their positions (positional orientation) and did not exchange reasons (reasoning orientation). Time pressure may strengthen this communication behavior as group members fail to understand the importance of explaining task-related issues to each other.

When time pressure increases toward the end of the project, the team has to organize itself with regards to its performance: What needs to be completed by whom and when? These actions require an internal leadership structure and a tacit knowledge of who knows what and has which skills. Personality characteristics such as openness are linked to role development from the beginning (Delice et al., 2019). Others such as extraversion are linked to being perceived as leader-like. The personality of single team members may translate into team emergent states like PS through single behavior or joint sensitivities (LePine et al., 2011) as we found to be true for neuroticism and agreeableness. Dynamics like this can be described in theoretical models as we have suggested above. Nevertheless, there might be critical incidents in situations, task requirements, or other individual

unique constellations that lead to experiences made by the team that have a stronger impact on the development of PS, and especially on the development of different PS perceptions in subgroups. Future research regarding team members' personality characteristics should focus more on the impact of personality and situation settings or consider subgroup dynamics through team members' perception of subgroups and subsequent differences in team members' PS.

Strengths and limitations

We addressed scholarly calls for more holistic research regarding time and changes in teams, more specifically regarding the formation and changes in PS (Roe et al., 2012; Edmondson and Lei, 2014). We refer to theories from diversity research and the model of team faultlines to provide a theoretical frame for how PS forms (Lau and Murnighan, 2005). Another strength of our study is the focus on the team itself, specifically on the effects of deep-level team diversity as antecedents for the development of PS. We help build the nomological net of how personality characteristics at the team level relate to team assessments of PS at different time points of a team project. Further, we add insights to previous research on select personality characteristics at the individual or team level (Edmondson and Mogelof, 2005; Ostermeier et al., 2020). Our study is one of the rare longitudinal team studies that follows teams over an entire project period. Finally, we used sophisticated methods such as multilevel growth curve modeling to predict initial levels and changes in PS based on this longitudinal design to match the dynamic nature of PS.

A few limitations of this study must be pointed out. These teams are characterized by demographic homogeneity (e.g., age, gender, educational background, and experiences in teamwork). A few teams may have been self-formed. Thus, our results should be generalized with caution. The study should be replicated with different kinds of teams in various organizational contexts; thereby controlling for other elements in the team's structure such as hierarchical differences or team member orientation, for example toward achieving the goal of the task (e.g., Acton et al., 2020). Nevertheless, our study's setting fulfilled the requirements for self-managed teamwork. As discussed by Lau and Murnighan (2005), we expect demographic faultline strength to have an even greater impact on PS in teams, whose members are more diverse in terms of gender, age, tenure, and educational backgrounds, than the teams in our study. Our findings are only partly in line with our assumptions on the effects of team personality on the dynamics of PS. Many hypothesized effects we did not find. Therefore, we encourage researchers to address this topic in future research, for example event-based, when critical circumstances or incidents such as voice make PS a priority in the team to identify factors in member personality characteristics that predict changes – such as the decrease over time that we found (*cf.* event system theory by Morgeson et al., 2015, applied to employees' voice in Li and Tangirala, 2021).

Implications

Our findings allow some preliminary conclusions about how and when to intervene more effectively to support the development of PS. Focusing on team properties and team member characteristics helps to predict a psychologically safe atmosphere for teamwork. We found that demographic faultline strength and some facets of deep-level diversity affected initial levels as well as changes in PS. Our study stresses the importance of states and changeable variables, namely task-relevant competencies, group dynamics of only two traits, neuroticism for initial levels, and agreeableness for changes in PS. The fact that PS decreased over time highlights the importance of time-related team interventions at the beginning, midpoint, and end (*cf.* Hackman and Wageman, 2005).

Therefore, at the beginning of a project, we recommend stressing pro-diversity beliefs in combination with enhanced task motivation to overcome the negative effects of team faultline strength (Meyer and Schermuly, 2012). The goal should be that team members develop positive and realistic expectations toward teamwork and will understand the value of diverse perspectives in teams. Later, during task work, we recommend training communication skills by stressing the importance of explaining one's view to the team and not holding on to positions without further elaboration (e.g., communication training for teams by Innami, 1994). This is especially advisable for teams of experts with high abilities because our results showed a decrease in PS in such teams over time. Toward the end, a reminder of the previous intervention (showing that diverse opinions are enriching and a strong focus on work is essential to improve output) could help in keeping the team open and not let it close too early to make final changes that could substantially improve the project. Future studies might explore the dynamics of subgroups from the subjective perspective of team members in bigger and more diverse teams, as well as how changes in PS affect team processes such as team performance, and how a decrease might be repaired or rebuilt for single members, or the team as a whole.

Conclusion

Our study shows that PS changes throughout a team project and how it depends on the team's faultline strength, diversity in team member neuroticism and agreeableness, and team member competencies. Our study contributes to research on PS by drawing from the theory of subgroups to explain team dynamics in PS and by shifting the focus from the leader (who is often seen as one primary source of influence) to the team as a whole. It introduces a theoretical framework for temporal dynamics in PS and provides empirical evidence on how PS forms and develops

over time. It demonstrates the consistently negative effect of demographic faultlines on PS and that only two of the five tested personality characteristics translated to PS on the team level. Finally, our findings stress the importance of time when designing interventions to enhance PS in teams.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

This article has been an intensive collaboration between the two authors. RG and CG jointly conceptualized the study, developed the hypotheses, and conducted the study at Chemnitz University of Technology. RG ran the analyses. RG and CG developed the manuscript together with RG writing the draft and CG providing feedback. Both authors significantly contributed to the article and approved the submitted version.

Funding

This research was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) project number 491193532 and Technische Universität Chemnitz. Any research material or data can be accessed via email from the RG.

Acknowledgments

We would like to thank Bryan L. Bonner and three reviewers for helpful comments on earlier versions of this manuscript.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated

organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Acton, B. P., Braun, M. T., and Foti, R. J. (2020). Built for unity: assessing the impact of team composition on team cohesion trajectories. *J. Bus. Psychol.* 35, 751–766. doi: 10.1007/s10869-019-09654-7
- Allen, N. J., and Hecht, T. D. (2004). The “romance of teams”: toward an understanding of its psychological underpinnings and implications. *J. Occup. Organ. Psychol.* 77, 439–461. doi: 10.1348/0963179042596469
- Anthony, D. B., Wood, J. W., and Holmes, J. G. (2007). Testing sociometer theory: self-esteem and the importance of acceptance for social decision-making. *J. Exp. Soc. Psychol.* 43, 425–432. doi: 10.1016/j.jesp.2006.03.002
- Argyris, C. (1969). The incompleteness of social psychological theory: examples from small group, cognitive consistency, and attribution research. *Am. Psychol.* 24, 893–908. doi: 10.1037/h0028880
- Barrick, M. R., Stewart, G. L., Neubert, M. J., and Mount, M. K. (1998). Relating member ability and personality to work-team processes and team effectiveness. *J. Appl. Psychol.* 83, 377–391. doi: 10.1037/0021-9010.83.3.377
- Barry, B., and Stewart, G. L. (1997). Composition, process, and performance in self-managed groups: the role of personality. *J. Appl. Psychol.* 82, 62–78. doi: 10.1037/0021-9010.82.1.62
- Becker, T. E. (2005). Potential problems in the statistical control of variables in organizational research: A qualitative analysis with recommendations. *Organ. Res. Methods* 8, 274–289. doi: 10.1177/1094428105278021
- Bell, S. T. (2007). Deep-level composition variables as predictors of team performance: a metaanalysis. *J. Appl. Psychol.* 92, 595–615. doi: 10.1037/0021-9010.92.3.595
- Bergmann, B., and Schaeppi, J. (2016). A data-driven approach to group creativity. *Harv. Bus. Rev.* 12, 43–62.
- 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*. eds. K. J. Klein and S. W. Kozlowski (San Francisco: Jossey-Bass), 349–381.
- Bliese, P. D. (2016). Multilevel modeling in R. Available at: http://cran.rproject.org/doc/contrib/Bliese_Multilevel.pdf
- Bliese, P. D., and Ployhart, R. E. (2002). Growth modeling using random coefficient models: model building, testing, and illustrations. *Organ. Res. Methods* 5, 362–387. doi: 10.1177/109442802237116
- Bonner, B. L., Sillito, S. D., and Baumann, M. R. (2007). Collective estimation: accuracy, expertise, and extroversion as sources of intra-group influence. *Organ. Behav. Hum. Decis. Process.* 103, 121–133. doi: 10.1016/j.obhdp.2006.05.001
- Boswell, W. R., Boudreau, J. W., and Tichy, J. (2005). The relationship between employee job change and job satisfaction: the honeymoon-hangover effect. *J. Appl. Psychol.* 90, 882–892. doi: 10.1037/0021-9010.90.5.882
- Bradley, B. H., Baur, J. E., Banford, C. G., and Postlethwaite, B. E. (2013). Team players and collective performance: how agreeableness affects team performance over time. *Small Group Res.* 44, 680–711. doi: 10.1177/1046496413507609
- Bunderson, J. S., and Boumgarden, P. (2010). Structure and learning in self-managed teams: why “bureaucratic” teams can be better learners. *Organ. Sci.* 21, 609–624. doi: 10.1287/orsc.1090.0483
- Bush, J. T., LePine, J. A., and Newton, D. W. (2018). Teams in transition: An integrative review and synthesis of research on team task transitions and propositions for future research. *Hum. Resour. Manag. Rev.* 28, 423–433. doi: 10.1016/j.hrmr.2017.06.005
- Carmeli, A., and Gittell, J. H. (2009). High-quality relationships, psychological safety, and learning from failures in work organizations. *J. Organ. Behav.* 30, 709–729. doi: 10.1002/job.565
- Chang, A., Bordia, P., and Duck, J. (2003). Punctuated equilibrium and linear progression: toward a new understanding of group development. *Acad. Manag. J.* 46, 106–117. doi: 10.5465/30040680
- Chen, G., Tjosvold, D., Zhao, H., Ning, N., and Fu, Y. (2011). Constructive controversy for learning and team effectiveness in China. *Asia Pac. J. Hum. Resour.* 49, 88–104. doi: 10.1177/1038411110391708
- Costa, P. T. J., and McCrae, R. R. (1994). “Set like plaster? Evidence for the stability of adult personality,” in *Can Personality change?* eds. T. F. Heatherton and J. L. Weinberger (Washington, DC: American Psychological Association), 21–40.
- Delice, F., Rousseau, M., and Feitosa, J. (2019). Advancing teams research: what, when, and how to measure team dynamics over time. *Front. Psychol.* 10:1324. doi: 10.3389/fpsyg.2019.01324
- Deng, H., Leung, K., Lam, C. K., and Huang, X. (2019). Slacking off in comfort: a dual-pathway model for psychological safety climate. *J. Manag.* 45, 1114–1144. doi: 10.1177/0149206317693083
- Doblinger, M. (2022). Individual competencies for self-managing team performance: a systematic literature review. *Small Group Res.* 53, 128–180. doi: 10.1177/10464964211041114
- Edmondson, A. C. (1999). Psychological safety and learning behavior in work teams. *Adm. Sci. Q.* 44, 350–383. doi: 10.2307/2666999
- Edmondson, A. (2002). “Managing the risk of learning,” in *International Handbook of Organizational Teamwork*. ed. M. West (Blackwell, London).
- Edmondson, A. C., and Lei, Z. (2014). Psychological safety: The history, renaissance, and future of an interpersonal construct. *Annu. Rev. Organ. Psych. Organ. Behav.* 1, 23–43. doi: 10.1146/annurev-orgpsych-031413-091305
- Edmondson, A. C., and Mogelof, J. P. (2005). “Explaining psychological safety in innovation teams,” in *Creativity and Innovation in Organizational Teams. 1st Edn.* eds. L. L. Thompson and H.-S. Choi (Psychology Press), 129–156.
- Festinger, L. (1950). Informal social communication. *Psychol. Rev.* 57, 271–282. doi: 10.1037/h0056932
- Ford, C., and Sullivan, D. M. (2004). A time for everything: how the timing of novel contributions influences project team outcomes. *J. Organ. Behav.* 25, 279–292. doi: 10.1002/job.241
- Frazier, M. L., Fainshmidt, S., Klinger, R. L., Pezeshkan, A., and Vracheva, V. (2017). Psychological safety: a meta-analytic review and extension. *Pers. Psychol.* 70, 113–165. doi: 10.1111/peps.12183
- Galecki, A., and Burzykowski, T. (2013). *Linear Mixed-Effects Models Using R: A Step-By-Step Approach*. New York: Springer.
- Garvin, D. A., Edmondson, A. C., and Gino, F. (2008). Is yours a learning organization? *Harv. Bus. Rev.* 86, 109–134.
- Gebauer, J. E., Sedikides, C., Wagner, J., Bleidorn, W., Rentfrow, P. J., Potter, J., et al. (2015). Cultural norm fulfillment, interpersonal belonging, or getting-ahead? A large-scale cross-cultural test of three perspectives on the function of self-esteem. *J. Pers. Soc. Psychol.* 109, 526–548. doi: 10.1037/pspp0000052
- George, J. M. (1990). Personality, affect, and behavior in groups. *J. Appl. Psychol.* 75, 107–116. doi: 10.1037/0021-9010.75.2.107
- Gerlach, R., and Gockel, C. (2018). We belong together: belonging to the principal's in-group protects teachers from the negative effects of task conflict on psychological safety. *School Leadership Manag.* 38, 302–322. doi: 10.1080/13632434.2017.1407307
- Gersick, C. J. G. (1988). Time and transition in work teams: toward a new model of group development. *Acad. Manag. J.* 31, 9–41. doi: 10.5465/256496
- Gockel, C., and Brauner, E. (2013). The benefits of stepping into others' shoes: perspective taking strengthens transactive memory. *Basic Appl. Soc. Psychol.* 35, 222–230. doi: 10.1080/01973533.2013.764303
- Goldberg, L. R. (1993). The structure of phenotypic personality traits. *Am. Psychol.* 48, 26–34. doi: 10.1037/0003-066X.48.1.26
- Grawitch, M. J., and Munz, D. C. (2004). Are your data nonindependent? A practical guide to evaluating nonindependence and within-group agreement. *Underst. Stat.* 3, 231–257. doi: 10.1207/s15328031us0304_2
- Graziano, W. G., Habashi, M. M., Sheese, B. E., and Tobin, R. M. (2007). Agreeableness, empathy, and helping: a person x situation perspective. *J. Pers. Soc. Psychol.* 93, 583–599. doi: 10.1037/0022-3514.93.4.583
- Hackman, J. R., Brousseau, K. R., and Weiss, J. A. (1976). The interaction of task design and group performance strategies in determining group effectiveness. *Organ. Behav. Hum. Perform.* 16, 350–365. doi: 10.1016/0030-5073(76)90021-0

- Hackman, J. R., and Wageman, R. (2005). A theory of team coaching. *Acad. Manag. Rev.* 30, 269–287. doi: 10.5465/amr.2005.16387885
- Harrison, D., Price, K., Gavin, J., and Florey, A. (2002). Time, teams, and task performance: changing effects of surface-and deep-level diversity on group functioning. *Acad. Manag. J.* 45, 1029–1045. doi: 10.5465/3069328
- Harvey, J. F., Leblanc, P. M., and Cronin, M. A. (2019). Beyond separate emergence: a systems view of team learning climate. *Front. Psychol.* 10:1441. doi: 10.3389/fpsyg.2019.01441
- Hayes, A. F., and Coutts, J. J. (2020). Use omega rather than Cronbach's alpha for estimating reliability. But.... *Commun. Methods Meas.* 14, 1–24. doi: 10.1080/19312458.2020.1718629
- Higgins, M., Dobrow, S. R., Weiner, J. M., and Liu, H. (2020). When is psychological safety helpful? A longitudinal study. *Acad. Manag. Discov.* 8, 77–102. doi: 10.5465/amd.2018.0242
- Hurtz, G. M., and Donovan, J. J. (2000). Personality and job performance: the big five revisited. *J. Appl. Psychol.* 85, 869–879. doi: 10.1037/0021-9010.85.6.869
- Innami, I. (1994). The quality of group decisions, group verbal behavior, and intervention. *Organ. Behav. Hum. Decis. Process.* 60, 409–430. doi: 10.1006/obhd.1994.1092
- Jehn, K. A. (1995). A multimethod examination of the benefits and detriments of intragroup conflict. *Adm. Sci. Q.* 40, 256–282. doi: 10.2307/2393638
- John, O. P., Donahue, E. M., and Kentle, R. L. (1991). *The Big Five Inventory—Versions 4a and 5a*. Berkeley, CA: University of California, Berkeley, Institute of Personality and Social Research
- John, O. P., and Robins, R. W. (1993). Determinants of interjudge agreement on personality traits: The big five domains, observability, evaluativeness, and the unique perspective of the self. *J. Pers.* 61, 521–551. doi: 10.1111/j.1467-6494.1993.tb00781.x
- Judge, T. A., Bono, J. E., Ilies, R., and Gerhardt, M. W. (2002). Personality and leadership: a qualitative and quantitative review. *J. Appl. Psychol.* 87, 765–780. doi: 10.1037/0021-9010.87.4.765
- Jung, T., and Wickrama, K. A. S. (2008). An introduction to latent class growth modeling analysis and growth mixture modeling. *Soc. Personal. Psychol. Compass* 2, 302–317. doi: 10.1111/j.1751-9004.2007.00054.x
- Knight, A. P. (2015). Mood at the midpoint: affect and change in exploratory search over time in teams that face a deadline. *Organ. Sci.* 26, 99–118. doi: 10.1287/orsc.2013.0866
- Kozlowski, S. W. J., and Bell, B. S. (2003). “Work groups and teams in organizations,” in *Handbook of Psychology: Industrial and Organizational Psychology*. Vol. 12. eds. W. C. Borman, D. R. Ilgen and R. J. Klimoski (London: Wiley), 333–375.
- Krumm, S., Mertin, I., and Dries, C. (2012). *Kompetenzmodelle*. Göttingen: Hogrefe.
- Lau, D. C., and Murnighan, J. K. (1998). Demographic diversity and faultlines: the compositional dynamics of organizational groups. *Acad. Manag. Rev.* 23, 325–340. doi: 10.2307/259377
- Lau, D. C., and Murnighan, J. K. (2005). Interactions within groups and subgroups: the effects of demographic faultlines. *Acad. Manag. J.* 48, 645–659. doi: 10.5465/amj.2005.17843943
- LePine, J. A., Buckman, B. R., Crawford, E. R., and Methot, J. R. (2011). A review of research on personality in teams: accounting for pathways spanning levels of theory and analysis. *Hum. Resour. Manag. Rev.* 21, 311–330. doi: 10.1016/j.hrmr.2010.10.004
- Levin, I. M., and Stokes, J. P. (1989). Dispositional approach to job satisfaction: role of negative affectivity. *J. Appl. Psychol.* 74, 752–758. doi: 10.1037/0021-9010.74.5.752
- Lewicki, R. J., Tomlinson, E. C., and Gillespie, N. (2006). Models of interpersonal trust development: theoretical approaches, empirical evidence, and future directions. *J. Manag.* 32, 991–1022. doi: 10.1177/0149206306294405
- Li, A. N., and Tangirala, S. (2021). How employees' voice helps teams remain resilient in the face of exogenous change. *J. Appl. Psychol.* 107, 668–692. doi: 10.1037/apl0000874
- Liang, J., Farh, C. L., and Farh, J. L. (2012). Psychological antecedents of promotive and prohibitive voice: a two-wave examination. *Acad. Manag. J.* 55, 71–92. doi: 10.5465/amj.2010.0176
- Marks, M. A., Mathieu, J. E., and Zaccaro, S. J. (2001). A temporally based framework and taxonomy of team processes. *Acad. Manag. Rev.* 26, 356–376. doi: 10.5465/amr.2001.4845785
- Marsh, H. W., Ellis, L. A., Parada, R. H., Richards, G., and Heubeck, B. G. (2005a). A short version of the self description questionnaire II: operationalizing criteria for short-form evaluation with new applications of confirmatory factor analyses. *Psychol. Assess.* 17, 81–102. doi: 10.1037/1040-3590.17.1.81
- Marsh, H. W., Trautwein, U., Lüdtke, O., Köller, O., and Baumert, J. (2005b). Academic self-concept, interest, grades, and standardized test scores: reciprocal effects models of causal ordering. *Child Dev.* 76, 397–416. doi: 10.1111/j.1467-8624.2005.00853.x
- Matthews, G., Deary, I. J., and Whiteman, M. C. (2009). *Personality Traits*. 3rd Edn. Cambridge: Cambridge University Press.
- Mayo, A. T., and Woolley, A. W. (2021). Variance in group ability to transform resources into performance, and the role of coordinated attention. *Acad. Manag. Discov.* 7, 225–246. doi: 10.5465/amd.2019.0231
- McAdams, D. P., and Pals, J. L. (2006). A new big five: fundamental principles for an integrative science of personality. *Am. Psychol.* 61, 204–217. doi: 10.1037/0003-066X.61.3.204
- McNeish, D. (2018). Thanks coefficient alpha, we'll take it from here. *Psychol. Methods* 23, 412–433. doi: 10.1037/met0000144
- Meyer, B., and Glenz, A. (2013). Team faultline measures a computational comparison and a new approach to multiple subgroups. *Organ. Res. Methods* 16, 393–424. doi: 10.1177/1094428113484970
- Meyer, B., and Schermuly, C. C. (2012). When beliefs are not enough: examining the interaction of diversity faultlines, task motivation, and diversity beliefs on team performance. *Eur. J. Work Organ. Psy.* 21, 456–487. doi: 10.1080/1359432X.2011.560383
- Meyer, B., Shemla, M., Li, J., and Wegge, J. (2015). On the same side of the faultline: inclusion in the leader's subgroup and employee performance. *J. Manag. Stud.* 52, 354–380. doi: 10.1111/joms.12118
- Morgeson, F. P., Mitchell, T. R., and Liu, D. (2015). Event system theory: An event-oriented approach to the organizational sciences. *Acad. Manag. Rev.* 40, 515–537. doi: 10.5465/amr.2012.0099
- Nakagawa, S., and Schielzeth, H. (2013). A general and simple method for obtaining R² from generalized linear mixed-effects model. *Methods Ecol. Evol.* 4, 133–142. doi: 10.1111/j.2041-210x.2012.00261.x
- Newman, A., Donohue, R., and Eva, N. (2017). Psychological safety: a systematic review of the literature. *Hum. Resour. Manag. Rev.* 27, 521–535. doi: 10.1016/j.hrmr.2017.01.001
- Okhuysen, G. A., and Waller, M. J. (2002). Focusing on midpoint transitions: an analysis of boundary conditions. *Acad. Manag. J.* 45, 1056–1065. doi: 10.2307/3069330
- Ostermeier, K., Davis, M., and Pavur, R. (2020). Personality configurations in teams: a comparison of compilation and composition models. *Team Perform. Manag.* 26, 227–246. doi: 10.1108/TPM-09-2019-0097
- Peeters, G., and Czapiński, J. (1990). Positive-negative asymmetry in evaluations: the distinction between affective and informational negativity effects. *Eur. Rev. Soc. Psychol.* 1, 33–60. doi: 10.1080/14792779108401856
- Pierro, A., Mannetti, L., De Grada, E., Livi, S., and Kruglanski, A. W. (2003). Autocracy bias in informal groups under need for closure. *Personal. Soc. Psychol. Bull.* 29, 405–417. doi: 10.1177/0146167203251191
- Preacher, K. J., Curran, P. J., and Bauer, D. J. (2006). Computational tools for probing interaction effects in multiple linear regression, multilevel modeling, and latent curve analysis. *J. Educ. Behav. Stat.* 31, 437–448. doi: 10.3102/10769986031004437
- R Development Core Team (2012). R: A Language and Environment for Statistical Computing (Version 2.15). Available at: <http://www.R-project.org> (Accessed August 2, 2012).
- Rammstedt, B., Koch, K., Borg, I., and Reitz, T. (2004). Entwicklung und validierung einer kurzskala für die messung der big-five-persönlichkeitsdimensionen in umfragen. *ZUMA Nachrichten* 28, 5–28.
- Roe, R. A., Gockel, C., and Meyer, B. (2012). Time and change in teams: where we are and where we are moving. *Eur. J. Work Organ. Psy.* 21, 629–656. doi: 10.1080/1359432X.2012.729821
- Schulte, M., Cohen, N. A., and Klein, K. J. (2012). The coevolution of network ties and perceptions of team psychological safety. *Organ. Sci.* 23, 564–581. doi: 10.1287/orsc.1100.0582
- Smith, K. G., Smith, K. A., Olian, J. D., Sims, H. P., O'Bannon, D. P., and Scully, J. A. (1994). Top management team demography and process: the role of social integration and communication. *Adm. Sci. Q.* 39, 412–438. doi: 10.2307/2393297
- Stokes, J. P., and Levin, I. M. (1990). The development and validation of a measure of negative affectivity. *J. Soc. Behav. Pers.* 5, 173–186.
- Taggar, S. (2002). Individual creativity and group ability to utilize individual creative resources: a multilevel model. *Acad. Manag. J.* 45, 315–330.
- Thatcher, S. M. B., and Patel, P. C. (2012). Group faultlines: a review, integration, and guide to future research. *J. Manag.* 38, 969–1009. doi: 10.1177/0149206311426187

Tsui, A., Egan, T., and O'Reilly, C. (1992). Being different: relational demography and organizational attachment. *Adm. Sci. Q.* 37, 549–579. doi: 10.2307/2393472

Tsui, A., and O'Reilly, C. (1989). Beyond simple demographic effects: The importance of relational demography in superior-subordinate dyads. *Acad. Manag. J.* 32, 402–423.

Ulloa, B. C. R., and Adams, S. G. (2004). Attitude toward teamwork and effective teaming. *Team Perform. Manag.* 10, 145–151. doi: 10.1108/13527590410569869

Van de Schoot, R., Lugtig, P., and Hox, J. (2012). A checklist for testing measurement invariance. *Eur. J. Dev. Psychol.* 9, 486–492. doi: 10.1080/17405629.2012.686740

Watson, W., Kumar, K., and Michaelsen, L. (1993). Cultural diversity's impact on interaction process and performance: comparing homogeneous and diverse task groups. *Acad. Manag. J.* 36, 590–602. doi: 10.5465/256593

Weiner, J., Francois, C., Stone-Johnson, C., and Childs, J. (2021). Keep safe, keep learning: principals' role in creating psychological safety and organizational learning during the COVID-19 pandemic. *Front. Educ.* 5:618483. doi: 10.3389/educ.2020.618483

Wiese, C. W., and Burke, C. S. (2019). Understanding team learning dynamics over time. *Front. Psychol.* 10:1417. doi: 10.3389/fpsyg.2019.01417