The Journal of Systems and Software 210 (2024) 111937

Contents lists available at ScienceDirect

The Journal of Systems & Software

journal homepage: www.elsevier.com/locate/jss

Exploring the relation between personality traits and agile team climate: Aggregating results from a twice replicated study in a telecom company

Sai Datta Vishnubhotla *, Emilia Mendes

Department of Computer Science, BTH - Blekinge Tekniska Högskola, Sweden

ARTICLE INFO

Keywords:
Personality traits
Team climate
Correlation analysis
Meta-analysis
Regression analysis
Agile software development

ABSTRACT

Context: Former literature revealed team performance is contingent on personality composition and interactive effects of team climate. While decades of research on personality prevails in software engineering, team climate remains sparsely researched.

Objective: In agile software development, individuals and interactions are key sources of agility. This study replicates a previous study and analyzes the relationship between five-factor-model personality traits and team climate dimensions among agile teams in a telecom company.

Method: A Web-based survey was replicated twice, first with 75 professionals from 12 teams in Sweden, followed by 46 professionals from seven teams in India. The data was used for correlation, regression analyses, and meta-analysis.

Results: We observed significant negative correlations between *neuroticism* and all the team climate dimensions. Meta-analysis identified a *significant medium-sized negative effect between neuroticism and participative safety*. Regression analysis showed *personality traits accounted for around 10 % of the variance in team climate dimensions*.

Conclusions: High neuroticism is not conducive to team climate as emotionally unstable members could impair team cohesion by being reactive and susceptible to stress. Managers assembling Scrum teams ought to mitigate higher neuroticism by counterbalancing it with an elevation of corresponding negatively correlated personality variables and providing support/training towards increasing the aforementioned variables.

1. Introduction

The software/software-intensive industry remains plagued with project failures, and many of the reasons for such outcomes are human-related (e.g., (Fernando Capretz, 2014; Capretz et al., 2017)). Given the complexity of software/software-intensive systems, companies are increasingly embracing teamwork in an effort to remain competitive in the global market (e.g., (O'Neill and Allen, 2011; Lindsjørn et al., 2016)). Furthermore, research shows that unsuitable team composition in software projects is one of the main drivers of project failures (e.g., (Gilal et al., 2018; Kollmann et al., 2009; Truong and Jitbaipoon, 2016, Software project,2023)); therefore, it is imperative to investigate how best to build high-performance software teams, without compromising team climate (Gilal et al., 2018, Coordination Challenges, 2023, Agile Development, 2023; Zolduoarrati et al., 2023).

Team climate represents team members' perceptions in relation to "the extent to which a team makes use of structures, policies, and practices supporting trust, cohesion, and innovativeness" (Berraies and Chouiref, 2022). Such shared perceptions can affect both team members' personal relationships and satisfaction, as well as their performance and the quality of their deliverables (Acuña et al., 2015; Fay et al., 2004; Acuña et al., 2008). Team climate has been shown to be an important facilitator of effective knowledge management in firms (Berraies and Chouiref, 2022). Research in the field of social psychology has provided evidence about team climate's association with personality, team performance, and team task characteristics (e.g., (Barry and Stewart, 1997; Molleman et al., 2004; St J. Burch and Anderson, 2004; Sumner and Molka-Danielsen, 2010)).

However, a recent tertiary study on human aspects in Software Engineering (SE) showed that only 13 studies in SE have investigated team climate (Zolduoarrati et al., 2023), which suggests that further investigation in SE is clearly needed. Such need becomes even more pressing given the expeditious rate of adoption of Agile Software Development (ASD) methods over the past decades (e.g., (Digital.ai 2021; Rodríguez et al., 2012; Ambler, 2008)), which increased the need for collaborative work (Lindsjørn et al., 2016). One of the core values in the Agile

* Corresponding author.
E-mail address: sai-datta.vishnubhotla@bth.se (S.D. Vishnubhotla).

https://doi.org/10.1016/j.jss.2023.111937
Received 8 May 2023; Received in revised form 11 December 2023; Accepted 17 December 2023
Available online 19 December 2023

Manifesto emphasizes on "individuals and interactions over processes and tools" (Fowler and Highsmith, 2001). So, a person's fit to a team and the ability to perform substantially affect a project's outcome (Trendowicz and Münch, 2009). In an agile team, members are expected to demonstrate capability in relation to methodological, social, and creative aspects while interacting with teammates. The capability of an individual is influenced by a compendium of components such as their personality, abilities, attitude, and practices (Vishnubhotla et al., 2018; Matturro et al., 2015; Feldt et al., 2010). Within the context of this research, we are investigating team climate within ASD, and team climate relates to one of the most investigated capabilities of software professionals – personality (Costa et al., 1992).

Personality is characterized as a set of individual differences and traits that significantly affect one's behavior (Matthews et al., 2003) and where consistent patterns of behavior are observed for each personality trait (Matthews et al., 2003). Therefore, several psychometric tests have been devised, where personality questionnaires are used in order to capture an individual's context-free behavior (Yilmaz et al., 2017). Studies have shown significant relationships between personality and other aspects such as teamwork (e.g. (Weinberg, 1971; Lee and Shneiderman, 1978; McCrae and Costa, 1989; Rothstein and Goffin, 2006)), team climate (e.g., (Vishnubhotla et al., 2020)), decision-making (e.g., (Mendes et al., 2019)), and team performance (e.g., (Soomro et al., 2016; Gila et al., 2014; Anderson et al., 2018; Soomro et al., 2015)), to name a few.

In our previous work (Vishnubhotla et al., 2020), we were the first to investigate specifically the relationship between personality and team climate, within an ASD context. In that study, we surveyed eight ASD teams and, by means of correlation and regression analyses, uncovered the association between Five Factor Model (FFM) personality traits (Costa et al., 1992; Cruz et al., 2015) and the perception of team climate, as represented using the Team Climate Inventory (TCI) (Anderson and West, 1998). Although statistically significant positive correlations were identified between personality traits and team climate dimensions, the regression analyses showed that the independent variable (personality trait) in each model presented small explanatory power on the dependent variable (overall team climate). That research was later expanded to other divisions of the collaborating company in order to better understand the phenomena under investigation. Therefore, given the importance of both personality and team climate to ASD and SE, this paper presents the results from two replications, thus investigating further the relationship between personality and team climate, and also within an ASD context.

Concerns about the reliability of empirical research results are quickly becoming endemic, and SE is no exception (Jørgensen et al., 2016). A Systematic Literature Review (SLR) on personality in SE by Cruz et al. (Cruz et al., 2011) indicated that among the set of empirical studies they analyzed, there has not been a wide replication. Replication is crucial in the case of empirical studies for consolidating knowledge that is potentially advantageous for new research and to pursue generalizable results about the state of practice (da Silva et al., 2014).

The idea of replication is to check whether the previously observed results hold. If they do, this implies that the results are more reliable, even without anything new being discovered (Acuña et al., 2015). Further, replication facilitates the realization of a larger sample by means of aggregation, which could increase the statistical power of models (compared to the original study) and lead to statistically significant results (Acuña et al., 2015). In summary, the underlying proposition is that confirmation through validation implies greater reliability and confidence in the results.

To date, there is limited research available that attempts to quantify empirically how various factors affect agile team climate (Vishnubhotla et al., 2020; Soomro et al., 2016). So, in this study, we emphasize on analyzing and reporting the findings from two replications performed in relation to our previous study (Vishnubhotla et al., 2020). The novelty of this study is two-fold. First, we validate the findings from the original study (Vishnubhotla et al., 2020) by using new and independent data sets acquired via replications. Next, we aggregate the results from the two replications with the ones from the original study by means of a meta-analysis of correlations.

It is regarded as a good practice in SE to replicate at the same site to find out whether the results are consistent (Acuña et al., 2015). So, with the support of our industrial collaborator, a large telecom company, we performed two instances of replication. While 75 members from 12 agile teams participated in the first replication of the survey, another 46 members from 7 agile teams were surveyed in the subsequent replication. Data gathered from the replications was used towards correlation and regression analyses, as performed in the original study (Vishnubhotla et al., 2020). Additionally, in this study, the results observed across samples were aggregated using meta-analysis (A. Santos et al., 2021), and the accuracy of the regression models developed was evaluated by a cross-validation procedure.

Insights from the replication would help diagnose other project planning problems and aid in revisiting the team selection strategies at our industrial collaborator's site. Furthermore, exploring the relationship between personality traits and team climate dimensions by considering the voices of team members would help towards evaluating and improving the team climate research in ASD.

The remainder of this paper is structured as follows: Section 2 presents the related work on personalities and team climate in SE. Section 3 presents the methodological details of the study, and Section 4 summarizes the results of correlation analysis, meta-analysis, and regression analysis from two replications. In Section 5, we present the threats to the validity of this study, and in Section 6, we discuss the key findings from our results. Finally, Section 7 states the conclusions of this research.

2. Related work

The related work first discusses the studies published in relation to investigating FFM personality traits. Further, the studies that explored factors influencing agile team climate are also discussed.

2.1. Studies investigating the personality of team members

Investigating personality characteristics of software professionals has been a popular research topic over the past decades (Cruz et al., 2015). Although there exist multiple tests that can be employed to assess the personality of individuals (e.g., Myers-Briggs Type Indicator,[1] HEXACO model of personality structure (Lee and Ashton, 2020)), previous studies have described the Five Factor Model (FFM) tests as prominent, valid, and reliable (Barrick and Mount, 1991; Barrick et al., 1998). Furthermore, the work by McCrae et al. (McCrae and Terracciano, 2005) details the results of cross-cultural tests in more than 50 different societies, which provide support for the 'universality' of the FFM. The FFM emphasizes on a structure that categorizes dimensions of differences in human personalities. The model categorizes five broad personality traits: openness to experience, conscientiousness, extraversion, agreeableness, and neuroticism (Johnson, 2014).

Among various instruments developed in relation to FFM, the NEO Five-Factor Inventory (NEO-FFI) (Matzler et al., 2008) and the Revised NEO Personality Inventory (NEO-PI-R) (Costa and McCrae, 2008) are two instruments that initially gained popularity. However, both of them are proprietary instruments (Salleh et al., 2014) that require training and a license for use (Kosti et al., 2014). On the other hand, the International Personality Item Pool (IPIP) is a freely available FFM-based set of items and scales (Goldberg et al., 2006). Some of the benefits of IPIP scales are that they do not require licensed accessors and facilitate statistical analysis by providing numerical scores for each factor.

One of the initial instruments created from the IPIP was Goldberg's

[1] https://www.myersbriggs.org/my-mbti-personality-type/mbti-basics/

300-item inventory (Goldberg, 1999), which was designed to measure constructs similar to those assessed by NEO PI-R (Costa and McCrae, 2008). Subsequently, to facilitate easy circulation of Goldberg's inventory over the Web, a version of that inventory named IPIP-NEO was created (Johnson, 2014). This automated instrument is considered much more efficient than paper-based instruments (Goldberg et al., 2006). However, as the IPIP-NEO inventory's length (300 items) was more than the original 240-item NEO PI-R inventory, it was not very convenient to integrate it with other psychological measures (Johnson, 2014). Due to this, later on, the inventory was simplified to create multiple smaller variants that were not only practically useful but also psychometrically acceptable (Donnellan et al., 2006).

Although shorter versions of the IPIP-NEO inventory with 20, 50, and 100 items exist, none of them were robust in terms of measuring the six facets associated with each of the FFM domains (Johnson, 2014). Thus, another version of IPIP-NEO that can reliably and validly represent five broad domains using 120 items was developed. In our original study (Vishnubhotla et al., 2020), we used survey methodology and employed the 120-item IPIP-NEO questionnaire as a data collection instrument to gather personality characteristics.

In relation to the related work within the context of this study, we limit our scope to discussing the studies that were published after our original study (Vishnubhotla et al., 2020) and those that exclusively focused on using the IPIP for assessing the personality characteristics in SE. Readers are referred to our previous work for a wider discussion on other instruments used in combination with the FFM (Vishnubhotla et al., 2020).

Sturdee et al. (Sturdee et al., 2022) conducted an exploratory study for identifying the personality traits of game developers. Survey methodology was employed to gather responses for the 50-item IPIP questionnaire from 123 game developers. The authors used one-way ANOVA to compare the findings with data from software developers. They observed that game development professionals tend to have higher neuroticism than other software development professionals and personnel in art, design, and production. Other characteristics like agreeableness, conscientiousness, openness, and extraversion were observed to be comparatively low for game developers.

Caulo et al. (Caulo et al., 2021) conducted an empirical study to investigate how personality traits affect the productivity of software developers within the context of distributed development of multiplatform apps in a GitHub project. By recruiting 31 master's students (grouped into 13 teams) with computer science background as subjects, the authors gathered responses for the 120-item IPIP-NEO questionnaire and used the metrics related to source code and commits as a proxy to measure productivity. A correlation analysis identified that the most productive participants were those with the highest scores for the personality traits of agreeableness and conscientiousness.

Mendes et al. (Mendes et al., 2021) investigated the relationship between decision-making style and personality within the context of software project development. A survey mechanism was used to gather data from 63 Brazilian software professionals. While the personality information was gathered using the 120-item IPIP-NEO questionnaire, the decision-making style of professionals was measured using a 30-problem questionnaire. Correlation analysis between decision-making style and personality facets identified seven statistically significant correlations. Further, a regression model developed using a backward elimination process to predict decision-making style selected only the agreeableness variable as a significant predictor to explain 4.2 % of the variation in decision-making style. The model accuracy was evaluated and deemed good enough.

Qamar and Malik (Qamar and Malik, 2020) conducted a replication study to evaluate the impact of software quality and team productivity on the team homogeneity index. This index was computed based on the personality data gathered in the form of responses from the 50-item IPIP questionnaire. The authors recruited 35 software professionals and 215 students with computer science background for the study. Besides computing the team homogeneity index, the authors also computed the weighted team homogeneity index to determine whether weights assigned to personality traits make any difference. A comparative analysis of the two indices indicated that the weighted index was more strongly correlated with team productivity and software quality in the case of teams comprising professionals.

Calefato and Lanubile (Calefato and Lanubile, 2022) emphasized on assessing the performance of general-purpose personality detection tools when applied to developers' e-mails retrieved from the public archives of the Apache software foundation. By using an electronic version of the 20-item IPIP questionnaire as a personality instrument, the authors gathered 50 valid responses by sending study invites to around 1000 email IDs from the Apache archives. This self-assessed personality data was regarded as ground truth and compared with the personality scores extracted by applying four Big Five based tools to developers' email dataset. Results from this study showed a decrease in performance when general-purpose tools are used out of the domain, as neither they agree with each other nor the self-reported personality scores. The results further suggested the need for personality detection tools specific for the SE domain.

Penzenstadler et al. (Penzenstadler et al., 2022) investigated the effects of a neuroplasticity practice on the attention awareness, well-being, perceived productivity, and self-efficacy of computer workers. In relation to this study, the authors initially investigated whether personality traits would show a difference in how the participants' awareness shifted. This initial investigation was done using the mini IPIP personality test. However, the IPIP test did not reveal anything in the analysis and did not lead to any conclusive evidence.

Among the recently published studies, we observed that multiple studies used IPIP for investigating personality characteristics among professionals (Sturdee et al., 2022; Mendes et al., 2021; Calefato and Lanubile, 2022). While two studies recruited students (Caulo et al., 2021; Francese et al., 2021), another two studies recruited both professionals and students (Qamar and Malik, 2020; Penzenstadler et al., 2022). Although some studies seem to be using the shorter versions of the IPIP questionnaire in an attempt to increase the low response rate of surveys in SE (Calefato and Lanubile, 2022), none of the shorter IPIP inventories that contained 20, 50, and 100 items were considered robust in terms of covering a wide range of facets. In essence, the shorter IPIP versions cannot measure the six facets associated with each of the five FFM domains (Johnson, 2014). We observed that only three of the recent studies (Caulo et al., 2021; Penzenstadler et al., 2022; Francese et al., 2021) used the 120-item IPIP-NEO questionnaire.

2.2. Studies investigating team climate with a focus on personalities

Team climate relates to the shared perception across members of a team regarding the organizational policies and practices (Anderson and West, 1998). Theoretical support and, later on, evidence-based support for its relationship with personality traits has been the focus of research of numerous studies in fields such as psychology, and for decades, detailed accounts of such studies are given in, for example, (Xu et al., 2019) and (Chatzi et al., 2022). As the focus of this study relates to personality traits and team climate within an ASD context, we will focus herein on the presentation of studies within such context.

The team climate for SE teams might substantially differ from that of the workforce belonging to other domains due to the differences in the nature of tasks and activities in software development teams (Soomro et al., 2016). The shared perception about policies and practices within a team not only influences the personal relationships among members but also affects the satisfaction of team members, the quality of software developed, and, in turn, the performance at the organizational level (Acuña et al., 2015; Fay et al., 2004; Acuña et al., 2008).

The TCI instrument was designed to inspect the four-factor theory of team climate (Anderson and West, 1998). This instrument was observed to demonstrate consistent psychometric properties on repeated validations (Mathisen et al., 2004; Ragazzoni et al., 2002). The four-factor theory associated with this instrument emphasizes on dimensions that are deemed essential for effective team functioning and propensity to innovation (Anderson and West, 1998) - team vision, participative safety, task orientation, and support for innovation.

Within SE, Soomro et al. (Soomro et al., 2015) initially investigated the relationship between personality traits, team climate, and team performance. The 120-item IPIP-NEO inventory was used to measure personality traits and the 38-item TCI instrument was employed for measuring team climate. Responses from 36 professionals were gathered by administering questionnaires to IT employees. Upon performing a correlation and regression analyses, authors observed a significant positive relationship between the extraversion personality trait and team climate. However, a closer look into their data collection mechanism informs us that the authors gathered personality and team climate information from completely disengaged people, i.e., from respondents belonging to different teams and different organizations. In such a context, collective analysis of personality scores and team climate scores of disengaged people do not make concrete coherence either at the team level or at the organizational level and, therefore, poses a threat to the validity of their results.

In a subsequent study by Soomro et al. (Soomro et al., 2016), an SLR was conducted to aggregate evidence on previously conducted studies investigating the relationship between personality and team climate & performance. This study reported about very limited research in relation to identifying what personality compositions lead to a better team climate. In order to address this research gap, in our previous study (Vishnubhotla et al., 2020), we investigated the association between FFM personality traits and the perception of climate within agile teams. The survey was conducted at a large telecom company. A total of 43 members from eight agile teams participated in this survey, and the acquired data was used towards a correlation analysis to investigate the relationships between personality traits and team climate dimensions, followed by multivariate regression analysis for model fitting based on the gathered data.

Analysis of our survey data (Vishnubhotla et al., 2020) identified a statistically significant positive correlation between openness to experience personality characteristic and support for innovation team climate dimension ($r = 0.31$). A similar significant positive correlation was observed between agreeableness characteristic and overall team climate ($r = 0.35$). In relation to the significant correlations, two regression models were developed. We observed that the openness to experience personality trait could explain only 9.7 % of the variance in the support for innovation dimension, and the agreeableness personality trait could explain only 12.4 % of the variance in the perceived team climate scores. Further investigation was clearly needed and is the subject of the research detailed herein.

By means of performing backward and forward snowballing over the SLR conducted by Soomro et al. (Soomro et al., 2016), in our previous survey study (Vishnubhotla et al., 2020), we identified a set of related SE studies that utilized the TCI instrument towards studying team climate at various contexts. In order to further identify the additional studies that investigated team climate in SE contexts from recent years, i.e., to find publications relating to team climate in SE after our survey study was published (the year 2020), we employed the forward-snowballing technique and reviewed the studies that cited Soomro et al.'s SLR (Soomro et al., 2016) in recent years, and our survey study (Vishnubhotla et al., 2020). This snowballing technique was iterated until no more relevant studies were found. This led to the identification of four studies, as elaborated next.

Dutra et al. (Dutra et al., 2020) initially developed an instrument to understand how different factors influence the organizational climate of ASD teams. In a subsequent study (Dutra and Santos, 2020), further investigations were done with respect to how organizations considered assessment of the organizational climate of agile teams and what were the benefits and the difficulties associated with such assessments. A qualitative study was conducted with five Brazilian organizations, and key personnel involved with organizational assessments were interviewed. The qualitative study identified 16 benefits and nine difficulties of organizational climate assessments. Their instrument for organizational climate was designed specifically to measure dimensions like communication, collaboration, and leadership that are not emphasized in the TCI; however, such results are understandable given that their goal was not, unlike the TCI, to understand the climate at the team level.

Lee and Chen (Lee and Chen, 2020) reported a conceptual study that uses propositional methodology with a review of existing literature pertaining to Software Process Tailoring (SPT), absorptive capacity, transactive memory system, and TCI to develop a theoretical model to foster SPT performance. The authors report that the team climate dimensions act as positive moderators in promoting a team's dynamic learning process. However, their study was solely conceptually established, and there were no empirical findings. Team climate was only considered as a contextual factor used to understand how it moderates the process to yield effective SPT performance.

Finally, Francese et al. (Francese et al., 2021) explored the relationship between personality traits and team climate within a distributed smart-working development context driven by the COVID-19 pandemic. This study recruited 53 graduate students with a background in computer science as subjects. The students were grouped into 19 teams, where they individually answered the 120-item IPIP-NEO questionnaire and TCI questionnaire. Their responses were used to carry out correlation and regression analyses. Correlation analysis revealed extroversion personality trait to be related to team climate. Task orientation was the only team climate dimension that was observed to satisfy the normality assumption. So, the authors adopted a model-fit approach and built a linear regression model to predict task orientation using extraversion as the independent variable. The extraversion variable could explain 13.96 % of the variance in task orientation scores.

Although the objectives and the analyses performed in Francese et al.'s study (Francese et al., 2021) look similar to those used in our previous study (Vishnubhotla et al., 2020), they recruited students as subjects and did not emphasize on industrial agile contexts like our study. Moreover, another crucial aspect that differentiates both studies is the domain emphasized. While the motivation behind the study by Francese et al. (Francese et al., 2021) was to better understand the relationship between personality and team climate factors in a distributed smart-working development environment, our previous study (Vishnubhotla et al., 2020) was driven by the need of the partnering company, and it focused on investigation within the context of a telecom company. Furthermore, they did not report any information about the effect sizes observed in their correlation analysis. While it would be ideal to compare replications and original studies in terms of effect sizes, incomplete reporting makes it difficult to understand the extent to which a replication is confirmatory and to what extent it yields additional knowledge to the SE community (Shepperd et al., 2018).

In the wake of multifold aspects, such as our industrial collaborator's need for investigating factors that can contribute to a better climate within their agile teams (Vishnubhotla et al., 2020), the lack of research in SE in relation to identifying factors affecting team climate, and the concern for contributing to the reliability and generalizability of our previous findings, this study replicates the investigation on the relationship between personalities and team climate (Vishnubhotla et al., 2020). To the best of our knowledge, this is the first study replicating the investigation in industrial agile contexts. The results obtained from the two replications and the original study are further aggregated by means of a meta-analysis of correlations. Knowledge from such an analysis can guide managers on how to compose better teams on the basis of team member personalities and their effect upon team climate.

3. Research method

The main goal of this study is to investigate, via replication, the relationship between personality traits and team climate factors of professionals working in agile teams of a telecom company. Like the original study (Vishnubhotla et al., 2020), this study also considers the personality traits from FFM as independent variables and the TCI's team climate dimensions & the aggregate variable Individual's Perceived Team Climate (IPTC) as dependent variables. Moreover, to address the research gap in relation to a very limited number of former SE studies reporting inter-correlations among personality traits (Vishnubhotla et al., 2020), this study also emphasizes on exploring the relationship among personality traits (similarly, among team climate factors also). Thus, the research questions guiding this study are as follows:

RQ.1) What is the relationship among personality traits of agile team members working in a telecom company?
RQ.2) What is the relationship among team climate factors of agile team members working in a telecom company?
RQ.3) What is the relationship between personality traits and team climate factors of agile team members working in a telecom company?

The first research question focuses on comparing and exploring the relationship among FFM personality traits (see Section 4.1.1 and Section 4.2.1), and the second research question focuses on exploring how the four team climate factors relate to one another (see Section 4.1.2 and Section 4.2.2). Finally, the third research question targets at exploring the relationship between personality traits and team climate factors (see Section 4.2.3, Section 4.3, and Section 4.4). The relationship among variables in the first and second research questions was investigated via descriptive analyses and inter-correlation analyses. Whereas the relationship between variables in the third research question was investigated by correlation analysis, meta-analysis, and regression analysis. While this study devotes more attention towards exploring relationships in the third research question, findings from the first two research questions would also provide inputs for research and practice.

Below, we describe each activity in our replication study using the same phases as used in the original study, so to illustrate similarities and/or differences between the studies. An overview of samples and types of analyses from the original and the current studies is presented in Fig. 1.

The two instances of replication from the current study were conducted within the context of a telecom company that provides global telecom and multimedia services (company A, from hereon). While the first replication was performed in connection with their Swedish division (SE_2) in September 2020, the second replication was conducted in association with their Indian division (IN_1) in June 2021. It is important to note that, unlike the original study, the two instances of replications detailed herein were conducted during the course of the COVID-19 pandemic. It can be observed from Fig. 1 that each sample comprised professionals associated with a different project (projects P1, P2, and P3 in Fig. 1). Due to the sensitive nature of the projects, we do not have permission from company A to disclose any information about them.

In relation to the types of analyses conducted, as shown in Fig. 1, except for the meta-analysis, the rest of the analyses between the original study and the current study are common. Except for the meta-analysis, the rest of the analyses reported in this study employ the two samples acquired via replication (SE_2 and IN_1). The current study additionally performs a meta-analysis where the correlation analysis results from the original study's SE_1 sample were aggregated with the correlation analysis results from the current study's SE_2 and IN_1 samples.

Both Swedish and Indian divisions hosted multiple agile teams that adhered to Scrum practices for developing software-intensive charging and billing systems for mobile networks. Two senior professionals from each division, undertaking the role of a product owner, helped us in the smooth execution of the study. These members were closely associated with the agile teams and were playing a strategic management role where one of the key responsibilities was recruiting members to teams. They were therefore chosen as the source of contact for our study.

In order to understand the personality characteristics of individuals and gather their perceptions with respect to team climate, we resorted to using survey methodology for gathering quantitative data. Like in the original study (Vishnubhotla et al., 2020), the 120-item IPIP-NEO personality test questionnaire was used to understand personality characteristics, and the 38-item TCI questionnaire was used to gather individuals' perceptions of team climate. While each item in the personality questionnaire was answered over a five-point Likert scale that ranged from *"very inaccurate"* to *"very accurate,"* every item in the TCI questionnaire ranged from *1 (strongly disagree/to a very little extent) through 3 (neutral/ to a moderate extent) to 5 (strongly agree/to a very great extent)*. In addition, a consent form was also included to comply with ethical principles and to acquire voluntary approval from members to participate in the research study. Besides information about the study's goal, participants' rights, and information on the storage and processing of collected data, the consent form also included demographic questions

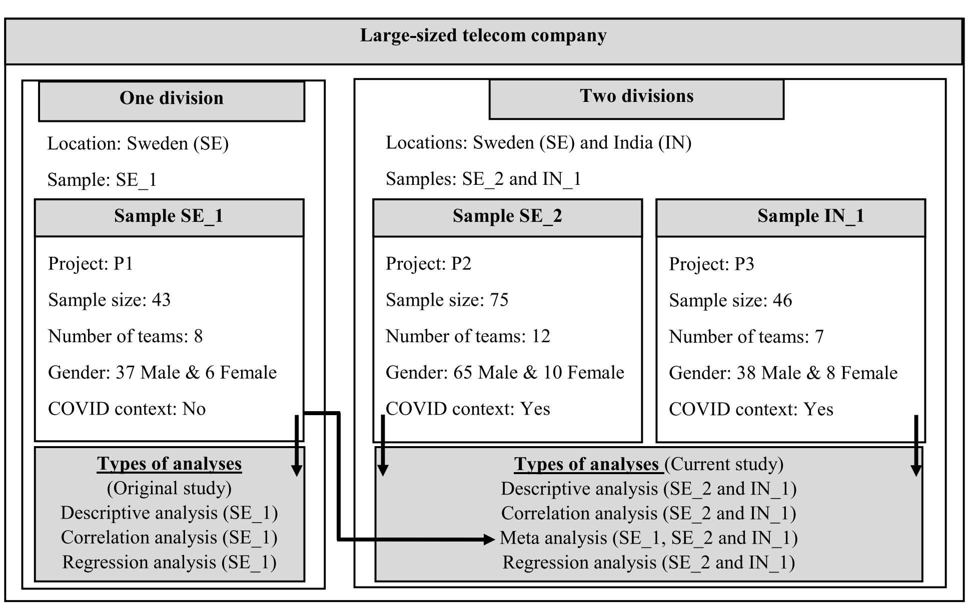


Fig. 1. Overview of samples and analyses from original and current studies.

to later aid in categorizing participants.

Although the original study's data collection technique involved a series of in-person interactive sessions (Vishnubhotla et al., 2020), it was not feasible to adopt a similar strategy for the two replications as the adversity of the COVID-19 pandemic during the times of the study forced legions of workers, including the employees at our collaborating company, to shift from office work to working from home (Russo et al., 2021). Therefore, we adapted to the prevailing conditions by employing a Web-based self-administered questionnaire that consisted of the consent form and the questions from the IPIP-NEO and TCI questionnaires. For maintaining the anonymity of subjects, we decided not to include any questions seeking the subject's name or email ID.

The survey was administered through the formsite website, which provides secure access and storage of responses. The flexibility offered by this platform for designing custom HTML and CSS forms was the main reason to choose it over the rest of the survey hosting websites. With respect to the survey questionnaire, multiple measures were taken to make sure that the respondents interpreted the questions correctly. A brief paragraph was included on the questionnaire's start page stating the study's aim and describing the types of questions presented in each section. To facilitate easy interpretation of questions, we stated that an explanation for certain phrases could be seen upon hovering the mouse pointer over the phrase (for example, "often feel blue" from IPIP-NEO).

Respondents were further presented the details of how the survey responses would be stored and processed. Here, it was informed that the responses would be stored anonymously and immediately after the survey, all the responses from the formsite platform would be downloaded and preserved securely on a system which would not be accessible to any person beyond our research team. While the respondents were given an opportunity to answer the questions in any order, all the questions in relation to IPIP-NEO and TCI were marked as mandatory as a precautionary measure to avoid the incidence of unanswered questions.

Before we started the actual data collection process, two professionals having experience of working in ASD teams from a different organization were requested to pilot our Web-based survey. In order to receive feedback on the questionnaire's presentation and to maximize its clarity, we asked these two professionals to inspect whether the language used in the questionnaire was simple and concrete. The suggestions from the professionals contributed towards improving the questionnaire's layout. Their suggestions, such as to include a description with tooltips for some more phrases and spreading the 120 questions of the IPIP-NEO across four pages instead of showing them all together, were incorporated into the questionnaire's final version.

Once the questionnaire was prepared for circulation, an email was sent out to the product owners with a brief text presenting the survey's aim and the link to the survey's instrument – an online questionnaire. After scrutinizing the questionnaire, the product owners subsequently issued their formal approval to proceed with the study and they selected participants for the study based on their availability and workload. The product owners from Sweden and India then set up separate virtual meetings via Zoom video conference call and invited our research team and all the selected professionals from respective divisions to the meeting.

In total, members from 19 agile teams participated in the virtual meetings. While the meeting with the Swedish division involved participation from 12 teams, the meeting with the Indian division comprised professionals from seven teams. Like the activities conducted within the in-person sessions of the original study (Vishnubhotla et al., 2020), the virtual meetings were initiated with a presentation where we briefly explained the overall goal of the study and how the findings of the study could be helpful. Further, the professionals were informed about what kind of information would be gathered. Finally, after informing the professionals that participation in the survey was voluntary, the Web link to the survey's questionnaire was sent as a message to everyone on the call and they were informed to fill in the responses at their own pace. During each virtual meeting, two members from our research group were present throughout the session in order to answer any questions regarding the study or questionnaire. On an average, the respondents took around 25 min to complete the questionnaire.

After each virtual meeting, the responses were downloaded as a spreadsheet from the formsite and a unique identifier was associated to each completed questionnaire. This was done by adding a new column to the spreadsheet. Next, the responses to the consent form that contained demographic questions were isolated from the responses to the IPIP-NEO and TCI questionnaires and were stored in a separate spreadsheet. The unique identifier was used to map the entries between the consent form responses and the rest.

Subjects:

A total of 121 software professionals agreed to participate in our survey. These are the total number of members from the 19 agile teams that were invited to participate and therefore, the response rate of the survey was also 100 %, as in the original study (Vishnubhotla et al., 2020). Among the 121 respondents, there were 75 subjects belonging to 12 teams in the sample acquired from the Swedish Division (SD), and 46 subjects who were part of seven teams from the Indian Division (ID). The SD subjects included 10 females (13.3 %) and 65 males (86.7 %). Whereas the ID subjects comprised of 8 females (17.4 %) and 38 males (82.6 %).

With respect to subjects' age, within the SD teams, most subjects were 31–35 years old (23 %), followed by members from the age group of 21–25 years (21 %). An overview of the subjects' age is presented in Table 1. In the case of ID teams, most subjects belonged to the age group of 31–35 years (46 %), followed by members who were 26–30 years old (24 %). It can be observed from Table 1 that more than half of the subjects in the current study were younger than 35 years. While 79 % of the respondents from ID teams were aged below 35 years, 52 % of the respondents from SD teams were younger than 35 years. Whereas in the original study's sample (Vishnubhotla et al., 2020), which was also from Sweden, 51 % of the respondents were under 35 years of age. Upon comparing the subjects' age between the original study (Vishnubhotla et al., 2020) and the current study, we can clearly see that a significantly large proportion of respondents from Indian teams were young compared to respondents from Swedish teams (respondents from the original study's sample and SD teams in this study).

Considering that sampling was not random, and respondents were selected based on their availability, one possible explanation for this situation could be, perhaps, the older professionals from India were holding more senior positions compared to those of the same age in Sweden and, therefore had less free time to participate in our study. A potential reason for this could be attributed to the differences in the wage levels of software employees between these countries (Rahman et al., 2021; Jalote and Natarajan, 2019). We hypothesize that the older professionals from India would be relatively more persuaded to look for managerial positions, which would yield them higher wage. Whereas the relatively higher wage levels in countries like Sweden could perhaps encourage professionals there to remain working as they are and lead to aiming for managerial positions at a relatively later age.

In relation to the subjects' roles, Table 2 presents the number of

Table 1
Distribution of respondents' age.

Age group	Number of professionals	
	Teams from Sweden	Teams from India
21 - 25 years	6 (8 %)	4 (9 %)
26 - 30 years	16 (21 %)	11 (24 %)
31 - 35 years	17 (23 %)	21 (46 %)
36 - 40 years	13 (17 %)	8 (17 %)
41 - 45 years	9 (12 %)	1 (2 %)
46 - 50 years	9 (12 %)	1 (2 %)
51 - 55 years	3 (4 %)	–
56 - 60 years	2 (3 %)	–

Table 2
Roles of software professionals.

Role	Number of professionals	
	Teams from Sweden	Teams from India
Designer	2 (2.6 %)	–
Design lead	1 (1.3 %)	–
Build master	1 (1.3 %)	–
Domain expert	2 (2.6 %)	3 (6.5 %)
Software developer	51 (68 %)	27 (58.7 %)
Project manager	1 (1.3 %)	1 (2.2 %)
CI engineer	1 (1.3 %)	–
Technical expert	6 (8 %)	4 (8.7 %)
Scrum master	1 (1.3 %)	1 (2.2 %)
Team lead	9 (12 %)	2 (4.3 %)
Tester	–	8 (17.4 %)

respondents in each role. Like in the original study (Vishnubhotla et al., 2020), it was observed that more than half of the subjects in this study had as role 'Software developer' (SD: 68 %; ID: 58.7 %), followed by (SD: team lead (12 %); ID: tester (17.4 %)), and technical expert (SD: 8 %; ID: 8.7 %).

Regarding the question asking subjects to indicate the country they felt they belong the most to, most SD subjects answered Sweden (84 %), followed by India (5.3 %), Lithuania (2.6 %), Spain (2.6 %), England (1.3 %), Iran (1.3 %), Macedonia (1.3 %) and Pakistan (1.3 %). In the case of ID, all subjects answered India.

In order to analyze the data from both samples in the same way as we did in the original study (Vishnubhotla et al., 2020), each person's responses for the personality test were initially entered into the online version of the IPIP-NEO instrument. This online version compares an individual's scores with reference personality data and generates scores that are percentile estimates after adjusting for age and gender (J.A. Johnson, Interpreting individual IPIP scale scores, (n.d.)). Furthermore, in our analysis, we used the same categorization as specified by the IPIP-NEO narrative report, where each person is classified as *low, average, or high* in a personality trait. In the case of responses to TCI, the scores of the four team climate dimensions were calculated by averaging the scores awarded to all the questions associated with a dimension.

After generating the scores for personality traits and team climate dimensions, we used the statistical techniques such as descriptive analysis (to summarize observations from personality and team climate scores), correlation analysis (inter-correlations and correlations between personality traits and team climate factors), meta-analysis (to aggregate the findings based on correlations from the original study and the two samples) and regression analysis (to understand significant predictors of team climate factors), and finally checked the accuracy of the prediction models. All the analyses were performed with the support of R-programming language and statistical software environment.

We used the guidelines proposed by Santos et al. (A. Santos et al., 2021) for analyzing the two iterations of replications in this study. We adhered to the four-step procedure outlined by them. However, note that we did not adopt in our study the final step from their guidelines. The final step is linked to conducting exploratory analyses to identify experiment level moderators and participant-level moderators. However, one of the limitations of exploratory analyses is that the moderator effects can be confounded if multiple simultaneous changes are made across replications (A. Santos et al., 2021). Since the replications in our study differ from the original study in terms of aspects such as the data collection mechanism (Web-based survey) and subject type (subjects recruited from different divisions & countries of company A), we chose not to conduct exploratory analyses in this study. The details of how different analyses from our study correspond to each of the three steps from the guidelines are presented next.

Step 1: Describe the characteristics of the participants using appropriate descriptive statistics and visualizations. This was accomplished as a part of the descriptive analysis presented in Section 4.1.
Step 2: Use consistent statistical techniques to pre-process, describe and analyse the data of each replication. This was accomplished by testing whether various assumptions were met in relation to correlation and regression analyses (Section 4.1, Section 4.2, and Section 4.4).
Step 3: Select suitable aggregation techniques to provide joint conclusions. This was accomplished by conducting a meta-analysis of correlations by including samples from the current and the original study (Section 4.3).

4. Results

This section presents the findings from our statistical analyses. Here, we report the findings of each analysis by comparing them with the original study's findings.

4.1. Descriptive analysis

4.1.1. Personality traits (RQ.1)

Each subject's percentile estimates for the five personality traits generated from the online IPIP-NEO test were used to compute basic statistics like mean, median, standard deviation, and Coefficient of Variation (CV). The details about the distribution of the percentile estimates for each personality trait across the two samples (we labeled the current study's sample from Sweden as SE_2 and the sample from India as IN_1) are presented in Table 3. The variations in personality traits' scores across the subjects can further be seen from the box-and-whisker plot for the SD teams (see Fig. 2) and the plot for the ID teams (see Fig. 3). Looking at the distribution of scores across the five traits, we can see that, except for the extraversion scores for the ID teams, the scores for the rest of the traits across both samples do not reveal any outliers. This indicates that, apart from one subject's extraversion score from the IN_1 sample, there were no other subjects across the two samples whose personality scores did not fall within the boxplot's standard quartile ranges. The descriptive statistics from Table 3 indicate that the subjects from both samples possess average levels of all five personality traits when compared to people of similar sex and age.

4.1.1.1. Teams from Sweden. From the distribution of personality scores within SE_2 presented in Fig. 2, we can observe that the openness to experience scores are mostly *low* for close to 50 % of the respondents (Q1 and Median). Conversely, close to 50 % of the agreeableness scores were *high* (Q3 and maximum). As for extraversion, conscientiousness, and neuroticism, most scores are within the *average* category; however, in the case of extraversion and neuroticism, slightly more than 25 % of the scores are *low*. In regard to the distribution of scores, conscientiousness is the only one that shows a close to normal distribution. Extraversion, neuroticism, and openness distributions are all right skewed, which informs us that there is higher conformity among the scores lying below the median. The opposite applies to Agreeableness.

The original study also recruited eight teams from Sweden (Vishnubhotla et al., 2020), and upon comparing the distribution of scores from the original study's sample (the original study's sample from Sweden is labeled as SE_1) with the current study's sample from Sweden (SE_2), it can be observed that the level of agreeableness was relatively higher in SE_2. However, an important point to note is that the level of neuroticism in SE_2 was relatively higher than that of SE_1. While the neuroticism trait recorded the lowest median score among all personality traits in the original study's sample, in the current study's SE_2 sample, the median for neuroticism is higher than the medians for two other traits (extraversion and openness to experience). By comparing the central tendencies, we further observed that the levels of extraversion

Table 3
Distribution of personality traits' scores in SE_2 and IN_1 samples.

Personality trait	Teams from Sweden (SE_2)				Teams from India (IN_1)			
	Mean	Median	Standard deviation	CV (%)	Mean	Median	Standard deviation	CV (%)
Extraversion	41.266	39	24.608	59.632	58.478	57	21.050	35.997
Agreeableness	60.906	68	25.020	41.079	62.021	64	25.360	40.888
Conscientiousness	54.373	56	26.676	49.061	68.978	75	24.361	35.317
Neuroticism	43.120	45	24.025	55.717	45.347	46	21.482	47.371
Openness to experience	33.946	32	21.802	64.225	43.173	46.5	23.236	53.820

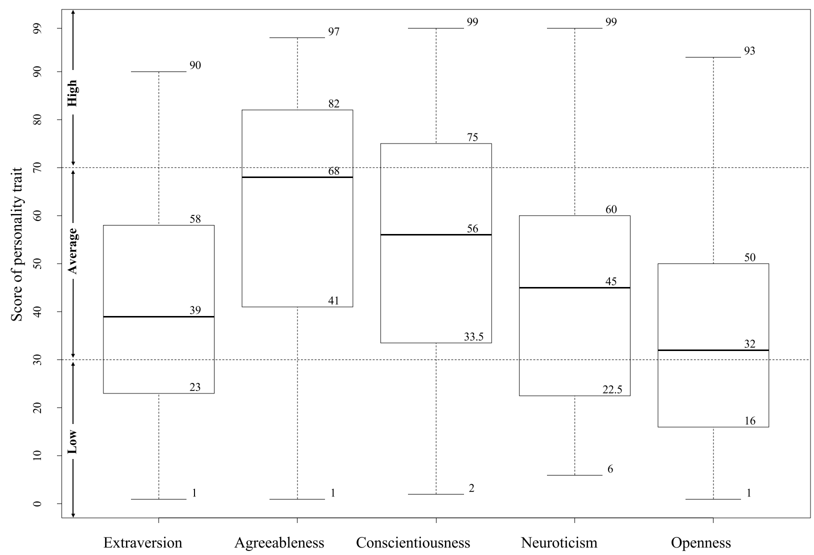


Fig. 2. Personality traits' scores in SE_2 sample.

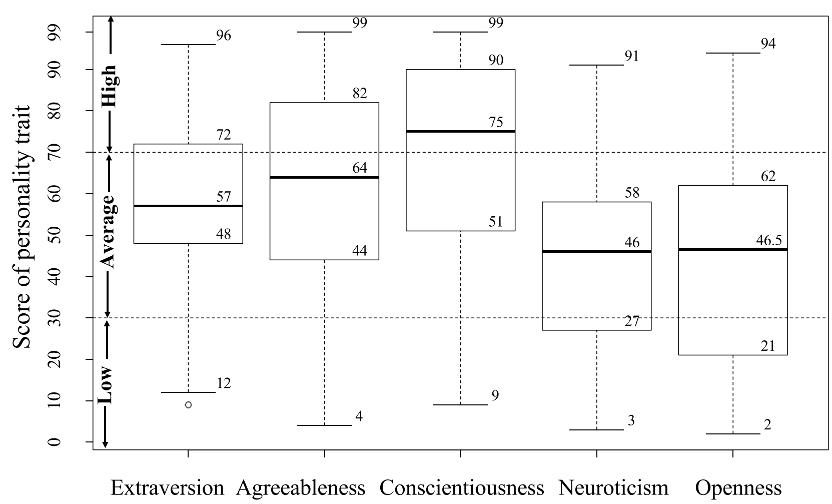


Fig. 3. Personality traits' scores in IN_1 sample.

and conscientiousness traits were significantly lower, and the level of agreeableness was significantly higher in the case of SE_2. Despite variations in the median values of various traits' scores, it is important to note that, both in SE_1 and SE_2, all the median levels fall under the average category (see Fig. 2), as presented in the IPIP narrative report (J. A. Johnson, IPIP-NEO narrative report, (n.d.)).

4.1.1.2. Teams from India. The distribution of personality scores within IN_1, as shown in Fig. 3, indicates that at least 25 % of the scores for neuroticism and openness to experience are *low* (minimum and some data points above Q1). Whereas, in the case of extraversion, agreeableness, and conscientiousness traits, most scores are either *medium* or *high*. In the case of conscientiousness, more than 50 % of the scores are *high* (Median, Q3, and maximum). Except for neuroticism and openness to experience, which suggest close to normal distributions, all the other distributions are left-skewed, which informs us that there is higher conformity among the scores lying above the median.

Upon comparing the distribution of personality scores among the two samples from the current study (see Table 3), we can notice that the medians for the three personality traits (extraversion, conscientiousness, and openness to experience) were relatively higher in IN_1. In specific, the median levels of extraversion and conscientiousness were significantly high in IN_1 in comparison to SE_2. This can also be seen upon comparing Fig. 2 and Fig. 3, where the scores of extraversion and conscientiousness traits were at least *average* for more than 75 % of the subjects in Fig. 3. However, the median level of agreeableness was slightly higher in SE_2, and in terms of neuroticism, the median levels of SE_2 and IN_1 were almost the same. Upon comparing the distribution of scores with the original study, the median levels of all five personality traits were relatively high in IN_1.

4.1.1.3. Observations from both samples. The distribution of scores from Table 3 reveals that the subjects from SE_2 and IN_1 samples possess relatively higher levels of agreeableness compared to the subjects from the original study (SE_1). Like the sample from SE_1, the median levels of scores with respect to all the five traits (see Fig. 2 and Fig. 3) fall in the range of scores that are classified as *average* by the IPIP-NEO narrative report. Therefore, we can understand that, compared to people of similar sex and age, the professionals from both samples are curious and concerned about others' needs but are, in general, unwilling to sacrifice their own responsibilities (average agreeableness). Not only do they like spending time with others, but they also take pleasure in having their alone time (average extraversion). Further, we can understand that the professionals from both samples are good at organizing, planning, and persevere their goals with determination (average conscientiousness). Although the professionals enjoy everyday work, they are still inclined towards trying new things (average openness to experience). The data shows that the emotional reactivity level of professionals is typical of the general population (Vishnubhotla et al., 2020), and they can, in general, cope with exasperating situations (average neuroticism) (Hogan et al., 1997).

An important observation from comparing the distribution of scores is that the level of neuroticism was seen to be higher in both SE_2 and IN_1 in comparison to SE_1. Since neuroticism is a trait associated with negative characteristics, the higher levels observed in SE_2 and IN_1 indicate that subjects were relatively more anxious and less calm.

Upon comparing the Coefficient of Variation (CV) values listed in Table 3 for SE_2 and IN_1, the CV for openness to experience scores appear to be higher than the rest of the distributions in both samples. This indicates that, for the openness to experience trait, dispersion in the scores around mean was greater than any other distribution.

Takeaway points:

- The median levels of all five personality traits were relatively high in IN_1 compared to the original study (SE_1).
- Subjects from the current study's samples (SE_2 and IN_1) possess relatively higher levels of agreeableness compared to those from the original study (SE_1).
- The higher levels of agreeableness observed in the current study's samples (SE_2 and IN_1) indicate that subjects were relatively more trustworthy, humble, and possess the ability to get along well.
- Upon comparing the samples from current study (SE_2 and IN_1) and the original study (SE_1), the level of neuroticism in SE_2 and IN_1 was relatively higher.
- The higher levels of neuroticism observed in the current study's samples (SE_2 and IN_1) indicate that subjects were relatively more anxious and worried.

4.1.2. Team climate factors (RQ.2)

The responses to the 38 questions in TCI were used to compute the scores for the four team climate dimensions, and the distribution of scores for these four dimensions is shown in Table 4 for SE_2 and IN_1.

4.1.2.1. Teams from Sweden. In relation to the SE_2, we can notice from the mean and median values reported in Table 4 and the box-and-whiskers plot in Fig. 4 that the distribution of scores for all the dimensions was negatively skewed, with more pronounced negative skewness for participative safety and support for innovation. This suggests that there was a higher conformity among the scores distributed above the median for all four dimensions. When looking into the measures of tendency from Table 4, for the four dimensions, we notice that participative safety presented the highest mean and median, and support for innovation showed the lowest mean and median. It is worth noting that a similar pattern was observed in the original study (SE_1) (Vishnubhotla et al., 2020). Almost 50 % of the scores for participative safety can be seen to be above 4 (median). In the case of team vision (see Fig. 4), we can observe the size of the box (interquartile range) in the box-and-whiskers plot to be compact, indicating greater conformity among those scores. The CV for team vision was 14.36 %, informing us of less dispersions in the scores around the mean.

4.1.2.2. Teams from India. With respect to IN_1, the details of the distribution of team climate scores are presented in Table 4 and in the form of a box-and-whiskers plot shown in Fig. 5. Like the observation from SE_2, the distributions of scores for all four dimensions were negatively skewed in IN_1, with more pronounced skewness for task orientation and support for innovation. The CV of 25.78 % for task orientation and 24.41 % in the case of support for innovation reveal that there were a lot of dispersions in the scores around the mean. The mean and the median values for participative safety were the highest (see Table 4), and the lowest mean and median were for task orientation. Finally, Fig. 5 informs us that 50 % of the scores in the case of participative safety and team vision can be seen to be above four (median).

4.1.2.3. Observations from both samples. Upon inspecting Table 4 and Figs. 3 and 4, we can see that within SE_2 and IN_1, the trends across the four team climate dimensions followed a similar pattern in terms of skewness. However, we can notice that the size of the boxes and whiskers in Fig. 4 is relatively compact when compared to those in Fig. 5. This informs us that there was higher conformity among the team climate scores within SE_2. Overall, insights from the distribution of team climate scores from both the samples reveal that, like SE_1, subjects considered their team atmosphere to provide them a safe forum for generating ideas where every team member's view was likely to be heard and acknowledged (high participative safety).

Takeaway points:

- In both SE_2 and IN_1 samples, the trends across the four team climate dimensions followed a similar pattern in terms of skewness. Moreover, a similar pattern was observed between Swedish samples from the current study (SE_2) and our previous study(SE_1).
- A higher conformity (compact box-plots) among team climate scores was observed in the SE_2 sample.

Table 4
Descriptive statistics for team climate for the SE_2 and IN_1 samples.

Team climate dimension	Teams from Sweden (SE_2)				Teams from India (IN_1)			
	Mean	Median	Standard deviation	CV (%)	Mean	Median	Standard deviation	CV (%)
Vision	3.866	3.909	0.555	14.365	3.802	3.909	0.623	16.389
Task orientation	3.756	3.857	0.643	17.143	3.568	3.714	0.920	25.787
Support for innovation	3.67	3.75	0.747	20.379	3.711	4	0.906	24.416
Participative safety	3.962	4	0.686	17.327	3.932	4.041	0.771	19.621

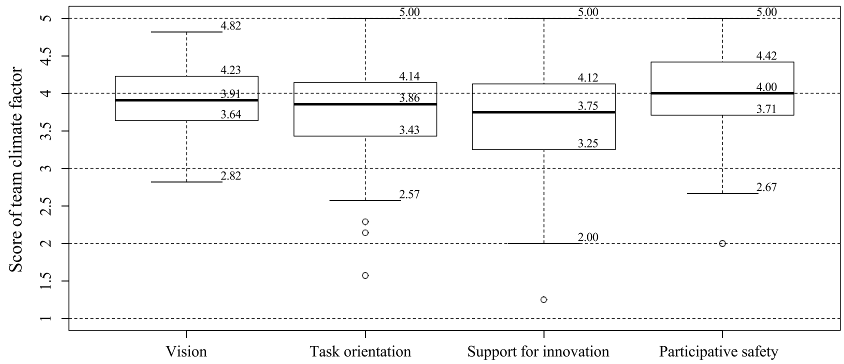


Fig. 4. Scores for Team climate dimensions for the SE_2 sample.

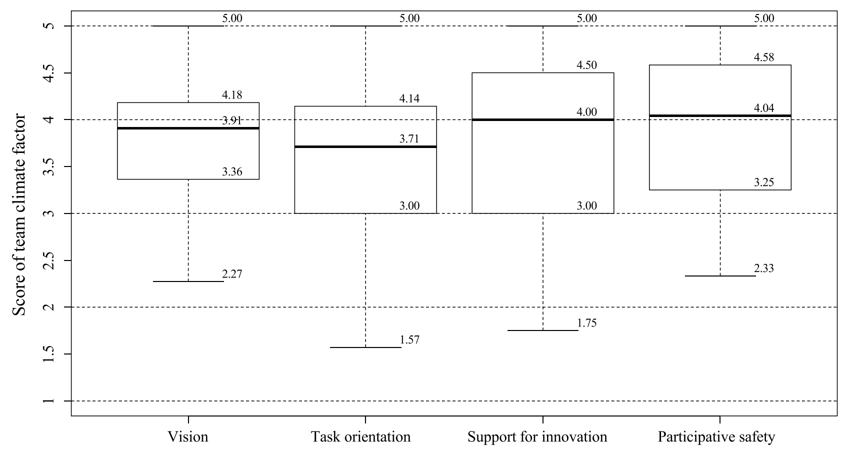


Fig. 5. Scores for the Team climate dimensions for the IN_1 sample.

- Distribution of scores within both SE_2 and IN_1 samples indicate that subjects considered their team atmosphere to provide them a safe forum for generating ideas (high participative safety).

While the team climate scores can be aggregated at the team level by demonstrating agreement among team members' scores, it is important to evaluate whether the data follows a normal distribution before performing parametric tests for computing the Intra-class Correlation Coefficient (ICC) and Pearson correlation coefficient (Vishnubhotla et al., 2020). Therefore, we performed the Shapiro-Wilk test on the data from the two samples, SE_2 and IN_1, to assess whether the data was normally distributed. We set the level of significance to 0.01, as done in the original study (Vishnubhotla et al., 2020) and also in other SE studies (Salman and Turhan, 2018; Rahmani and Khazanchi, 2010; Silva et al., 2019).

4.1.2.4. Testing for normality. Regarding the SE_2 sample, the Shapiro-Wilk test showed that, except for the agreeableness trait, all the other four personality traits presented scores that complied with a normal distribution. That is, with respect to extraversion ($p = 0.049$), conscientiousness ($p = 0.027$), neuroticism ($p = 0.013$), and openness to experience ($p = 0.052$) traits, the data were normally distributed. We transformed the scores of agreeableness variable using different transformations such as square root, cube root, and logarithmic transformations; however, none contributed to bringing the distribution closer to a normal shape distribution. In the case of team climate scores, all were normally distributed, with the following scores for team vision ($p = 0.047$), task orientation ($p = 0.036$), support for innovation ($p = 0.017$), and participative safety ($p = 0.064$). The scores for the IPTC variable, computed by averaging the scores of the team climate dimensions, also followed a normal distribution ($p = 0.055$).

As for the IN_1 sample, the Shapiro-Wilk test indicated that the scores of all the personality traits, except for conscientiousness, were normally distributed. The p-values observed across the personality traits are extraversion ($p = 0.46$), agreeableness ($p = 0.063$), neuroticism ($p = 0.83$), and openness to experience ($p = 0.074$). Here, we also applied several transformation functions to bring the scores of the conscientiousness variable closer to a normal distribution; however, no significant improvement in its normality was observed. With respect to the team climate dimensions, all the scores were normally distributed, with the following p values: for team vision ($p = 0.623$), task orientation ($p = 0.162$), support for innovation ($p = 0.028$), and participative safety ($p = 0.013$). Finally, the IPTC scores also followed a normal distribution ($p = 0.355$).

Takeaway points:

- The scores of all the four team climate dimensions across both the samples from the current study (SE_2 and IN_1) were normally distributed.
- In the SE_2 sample, the scores of four personality traits (except for the agreeableness trait) complied with a normal distribution. Similarly, in the IN_1 sample, except for the conscientiousness trait, the scores of the rest of the four traits were normally distributed.

4.1.2.5. Aggregation of scores. Based on the observations from the Shapiro-Wilk test, like the original study (Vishnubhotla et al., 2020), the ICC(1) indices were computed for the scores of personality traits and team climate dimensions by setting the significance level for the F-test to 0.05. In SE_2, except for the scores of conscientiousness, the ICC(1) indices for the rest of the traits fell below 0.20, which is considered as a threshold over which aggregation of scores is justified (Acuña et al., 2008). So, most personality traits could not be aggregated to the team level. In relation to team climate dimensions, the scores of task orientation (0.28), participative safety (0.32), and IPTC (0.29) factors were observed to be above the threshold, and therefore the scores within teams from SE_2 were aggregated for these factors. The average scores for task orientation and participative safety, together with the average team climate scores (IPTC) for the 12 teams, are presented in Fig. 6.

In the case of IN_1, the ICC(1) indices for all the personality traits were observed to be below 0.20, and hence, the aggregation of these scores was also not possible. However, with respect to the scores of team climate dimensions, team vision (0.21), task orientation (0.52), support for innovation (0.43), participative safety (0.53), and IPTC (0.47) were observed to be above the threshold level. The average scores for the four team climate traits, together with the IPTC average for the seven teams in IN_1, are presented in Fig. 7.

Upon inspecting the team level characteristics from both the samples SE_2 and IN_1 (Fig. 6 and Fig. 7), we can notice the average values of team climate scores to be falling within the range of 3.5 to 4.5 (except for the I.T3 and I.T5 teams), thus indicating the perceived team climate among most teams from the two samples to be ranging between positive to highly positive. This was in line with the observations from the original study's sample SE_1.

4.1.3. Observations and implications

Comparison of median levels of personality traits' scores between the Swedish samples from the original study (SE_1) and the current study (SE_2) indicated higher levels of agreeableness and neuroticism and lower levels of extraversion and conscientiousness in the latter sample. On the other hand, the median levels of all the scores were relatively high in the Indian sample (IN_1). Such variations in trends across the samples could mainly be attributed to the change in context associated with the execution of each survey instance. Although all the samples were acquired from two divisions of the same company, each sample consisted of different groups of people working on different projects. We also believe other factors like age, ethnicity, work culture, and contemporary factors such as the COVID-19 pandemic to be potential reasons behind the shift in trends (Kang et al., 2006). Therefore, additional data via further replications would be imperative to analyze the aforementioned trends; regarding future studies inspecting personalities and team climate characteristics in an organization, we recommend executing investigations in different contexts as that would help in acquiring robust knowledge and aid towards generalizing findings to a wider group.

Within ASD, tasks such as generative design, documentation, continuous collaboration in relation to developing code, refactoring, prioritizing customer needs, testing, and frequent code releases all entail a highly charged and stimulated team environment. Practices such as pair programming also call for greater understanding among team members. In such collaborative environments, while low to moderate levels of conflict could contribute towards enhancing the performance of an agile team, high levels of conflict are very likely to be detrimental to team effectiveness (Domino et al., 2003). The ability to cope with conflicts depends on one's personality (Balijepally et al., 2006).

In agile teams, there is a risk of conflicts occurring due to disagreements and differences of opinion among team members (Balijepally et al., 2006). The personalities of team members are an additional dimension, apart from skills and experience, which managers should perhaps consider while assembling teams. By guaranteeing anonymity and secure storage of the sensitive personality data, as detailed in Section 3, investigating the personality profiles of both successful and not so successful teams would be a logical first step. This approach would lead to the gathering of cumulative evidence that managers could use to fine-tune their team-building criteria and also to inform additional support towards improving existing personality traits that are not so conducive towards a good team climate.

Two approaches where team members' personality traits can be used for establishing agile teams are, elevation and variability (Balijepally et al., 2006). Elevation of a trait is associated with measuring the aggregated or average of individual scores within an agile team. It is implicit in this approach that the high score of a team member on a particular trait would compensate for the low score of another member. Thereby, the addition of an individual would potentially affect a team. On the other hand, variability indicates the homogeneity or heterogeneity of a particular trait in a team. The CV, discussed in Section 4.1.1.3, is a measure of dispersion of scores around the mean. By comparing the CV for various configurations of members, dispersions in the scores of a trait can be monitored.

Previous SE studies (Balijepally et al., 2006) and (R. Baumgart, M. Hummel, Personality traits of scrum roles in agile software development teams - a qualitative analysis, (n.d.)) discussed the positive impact of personality traits such as agreeableness and conscientiousness towards establishing a successful agile team. On the other hand, low levels of neuroticism were reportedly related to a relaxed team atmosphere and improved coordination and stability within teams. Elevation in agreeableness is expected to promote cooperation, open communication, and compliance with team goals. However, variability in agreeableness is considered detrimental to the performance of a team, as even one

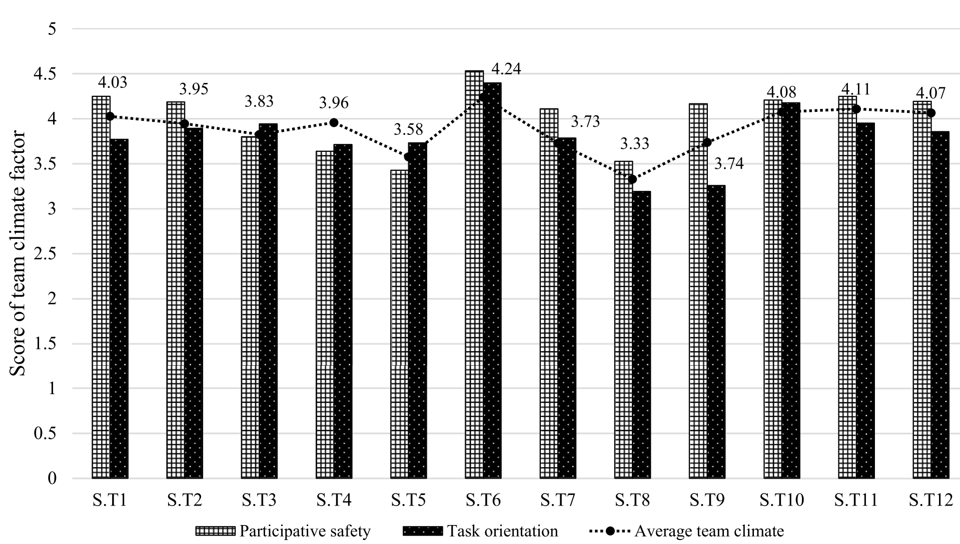


Fig. 6. Team climate trends across 12 teams from SE_2 sample.

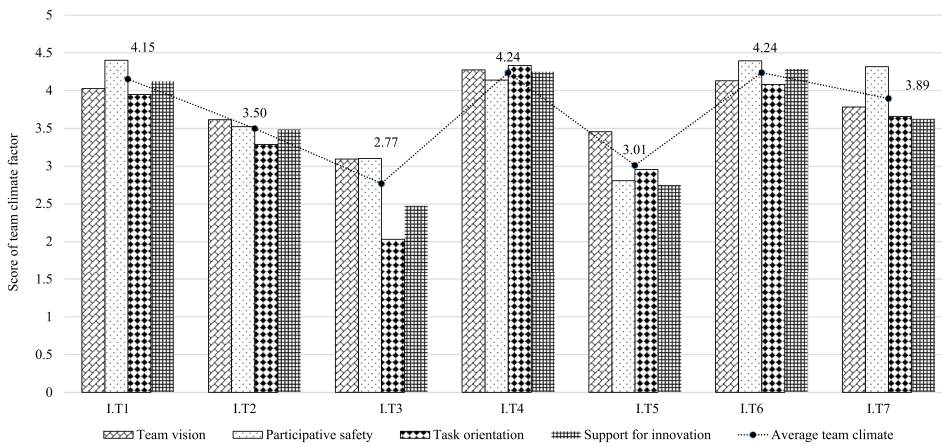


Fig. 7. Team climate trends across 7 teams from IN_1 sample.

disagreeable team member could disrupt cooperation. In agile teams, the presence of agreeable team members, i.e., members whose agreeableness scores are not very different from each other, would be highly desirable. Elevation in conscientiousness is expected to promote perseverance towards a team's goal completion and commitment to tasks. Conversely, low conscientiousness may lead to social loafing. Variability in conscientiousness within a team is expected to lower team performance as members high on this trait may detest loafing by low conscientious members. In the case of neuroticism, variability may not be conducive to team working as even one emotionally unstable person could impair the performance of the team by affecting cohesion and cooperation (Balijepally et al., 2006). Such unbalances in conscientiousness and neuroticism are not to be understood as means for removing team members; these should be employed as means to provide members with the necessary support and training in order to improve their conscientiousness/lower their neuroticism.

In order to monitor the conformity on the perceived climate among team members, managers could use team members' IPTC scores to generate box-and-whiskers plots. In such plots, small-size boxes and whiskers indicate a more homogeneous distribution of scores, which translates as a higher level of conformity among team members' perceptions of climate. Moreover, the data points falling outside the box-and-whiskers plot would indicate outliers, which could prompt a more detailed discussion with the team(s) for which outliers were identified. Note that outliers should not be excluded as they affect the climate score computed at the team level. In agile teams, the presence of members whose perception of overall team climate is not very different from others would be highly desirable. In order to have an aggregated team climate score, managers can perhaps compute the ICC(1) index as outlined in Section 4.1.2.5. Such an aggregated score gives a better idea of the range under which the team-level perceptions of climate fall. The scores falling within the range of 3 to 5 would indicate the perceived team climate to be ranging between positive and highly positive (Vishnubhotla et al., 2020).

4.2. Correlation analysis

In order to investigate correlations between personality traits, correlations between team climate domains, and then correlations between personality traits and team climate dimensions, we employed the Pearson correlation test, which is appropriate for use with data measured at least on an interval scale. Similar to the original study (Vishnubhotla et al., 2020), we refrained from using any of the non-parametric methods for correlation analysis as they tend to show higher than real values for distributions with the presence of tied data, like the scores of different variables in our samples (Klebanov, 2006).

The correlation analyses presented herein relate mainly to the data collected via the two instances of replication. That is, correlation coefficients between various variables from the current study's samples SE_2 and IN_1 are primarily presented and discussed in the forthcoming sub-sections. This is done to emphasize and accommodate for findings from the new samples and to make the tables with correlation matrices more concise and readable. However, readers can find the results of correlation analysis from the original study's sample SE_1 listed under Appendix B. Nonetheless, it is important to note that observations from the comparison of significant correlations between the original study's sample (SE_1) and the current study's samples SE_2 and IN_1 are discussed under each sub-section below.

Before proceeding with the correlation analyses, checks were performed to ensure the assumptions for Pearson correlation were met. This included the Shapiro-Wilk test, plotting scatter plots to investigate the relationship between variables, and using the Q-Q plots to inspect whether variables under consideration were normally distributed. The details of correlation coefficients observed between variables and the information on respective effect-sizes based on Cohen's rules, as employed in the original study (Vishnubhotla et al., 2020), are presented next.

4.2.1. Relationship between personality traits (RQ.1)

In the light of very limited studies reporting the inter-correlations among the personality traits (Vishnubhotla et al., 2020), we start the correlation analysis by emphasizing on the inter-correlations for both samples SE_2 and IN_1. The Pearson correlation coefficient (r) values for various combinations of normally distributed personality scores are presented in Table 5. Further, the level of significance for correlations is also highlighted in Table 5.

Within SE_2, we observed medium-sized positive effects between extraversion and two personality variables (conscientiousness ($r = 0.35$) and openness to experience ($r = 0.33$) from Table 5). Both correlations were significant at the 0.01 level. These positive correlations inform us that, within the sample SE_2, an increase in a person's ability to socialize and express (extraversion) corresponded to a rise in the ability to plan and organize effectively (conscientiousness) and a rise in the ability to imagine and try new things (openness to experience).

Besides, we also observed large-size negative effects between neuroticism and two other traits (extraversion (r = 0.51) and conscientiousness ($r = 0.49$) from Table 5). In both cases, the correlations were significant at the level of 0.0001. Since neuroticism, as a trait, is associated with negative characteristics, the large size negative effects indicate that, within SE_2, a rise in the level of a person's anxiety (neuroticism) corresponded to a sharp fall in the ability to interact socially (extraversion) and a major decline in the level of confidence

Table 5
Pearson correlation matrix for personality traits.

	Extraversion		Agreeableness		Conscientiousness		Neuroticism	
	SE_2	IN_1	SE_2	IN_1	SE_2	IN_1	SE_2	IN_1
Agreeableness	–	0.03						
Conscientiousness	0.35**	–	–	–				
Neuroticism	−0.51***	−0.18	–	−0.14	−0.49****	–		
Openness	0.33**	0.40**	–	0.45**	0.13	–	−0.19	−0.29

**** $p < 0.0001$.
*** $p < 0.001$ and.
** $p < 0.01$.

(conscientiousness).

In the case of IN_1, we observed medium-size positive effects between openness to experience and two personality traits (extraversion ($r = 0.40$) and agreeableness ($r = 0.45$) from Table 5). The correlations were significant at the 0.01 level. These two positive correlations inform us that among the seven teams from India, an increase in team members' ability to be imaginative and attempt new tasks corresponded to a rise in the members' ability to express and be optimistic (extraversion) and a rise in member's positive peer relationships (agreeableness).

By comparing the findings from SE_2 with the original study, it was interesting to see the two medium-sized positive effects in relation to extraversion and the two large-size negative effects in relation to neuroticism being replicated across the teams from Sweden within company A. When the findings from IN_1 were compared with the original study, the two medium-sized positive effects in relation to openness to experience were noticed to be replicated.

Via the comparison of the effects across the SE_1, SE_2, and IN_1 samples, we saw that the medium-size positive effect between openness to experience and extraversion ($r = 0.33$ in SE_2, 0.40 in IN_1 and 0.40 in SE_1) replicated across all the three samples, where the correlations were significant at the 0.01 level. This replication of effects from the current study adds more evidence to the findings from the original study and informs us that such an association is highly likely to be observed among other teams spread across different divisions within company A.

4.2.2. Relationship between team climate factors (RQ.2)

In relation to the samples SE_2 and IN_1, results showed that all four team climate variables were statistically significantly positively correlated with each other at the 0.01 level. The correlation coefficients from this inter-correlation analysis are presented in Table 6. Among the various positive effect sizes, we observed a large size positive effect between task orientation and support for innovation with respect to both SE_2 and IN_1 samples ($r = 0.69$ in SE_2 and $r = 0.84$ in IN_1). In both cases, the correlation was significant at the 0.0001 level, and the positive direction of the correlation indicates that, within both samples, an increase in an individual's perceived level of team effort exerted towards achieving excellence (task orientation) corresponded to a strong rise in the individual's perceived level of idea sharing within a team (support for innovation).

It is worth noting that the large effect size in relation to task orientation and support for innovation variables ($r = 0.62$ and $p < 0.0001$) and other effect sizes identified in the original study were also present in both SE_2 and IN_1. This suggests that such an association across team climate traits is highly likely to be observed among teams from different divisions within company A.

4.2.3. Relationship between personality traits and team climate dimensions (RQ.3)

The Pearson correlation was further applied over the scores of normally distributed variables to explore the relationship between personality traits and team climate dimensions. The IPTC scores computed by averaging the scores of all the team climate traits were also included in this correlation analysis. The correlation coefficients observed in this analysis are presented in Table 7.

In relation to SE_2, we noticed a significant negative correlation between neuroticism and the rest of the team climate variables (team vision ($r = 0.24$), task orientation ($r = 0.25$), support for innovation ($r = 0.27$), participative safety ($r = 0.25$) and IPTC ($r = 0.30$)). The five correlations were significant at the 0.05 level. The negative direction of the relationship between neuroticism and other variables indicates that a rise in the level of a person's anxiety (neuroticism) corresponded to a decrease in the perceived levels of team climate aspects such as the clarity of goals set within a team (team vision), the effort exerted by the team towards achieving excellence (task orientation), idea sharing within the team (support for innovation), the team as a safe, non-threatening forum for discussing ideas (participative safety) and the overall perception of climate within a team (IPTC).

With respect to IN_1, a medium-sized negative effect was seen between neuroticism and participative safety variables. The negative correlation was significant at the 0.05 level. Neuroticism is related to a person's state of emotional stability. While a person who is less neurotic tends to appear confident and poised, a highly neurotic person is likely to be apprehensive and insecure (Vishnubhotla et al., 2020). On the other hand, participative safety relates to what a person feels about the level of trust within a team while expressing one's opinions and ideas (Vishnubhotla et al., 2020). The negative direction of the relationship in IN_1 informs us that the increase in a person's anxiety levels corresponds to the fall in the person's perception about the team being interpersonally non-threatening and safe for sharing ideas.

Upon comparing the coefficients from the two samples in Table 7, we noticed that neuroticism was the only personality variable where significant correlations were observed in relation to team climate variables.

Table 6
Pearson correlation matrix for team climate factors.

	Team vision		Task orientation		Support for innovation	
	SE_2	IN_1	SE_2	IN_1	SE_2	IN_1
Task orientation	0.51****	0.80****				
Support for innovation	0.55****	0.66****	0.69****	0.84****		
Participative safety	0.50****	0.60****	0.50****	0.66****	0.57****	0.68****

****$p < 0.0001$.
***$p < 0.001$ and.
**$p < 0.01$.

Table 7
Pearson correlation matrix for personality traits and team climate factors.

	Team vision		Task orientation		Support for innovation		Participative safety		IPTC	
	SE_2	IN_1	SE_2	IN_1	SE_2	IN_1	SE_2	IN_1	SE_2	IN_1
Extraversion	0.13	0.21	0.12	−0.01	−0.04	0.12	−0.04	0.23	0.02	0.16
Agreeableness	–	0.18	–	0.01	–	0.01	–	0.02	–	0.06
Conscientiousness	0.19	–	0.07	–	−0.04	–	0.17	–	0.12	–
Neuroticism	−0.24*	−0.09	−0.25*	−0.12	−0.27*	−0.06	−0.25*	−0.34*	−0.30**	−0.19
Openness	−0.02	0.06	−0.06	−0.03	−0.07	0.09	−0.09	0.23	−0.10	0.11

**$p < 0.01$ and.
*$p < 0.05$.

Further, we noted that the medium size negative effect between neuroticism and participative safety variables to be common in both samples. One of the main contributions of this study is to assess whether the relationships identified in our original study (Vishnubhotla et al., 2020) are corroborated using new and independent data sets. The current study identified numerous relationships; however, the comparison between the correlation coefficients in Table 7 and the findings from the original study showed no common relationship between these studies. Given that the development contexts for the different studies varied, we believe that an avenue for further investigation would be to replicate these studies in contexts that are as similar as possible to the original study (Dybå et al., 2012).

Takeaway points:

- In the case of the SE_2 sample, a medium-sized negative effect was observed between neuroticism and all the team climate dimensions, where the correlations were significant at 0.05 level.
- A medium-sized negative effect between neuroticism and participative safety variables was observed to be common to both the samples from the current study (SE_2 and IN_1).
- The negative relationship indicates that an increase in a person's anxiety levels (neuroticism) corresponds to a fall in the person's perception about the team being a safe forum for sharing ideas (participative safety).

4.2.4. Observations and implications

Based upon inter-correlations among personality variables, we noticed the medium-sized positive effect between openness to experience and extraversion variables to be common across the sample from the original study and the two samples in the current study. Whereas in the case of team climate variables, a large size positive effect was commonly observed between all the variables. Such uncovered common relationships would have implications for both research and practice. The replication of the effect across samples pronounces more confidence in the original study's findings, and that contributes to the personality and team climate research in ASD. While it is highly likely to observe such an effect in other divisions of company A, the replication of the effect presents a strong case for formulating a hypothesis and testing further if the relationships would hold within the context of other telecom companies.

Under the industrial contexts where the relationship between personality variables is significant, managers can use that information while assembling a team. In cases where the elevation of a particular personality trait (see Section 4.1.3) is essential, the elevation of the corresponding positively correlated variable can be performed to indirectly influence the level of a particular trait. In the presence of a negative effect between personality variables, the elevation of a particular trait could be accomplished by demotion/reduction in the corresponding negatively correlated variable. Such changes in personality traits are feasible and even within a short timeframe (Allemand and Flückiger, 2017; Stieger et al., 2021).

When all the variables are positively correlated to each other, as in the case of team climate factors, a rise in the level of one of the factors would correspond to a boost in the levels of remaining factors and, ultimately, contribute towards an upsurge in an individual's perceived level of overall team climate. This would mean if a team manager could work towards improving one of the team climate factors, i.e., for example, if a manager could increase transparency and communication in a team by providing a safe forum for every team member to openly express their concerns, this would lead to a rise not only in the perceived level of participative safety within a team but also in the perceived level of rest of the team climate factors and indirectly contribute towards improving the overall perceived climate within the team.

Although some of the relationships identified via inter-correlations (in samples SE_2 and IN_1) within the personality and team climate variables were observed to be common with the original study, the analysis of correlations between personality and team climate variables did not result in any common relationship with the original study. Since each sample recruited in this study consisted of a whole new group of professionals working on different projects (different from the original study), where other factors like age, ethnicity, and working conditions of professionals also varied, not observing the same results despite recruiting one of the samples from the same division was not surprising. The fact that different results were observed across samples from the same organization, although it may suggest a threat to the external validity of the results, shows the importance of context. The samples SE_2 and IN_1, both of which were acquired during the context of pandemic thrust work from home, showed a significant negative correlation between neuroticism and team climate factors. As stress-related factors showed the most significant harm when working from home during the pandemic (Russo et al., 2021), we strongly believe the aforementioned negative correlation emerged due to a rise in neuroticism, which was a consequence of a contextual stress factor. The variations in the results of correlation analyses between current and original studies clearly imply the need for investigating different contexts.

4.3. Meta-analysis (RQ.3)

In the previous section, multiple effect sizes were observed across different samples. To better understand the magnitude of an effect and identify trends that can influence future research, we performed a meta-analysis. This analysis helps in aggregating the findings from multiple samples and in identifying statistically significant results. One of the contributions of this study is to aggregate the results from two replications with the ones from the original study by means of a meta-analysis of correlations.

To aggregate the findings from SE_1, SE_2, and IN_1, we conducted separate meta-analyses for each combination of personality and team climate variables across the three samples. The meta-analysis technique of correlations based on Z transformation (Acuña et al., 2015) was employed in our analysis and was implemented by R programming language.

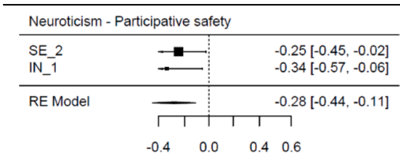
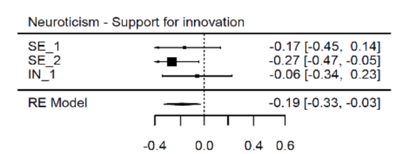
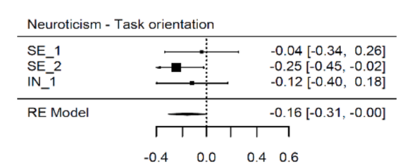
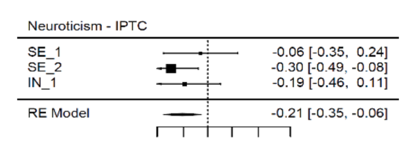
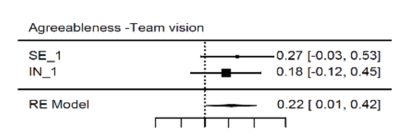
To estimate the meta-analytic effect size, in general, two models, namely the fixed-effects model and the random-effects model, are popular among SE studies. While the fixed-effects model assumes a fixed and unknown population effect size, the random-effects model assumes a stochastic and unknown population effect size distribution. In the fixed effects model, all the included studies are seen as drawn from the same population, and the variances in effect sizes between studies are associated with subject variability. Whereas, in the case of the random effects model, it accounts for situational variables and unknown factors that were not taken into consideration in the analysis. In this model, the variances in effect sizes are considered due to subject variability and because of inter-study variability. Based on the recommendations for meta-analyzing results of SE replications, for the meta-analyses in this study, we chose the random effects model over fixed effects procedures to account for heterogeneity in effect sizes beyond those produced by sampling error (A. Santos et al., 2021).

The forest plot that depicts the estimated common effect and gives a visual suggestion of the level of study heterogeneity was generated for every combination (except in two cases) of personality and team climate variables. Since the relationship between conscientiousness - participative safety and agreeableness - participative safety were studied in a single sample, a meta-analysis could not be performed. Due to space limitations, in this section, we present only the cases where a personality variable was observed to be a direct determinant of a team climate factor, i.e., we discuss only the cases where the pooled effect of the meta-analysis was significant. The details of such cases are presented in Table 8, and the information for the rest of the cases is presented in Appendix A.

In each forest plot presented in Table 8, the left column presents the list of labels for study samples whose individual effects were aggregated using a meta-analysis. The right column of the forest plot presents the observed effect size corresponding to each sample and its 95 % confidence interval. The box in the middle of each horizontal line (confidence interval) represents the point estimate of the effect for a single sample. The size of the box is proportional to the weight of the study in relation to the pooled estimate. The lozenge represents the overall effect estimate of the meta-analysis under the random effects model (RE model). The overall effect estimate, along with its p-value, are further listed under the 'Estimate' column of Table 8. The lozenge's width represents the 95 % confidence interval around the point estimate of the pooled effect. Further, the last column of Table 8 presents the Q statistic together with its p-value and the I^2 statistic, which are all indicators of heterogeneity. While a significant Q rejects the null hypothesis of homogeneity, a value of zero for the I^2 statistic indicates no observed heterogeneity.

The meta-analysis results from Table 8 suggested that there was a significant medium-size negative effect ($r = 0.292$ and $p < 0.01$) between the neuroticism and participative safety variables. It can be observed from Table 7 that although the effect was studied in only two cases (SE_2 and IN_1), both cases displayed significant negative correlations for this relationship. Further, the Q statistic was not significant and does not suggest that SE_2 and IN_1 are heterogeneous. The surveys in SE_2 and IN_1 were conducted in different countries where the sample sizes were different, and the ages and the roles of the subjects also varied. This statistical significance of the combined effect may now be

Table 8
Results from meta-analysis where significant correlations were observed.

Forest plot	Estimate		Heterogeneity		
	Correlation	p	Q	p	I^2
	−0.292**	0.0017	0.262	0.608	0.0 %
	−0.189*	0.0183	1.282	0.526	0.0 %
	−0.162*	0.0432	1.297	0.522	0.0 %
	−0.212**	0.008	1.624	0.443	0.0 %
	0.227*	0.038	0.186	0.665	0.0 %

p -value: '***' 0.001, '**'0.01 and '*' 0.05.

interpreted as evidence that, within a telecom company adhering to agile practices, an increase in neuroticism significantly decreases the perceived level of participative safety. The pooled negative effect confirms a commonsense dictate during teamwork; that is, team members who are anxious and worry about things have a lower level of perception regarding information sharing and acceptance within their team.

Besides participative safety, significant meta-analytic effects were observed between neuroticism and most of the team climate variables. While a significant small-size negative effect was seen with task orientation ($r = 0.162$ and $p < 0.05$), a similar pooled effect was detected in relation to support for innovation ($r = 0.189$ and $p < 0.05$). In both cases, the non-significant Q statistic and the I^2 statistic of zero inform us that there was no heterogeneity among studies. The meta-analytic effects inform us that a rise in the neuroticism level of a person corresponds to a slight decline in the person's perception of team characteristics like task orientation and support for innovation.

A relatively high (but still low) negative effect was observed between neuroticism and average team climate (IPTC) variables ($r = 0.212$ and $p < 0.01$). The Q statistic in this case was non-significant, and the I^2 statistic of zero informs us that SE_1, SE_2, and IN_1 are not heterogeneous. Although the negative relationship was noted to be significant only in SE_2, this effect was significant at a relatively higher level ($p < 0.01$), which informs us that compared to the former two small-size effects, there is relatively a higher probability in having this effect within a telecom company adhering to agile practices. The pooled negative effect presents evidence that within a telecom company like company A adopting agile practices for developing software, a rise in a person's neuroticism levels corresponds to a decline in the person's perception of the overall climate within the team.

While the effects associated with the neuroticism variable were all negative, indicating detrimental effects, we identified a significantly small size positive effect ($r = 0.227$ and $p < 0.05$) between agreeableness and team vision variables. Although a significant relationship was not seen in individual analyses associated with SE_1 IN_1, aggregation of the effects using meta-analysis led to the identification of a significant small-size positive effect. The Q and I^2 statistics do not provide any evidence for heterogeneity. Team vision relates to whether everyone within a team is clear and knows what they are doing and if they are on the right track to succeed. The meta-analysis of agreeableness and team vision variables manifests that a team member's ability to get along well with others corresponds to a better perception of team vision.

Although SE_1 picked significant positive correlations between agreeableness - IPTC variables and openness to experience – support for innovation variables, the results of our meta-analysis based on observations from three samples suggest that agreeableness and openness to experience traits are respectively not direct determinants of overall team climate (IPTC) and support for innovation dimensions. However, more replications would be necessary to establish confidence in this observation.

Takeaway points:

- The results from the meta-analysis suggest a medium-sized negative effect between the neuroticism and participative safety variables, where the effect was significant at 0.01 level.
- The meta-analysis over three samples further demonstrates neuroticism trait to have a relatively high (but still low) negative effect over the average team climate variable (IPTC) and a small size negative effect with task orientation and support for innovation variables.
- Despite the absence of a significant relationship when induvial samples (SE_1 and IN_1) were analyzed, meta-analysis identified a significant small-size positive effect between agreeableness and team vision variables.

4.3.1. *Observations and implications*

The meta-analysis executed in this study identified a significant small-size positive effect between agreeableness and team vision variables, although no significant relationship was observed in analyses of respective samples. On the other hand, despite noticing significant correlations between agreeableness - IPTC variables and openness to experience – support for innovation variables in the original study, our meta-analysis could not find any pooled effect in those cases. While the presence of a pooled effect shows evidence for a potential relationship between variables, the relationships identified in this study cannot yet be confidently and fully generalized to other organizations as they are still preliminary and are based on analyzing only three samples from a single company. It is also important to note that the absence of a significant effect from our meta-analysis cannot entirely rule out the possibility for a potential relationship among the variables. So, before applying our findings to practice and for further research in this direction, it is crucial to replicate the original study under different industrial contexts and aggregate all prior studies to establish confidence in our observations.

4.4. *Regression analysis (RQ.3)*

We performed regression analysis in extension to the correlation analysis for investigating whether the values of personality traits (Independent Variables IVs) can predict the values of team climate dimensions (Dependent Variables DVs). Since the results of the correlation analysis (see Table 7) show that five dependent variables (team vision, task orientation, support for innovation, participative safety, and IPTC) were significantly associated with a corresponding single independent variable (neuroticism), like the original study (Vishnubhotla et al., 2020), a linear regression model was used to further investigate the nature of the relationship between each pair of these independent and dependent variables. Therefore, based on the six significant correlations observed in Table 7, we developed six linear regression models to study the relationship between various IVs and DVs.

Prior to building the regression models, in each case, we tested whether assumptions to justify the use of linear regression were met. The assumptions relate to the normal distribution of the residuals, independence of the errors, and homoscedasticity of the errors (Usman et al., 2018). The Shapiro-Wilk test was used to validate the assumption related to the normal distribution of residuals, and the Durbin-Watson test was used to check the independence of residuals. Whereas the Breusch-Pagan and Koenker tests were used to test the constant variance (Vishnubhotla et al., 2020). The summary of the models, together with the observations from the tests, are presented next.

The summaries of linear regression models in relation to the six significant correlations observed from Section 4.2.3 are presented in Table 9. Under each model, estimates are presented for intercept and slope terms of the linear regression model. Further, the table also displays the coefficient of determination (R^2) together with the F-statistic.

In relation to the tests for validating regression assumptions, we observed that the null-hypothesis of normality for the Shapiro-Wilk test was not rejected in any of the cases (model1($p = 0.051$), model2($p = 0.042$), model3($p = 0.032$), model4($p = 0.092$), model5($p = 0.019$) and model6($p = 0.287$)) indicating that the residuals in each model were normally distributed. Next, the Durbin-Watson test for independence of residuals showed that the test statistics (model1(2.235), model2(2.279), model3(2.385), model4(2.258), model5(2.348) and model6(1.511)) were within the range of 1.5 to 2.5. Further, the results from Breusch-Pagan tests (model1(0.801), model2($p = 0.809$), model3($p = 0.818$), model4($p = 0.839$), model5($p = 0.971$) and model6($p = 0.142$)) and the Koenker tests (model1(0.820), model2($p = 0.849$), model3($p = 0.854$), model4($p = 0.854$), model5($p = 0.978$) and model6($p = 0.067$)) across all the models showed p-values above the significance level of 0.05 indicating that the null hypothesis for homoscedasticity was not rejected. Overall, the results from the tests indicate that linear regression assumptions were not violated with respect to any model.

Inspecting the summaries of regression models reported in Table 9,

Table 9
Summary of regression models for predicting team climate dimensions.

	Model 1: Team vision (SE_2)			Model 2: Task orientation (SE_2)			Model 3: Support for innovation (SE_2)			Model 4: Participative safety (SE_2)			Model 5: IPTC (SE_2)			Model 6: Participative safety (IN_1)		
	Est.	SE	p	Est.	SE	p	Est.	SE	p	Est.	SE	p	Est.	SE	p	Est.	SE	p
Constant	4.032	0.169	<0.001	4.047	0.149	<0.001	4.017	0.166	<0.001	4.284	0.142	<0.001	4.106	0.123	<0.001	4.487	0.254	<0.001
Neuroticism	−0.005	0.002	<0.05	−0.006	0.003	<0.05	−0.007	0.003	<0.05	−0.006	0.002	<0.05	−0.006	0.002	<0.01	−0.012	0.005	<0.05
R^2	0.054			0.063			0.067			0.071			0.091			0.116		
F	4.89			4.94			5.252			5.638			7.333			5.774		
p	0.021			0.029			0.024			0.020			0.008			0.0205		

we can see that in all the models, the intercept and slope values were significant at the 0.01 level. This informs us that, in each case, the addition of an IV to the respective model was significant. Moreover, the F-statistic in each case was greater than one indicating a real relationship between respective IV and DV. We can further observe that due to the detrimental effect of neuroticism, the DV in each model is negatively associated with the corresponding IV.

Upon comparing the coefficients of determination, we can see the value to be relatively high for model six. The R^2 value for model six was 0.116, indicating that for every one-point increase in the score of the neuroticism variable, the participative safety average score goes down by 0.116 points. Among the models associated with the SE_2 sample, the R^2 value was observed to be relatively high for model five, which predicts overall team climate, followed by the models to predict participative safety, support for innovation, task orientation, and team vision.

Between model six and model four, where the DV is participative safety, the coefficients of determination from Table 9 indicate that in model six, the neuroticism variable could better explain the variance found in participative safety scores. While 7.1 % of the variance in participative safety scores could be explained by the neuroticism variable from the SE_2 sample, in the case of sample IN_1, the neuroticism variable could explain 11.6 % of the variance in participative safety scores.

Although the estimates of intercept and slope in each model were significant, the smaller values of the coefficient of determination across all the models imply that neuroticism (IV) alone could not account for the majority of the variance in the DV. Results from our regression analysis showed that the IV from each model could explain only around 10 % (or less) of the variance in the DV. This suggests that the DV in each case is highly likely to be additionally explained by some other factors in addition to neuroticism.

In order to complement our results from regression analysis, a cross validation approach was adopted to finally investigate the models' prediction accuracy. We employed the five-fold cross validation approach as demonstrated by Mendes et al. (Mendes et al., 2021). In this approach, each sample is split into five-folds and ultimately used to compare the predicted model to the median model, as shown in Fig. 8 (Mendes et al., 2021).

The five-folds shown in Fig. 8 were used to build the median model and the predicted model. While the median model was created by filling a fold's median values in the place of predicted values, the predicted model was created by reserving one-fold for testing and using the remaining folds for training at each iteration. The training set was used for generating an equation, which is tested by the remaining fold.

The accuracy is measured with the help of three measures, such as the Mean Magnitude of Relative Error (MMRE), Mean Absolute Residual (MAR), and percentage of predictions within 25 % of error called Pred (25). The definitions and interpretations of these measures are presented in (Mendes et al., 2021) and (Shepperd and MacDonell, 2012). Further, a paired T-test with a significance level set to 0.05 was used to decide whether there were any differences between the predicted and median models. The accuracy of the six regression models has been assessed using the five-fold cross validation method, and the results are presented in Table 10.

With respect to the sample SE_2, The MMRE values across the five models in Table 10 ranged from 9.24 % to 26.06 %. The closer MMRE is to zero, the greater the accuracy. The MMRE value of up to 25 % indicates good accuracy (Shepperd and MacDonell, 2012). All the five models associated with sample SE_2 have MMRE values around 25 %, indicating good accuracy. The MAR statistic shows errors in absolute numbers that range from 0.328 to 0.762. With respect to MAR, smaller values indicate higher accuracy. Considering that the DV in the six models ranges from 1 to 5, an error between 0.3 to 0.7 units is small. The Pred(25) values show that at least 80 % of the prediction error is lower than 25 %. The closer the Pred(25) value is to 100 %, the greater the accuracy. Overall, from the statistics, we can understand that the

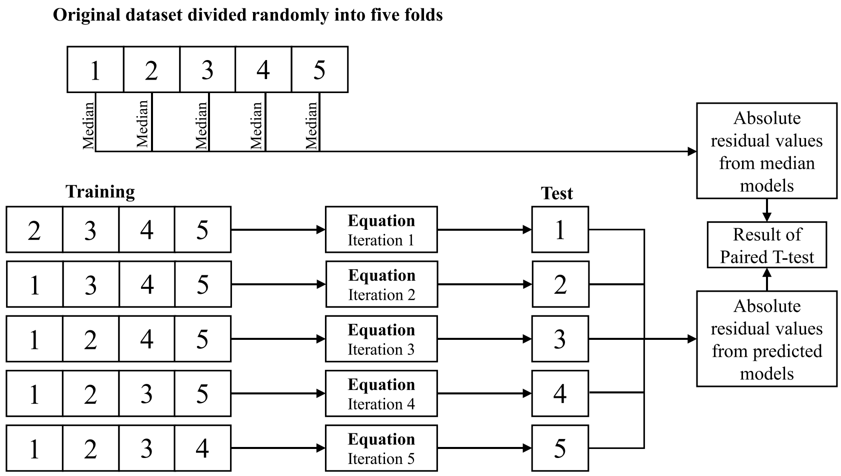


Fig. 8. The five-fold cross validation procedure.

accuracy is good for models one to six.

Whereas, in the case of model six, that is associated with sample IN_1, we observed that the MMRE values ranged from 11.49 % to 22.67 %. The MAR values ranged from 0.490 to 0.800, and the Pred(25) values show that at least 66.66 % of the prediction error is lower than 25 %. From the values of MMRE, MAR, and Pred(25), we believe that the accuracy of this model is substantial but relatively low in comparison to the other models.

It can be observed that for model five, where the overall team climate is predicted, the MMRE values range from 9.24 % to 12.22 %. The MAR values ranged from 0.328 to 0.473, and the Pred(25) values show that at least 86.66 % of the prediction error is lower than 25 %. This shows model five has greater accuracy compared to the rest.

In order to test the difference between the median model and the prediction model, we conducted a paired T-test to compare the absolute residual values from the median model and the same values produced by the predicted models. However, in relation to all six models we did not find a significant difference between the predictions obtained using the proposed model and the median model (model1: t(74)= 0.44 and $p = 0.61$, model2: t(74)= 0.47 and $p = 0.63$, model3: t(74)= 0.88 and $p = 0.38$, model4: t(74)= 0.28 and $p = 0.77$, model5: t(74)= 0.16 and $p = 0.86$ and model6: t(45)= 0.45 and $p = 0.65$).

Takeaway points:

- Regression analysis showed that the independent personality variable in five out of six models could account for only less than 10 % of the variance in the team climate variable (11.6 % in model six).
- Among the six regression models that were developed in this study, the accuracy was observed to be relatively high for the model predicting IPTC (model five).
- The comparison of the median model and regression-based model using paired T-test did not show any statistical difference between them, indicating predictions could be obtained by using either model.

4.4.1. *Observations and implications*

Since the independent personality variables in our regression analysis could account for 10 to 11 % variance in dependent team climate variables, it is important to note that the relationships detected in our study might or might not materialize in other agile teams associated with an industrial context that is similar to our study. In this regard, since there have not been many studies that emphasized on studying the relationship between personality and team climate factors, it is important for future research in this direction to replicate our original study and find out to what extent the relationships would be valid in other contexts. In specific, long-term inspection of personalities and team climate characteristics within one context can be a helpful strategy for acquiring more detailed knowledge necessary to establish robust prediction models. Finally, the relationships uncovered so far could serve as a starting point for formulating hypotheses that can be vividly tested in different industrial agile contexts.

In light of the relationships uncovered in this study, we recommend software organizations to maintain a personality and team climate repository of their employees and vividly incorporate those aspects into their decision-making process. Such a move would not only contribute to selecting personnel towards a better team climate and ultimately benefit their employees but also would provide inputs to organizations about offering training for their employees to improve the climate within teams ((Chad) Chiu et al., 2021).

5. Threats to validity

Research validity is a crucial aspect as it determines to what extent the results observed from our study pertain to the real world. This section discusses some of the threats that could affect the validity of our results and further outlines how we tried to mitigate some of them.

Evaluation apprehension and response bias: Evaluation apprehension relates to subjects trying to portray themselves as better when asked to share their opinions in relation to personality and team climate questions. Response bias relates to a subject's tendency to respond inaccurately. In our study, we mitigated these threats by explicitly informing the subjects that there were no correct or wrong answers to the questionnaire items. The subjects were further informed about how the data would be stored and managed and were notified that the raw data would not be shared with anyone from company A. Most importantly, the subjects were informed that the survey data would be anonymized and were clearly instructed about their right to opt-out of participating in our study.

Construct validity: This threat is associated with issues that could arise due to the improper design of the survey instrument. Although the reliability and validity of the IPIP-NEO and TCI questionnaires were rigorously evaluated by several studies (Vishnubhotla et al., 2020), since we transformed both questionnaires to be part of a Web survey, it is important to ensure that the survey instrument measures what it is intended to measure. So, we requested two professionals with experience of working in ASD teams to pilot our Web-based survey and assess the survey instrument. Their suggestions on the presentation and clarity of the instrument were duly addressed.

Internal validity: This threat relates to issues due to irrelevant respondents who could introduce a bias or systemic error in the study results. The following steps were taken to mitigate this threat: (1) The

Table 10
Accuracy of the regression models predicting team climate dimensions.

Case	Model	Fold	Equation		Accuracy Measures		
			Neuroticism	Constant	MMRE(%)	MAR	Pred(25)(%)
IV: Neuroticism	Predicted model	1	−0.005231	4.088320	10.82 %	0.352	93.33 %
DV: Team vision		2	−0.005402	3.975913	22.18 %	0.613	86.66 %
Sample: SE_2		3	−0.01027	4.08819	14.45 %	0.632	86.66 %
(Model 1)		4	−0.006015	4.136182	12.67 %	0.351	93.33 %
		5	−0.006314	4.038702	11.46 %	0.389	86.66 %
	Median model	1	–	–	15.32 %	0.415	86.66 %
		2	–	–	13.21 %	0.429	93.33 %
		3	–	–	23.91 %	0.648	80 %
		4	–	–	12.24 %	0.421	86.66 %
		5	–	–	11.55 %	0.503	93.33 %
IV: Neuroticism	Predicted model	1	−0.005403	4.018341	11.33 %	0.371	86.66 %
DV: Task orientation		2	−0.01037	4.16748	15.42 %	0.515	93.33 %
Sample: SE_2		3	−0.005539	3.988194	24.38 %	0.762	86.66 %
(Model 2)		4	−0.006125	4.048802	11.97 %	0.381	93.33 %
		5	−0.006664	4.020514	10.41 %	0.409	93.33
	Median model	1	–	–	13.27 %	0.428	86.66 %
		2	–	–	12.12 %	0.390	93.33 %
		3	–	–	26.69 %	0.742	60 %
		4	–	–	11.82 %	0.380	93.33 %
		5	–	–	10.85 %	0.428	86.66 %
IV: Neuroticism	Predicted model	1	−0.006793	4.008844	15.68 %	0.501	80 %
DV: Support for innovation		2	−0.009533	4.088420	20.42 %	0.628	86.66 %
Sample: SE_2		3	−0.006984	3.941783	13.61 %	0.501	86.66 %
(Model 3)		4	−0.007148	4.054913	26.06 %	0.590	80 %
		5	−0.008326	3.990937	13.52 %	0.528	86.66 %
	Median model	1	–	–	16.24 %	0.533	86.66 %
		2	–	–	19.98 %	0.583	80 %
		3	–	–	14.37 %	0.475	86.66 %
		4	–	–	24.40 %	0.575	86.66 %
		5	–	–	11.92 %	0.441	93.33 %
IV: Neuroticism	Predicted model	1	−0.003932	4.185321	17.45 %	0.552	86.66 %
DV: Participative safety		2	−0.009408	4.399264	10.80 %	0.382	93.33 %
Sample: SE_2		3	−0.006086	4.196182	12.45 %	0.545	100 %
(Model 4)		4	−0.005302	4.247200	10.60 %	0.369	86.66 %
		5	−0.009627	4.386344	14.31 %	0.542	93.33 %
	Median model	1	–	–	18.85 %	0.583	80 %
		2	–	–	9.29 %	0.338	93.33 %
		3	–	–	13.92 %	0.561	86.66 %
		4	–	–	11.96 %	0.416	86.66 %
		5	–	–	12.36 %	0.455	93.33 %
IV: Neuroticism	Predicted model	1	−0.005079	4.065923	12.15 %	0.370	86.66 %
DV: IPTC		2	−0.009406	4.204690	12.22 %	0.402	93.33 %
Sample: SE_2		3	−0.005826	4.029788	12.10 %	0.473	86.66 %
(Model 5)		4	−0.005797	4.103526	10.25 %	0.328	86.66 %
		5	−0.007955	4.127688	9.24 %	0.372	93.33 %
	Median model	1	–	–	13.70 %	0.419	93.33 %
		2	–	–	10.42 %	0.357	93.33 %
		3	–	–	12.84 %	0.473	86.66 %
		4	–	–	11.15 %	0.364	86.66 %
		5	–	–	8.79 %	0.313	100 %
IV: Neuroticism	Predicted model	1	−0.012299	4.544502	22.67 %	0.800	66.66 %
DV: Participative safety		2	−0.009903	4.381422	15.77 %	0.524	77.77 %
Sample: IN_1		3	−0.012015	4.489874	15.56 %	0.507	77.7 %
(Model 6)		4	−0.011800	4.428124	11.49 %	0.490	100 %
		5	−0.015308	4.600654	19.78 %	0.712	70 %
	Median model	1	–	–	18.73 %	0.740	55.55 %
		2	–	–	19.11 %	0.583	66.66 %
		3	–	–	14.44 %	0.462	77.77 %
		4	–	–	11.80 %	0.462	88.88 %
		5	–	–	20.05 %	0.658	70 %

survey homepage clearly mentioned that it was intended for gathering personality and team climate-specific information from people working in agile teams. Besides, respondents were asked to specify their role and team ID. The questions in relation to role and team ID were answered by all the respondents. (2) To minimize the evaluation apprehension among respondents, we informed them about anonymizing results and emphasized that there were no right or wrong answers to the questionnaire items.

External validity: One of the main problems related to the generalizability of the findings is the size of the sample or the number of respondents who took part in our study. Since the recruitment of subjects to our survey was handled by product owners from company A, our research team did not have an opportunity to identify more subjects. As this study investigates the case of one particular company, the statistical inferences that can be made from this study to a population are limited. Therefore, no inference statistics for comparing sample means and medians with regard to statistical significance were used.

Although correlation and regression analyses, tools of statistical inference, were used in this study, the interpretation of those analyses has to be done with great care as the results come from a single company, and no random sampling with regard to a population has been conducted. The main purpose of regression was to investigate the effect of personality characteristics of agile team members on their perception of team climate at company A. Companies with similar contexts might make similar observations, but an inference to the population of all telecom companies based on the regression would be misleading.

The results from our study can tentatively be generalized to agile teams working under a similar context. However, to generalize the findings for wider contexts, our research would need to be expanded to other companies in the telecom domain and also, at a later stage, to companies outside such domain.

It can be noticed that the majority of subjects from the SE_2 sample indicated as belonging to Sweden and all the subjects from IN_2 indicated as belonging to India. Since the majority of subjects within each sample belong to a specific country, this might pose a threat to the generalizability of our results. Recruiting a large and diverse sample consisting of people with different ages, ethnicities, and work cultures could help overcome this issue.

With respect to the correlation and regression analysis performed in our study, the neuroticism variable was observed to have small and medium effect sizes on some of the team climate variables. Although the effects were significant, the small and medium effect sizes mean that the associations observed in our study might/might not occur in some agile teams under similar contexts. Thus, the results of this study should be treated with "a grain of salt."

Conclusion validity: This threat is concerned with the correctness of conclusions regarding relationships in the data analyzed (Trochim and Donnelly, 2006). The threats from this category related to our study are the low reliability of measures due to noise and low statistical power. In order to mitigate the first threat, we organized seminars with product owners after analyzing responses from each sample to verify the consistency of the gathered data. The two samples recruited in our study only had 75 and 46 subjects, respectively. While the size of one of the samples (IN_1) was similar to that of the original study(SE_1) (Vishnubhotla et al., 2020), the second sample was relatively large (SE_2). A post-hoc power analysis over the SE_2 sample (employing the G*power tool) using our set alpha (0.05), estimated effect size, and actual sample size resulted in a post-hoc power of 0.75 and was observed to be within an acceptable range (0.70 to 0.90) (Maier and Lakens, 2022). Whereas in the case of the IN_1 sample, the relatively small sample size, combined with an alpha at 0.05 and a medium effect size, led to a post-hoc power of 0.54. We do acknowledge that such small power is far from ideal and should therefore be interpreted with caution; however, the issue of small sample size does not apply solely to our study (e.g., (Ampatzoglou et al., 2020)). This is a common issue in SE, as many industrial settings do not have large sample participants to collaborate in joint research projects.

Confirmability: This threat relates to the degree to which the results could be confirmed or corroborated by others. We mitigated this threat by organizing a seminar with professionals from our partnering company who are responsible for managing the recruited teams (product owners). In this seminar, we discussed the results for validation and collected their feedback. Overall, our study received a positive response.

6. Discussion

Replications are known to play a key role in empirical SE as they facilitate improving confidence and assessing the reliability of results (Shepperd et al., 2018). By replicating the original survey, we put together data about how the personality traits and perception of team climate factors vary across different divisions of a company. Each such replication will aid in incrementally pooling data by considering the practical issues associated with different organizational cultures and can bring several findings to team climate research within ASD. Further, the combination of data from different countries and contexts can shed some light on the possible differences or similarities regarding the influence of personality traits on agile team climate.

Upon comparing the measures of central tendency for the scores of personality traits, we can observe that in the sample SE_2, the level of agreeableness was relatively higher, followed by the level of conscientiousness and other traits. Whereas, in the IN_2 sample, the level of conscientiousness was relatively higher, followed by agreeableness and other traits.

Agreeableness has been recognized as an important factor for team progress (Balijepally et al., 2006) and team performance (R. Baumgart, M. Hummel, Personality traits of scrum roles in agile software development teams - a qualitative analysis, (n.d.)). Lower level of agreeableness in teams was termed detrimental to team performance as the presence of even one disagreeable member can disrupt the cooperation within a team. In agile teams, the presence of agreeable team members who are not very different from each other on agreeableness levels is desirable (Balijepally et al., 2006). The situation in both SE_2 and IN_1 was in line with this, as we can see from Fig. 2 and Fig. 3 that more than 75 % of the respondents had at least an average level of agreeableness.

On the other hand, conscientiousness reportedly influences individual job performance (Balijepally et al., 2006), contributes towards work satisfaction and promotes perseverance towards team goal completion (Neuman and Wright, 1999). Low levels of conscientiousness within a team are expected to lower team performance as members high on this trait may detest loafing by low conscientious members (Balijepally et al., 2006). This was clearly not the case in both SE_2 and IN_1 samples. While the conscientiousness levels were at least on an average level for more than 75 % of the subjects in SE_2, in IN_1, the conscientiousness levels were high for more than 50 % of the subjects. Conscientiousness is a positive personal trait for any work situation, including agile teams (Balijepally et al., 2006).

However, in our analysis results showed that the level of neuroticism was high in both SE_2 and IN_1 samples when compared to the levels observed in our original study (Vishnubhotla et al., 2020). While low levels of neuroticism are related to improved coordination and stability within a team (Neuman et al., 1999), task cohesion, and team performance (Barrick et al., 1998). High levels of neuroticism may not be conducive to teamwork as an emotionally unstable member could impair team performance by affecting cohesion and cooperation among members (Balijepally et al., 2006). Although Fig. 2 and Fig. 3 shows that the median for neuroticism scores fall under average levels with respect to reference personality data from respective countries, the fact that the median levels were high in SE_2 and IN_1 compared to the case in original study gives a hint that the stress levels of subjects were relatively high. The rise in stress levels could be partly attributed to situations such as lockdowns and working from home due to the COVID-19 pandemic during the time of execution of the survey. A study on understanding software developers' well-being and productivity during the pandemic times observed stress as a detrimental factor to respondents' well-being and mentioned that stress-related factors showed the most significant harm when working from home during the pandemic (Russo et al., 2021).

In relation to the team climate scores observed in SE_2 and IN_1 samples, the scores ranged from 3.5 to 5 (see Fig. 4 and Fig. 5), indicating the perceived team climate among all the teams to vary between positive to highly positive. Comparison of scores between the original study and the current study revealed that among the four team climate traits, the median levels of participative safety trait were relatively high across all the samples. The higher levels of participative safety indicate that subjects display higher tendencies for avoiding cognitive conflicts that could arise during discussions (Reiter-Palmon et al., 2012). The higher levels of participative safety also inform us about the subjects' state of feeling safe to share opinions within their teams, and such an environment could persuade members to be more expressive, which consequently leads to proposing novel ideas (Acuña et al., 2015). Since TCI is observed to be sensitive to differences in climate within teams from different contexts (Sumner and Molka-Danielsen, 2010), like in our original study (Vishnubhotla et al., 2020), the team climate results seen in the current study would be particularly relevant for organizations associated with the telecom domain that adopt scrum methodology for software development.

From examining the inter-correlations between the scores of personality traits, we identified the medium-sized positive effect between openness to experience and extraversion scores to be common across all three samples, where the correlations were significant at the level of 0.01. Further, from the inter-correlations between the scores of team climate dimensions, we observed that all the effect sizes observed in the original study, including the large size effect in relation to task orientation and support for innovation scores, were replicated in both samples. This informs us that such highly significant associations are very likely to be observed among agile teams from other organizations.

Unlike the original study, the correlation analysis from the current study identified significant correlations between neuroticism and multiple team climate dimensions. However, the results of the replications did not reproduce the results of the original survey with respect to the correlations between personality traits and team climate factors. So, in this context, it is worth noting that the findings from a baseline or original study shall not be regarded as a result that needs to be reproduced but shall be viewed only as a small piece of evidence within a larger picture, which eventually emerges after assembling many small pieces to complete the puzzle (A. Santos et al., 2021). Moreover, apart from the case of reproducing results of a previous study, where one might hope to find complete agreement, researchers, in general do not expect to find identical results between replications (Shepperd et al., 2018). Nevertheless, the variations in the design of the two studies and the differences in the sample of professionals allow us to conclude that the two results are compatible, increasing our confidence in the existence of the correlations.

Among the significant correlations, we noticed that the correlation between neuroticism and participative safety was significant in both SE_2 and IN_1 samples. In both cases, a medium size negative effect was observed. On the other hand, in the SE_2 sample, a significant medium-size negative effect was identified between neuroticism and overall perception of team climate (IPTC). The negative correlation in the aforementioned cases seems coherent because neuroticism is the tendency to have negative emotional experiences such as sadness, anxiety, and depression (Balijepally et al., 2006). Neurotic people tend to show emotionally unstable behavior and are reactive and susceptible to stress. So, neurotic members are typically less enthusiastic and not so confident about sharing ideas. Consequently, they tend to feel that their ideas are not likely to be received or acknowledged well within their team and, thereby, cannot perceive their team as a safe platform for generating ideas. These feelings, in turn, could adversely affect their perception of the overall climate within their team.

The aggregation of results from the current replications and original study via meta-analysis led to identifying five cases where a significant association between personality traits and team climate factors was observed. Although four out of five effects from the resultant meta-analytic estimates were of small size, it is worth noting that we did not observe a positive and negative effect size in identical replications (see forest plots from Table 8). This gives us a good idea of the direction of the relationship between the variables.

Further, our meta-analysis could not find any pooled effect in the case of agreeableness - IPTC variables and openness to experience – support for innovation variables, among which a significant correlation was observed in our original study. It is important to bear in mind that the absence of a significant effect from our meta-analysis cannot entirely rule out the possibility for a potential relationship among the variables. The findings from our meta-analysis should only be seen as a partial view of the results within the population. It would be possible to visualize the bigger picture only after collecting large chunks (i.e., running replications with large samples) or assembling many small pieces (like replications in this study) of evidence and treating them with appropriate methods (like meta-analysis). From this perspective, every tiny replication is key in SE, provided the replication is of comparable quality to the original study.

The meta-analysis in this study aggregated data from a set of three samples stemming from two countries. We would like to emphasize that the relationships and effects observed in our study may or may not be attained in other countries. This is because, in any country, the software industry needs are influenced by a set of context variables such as the maturity level of the software industry, the type of software products developed in a particular region, socio-economic conditions, etc. (Garousi et al., 2019). It is, therefore, crucial to replicate the original study under different conditions and aggregate all prior studies to define a more precise estimate of the effect size in question. Consequently, a small meta-analysis, as reported in this study, could perhaps serve as an initial step in the continuously cumulating meta-analysis technique (Braver et al., 2014).

Results from the regression analyses showed that in each of the six models that were constructed to predict team climate variables, only a single personality variable turned out to be a statistically significant independent variable. Results from our regression analysis showed that the independent personality variable in five out of six models could account for only less than 10 % of the variance in the team climate variable (11.6 % in the sixth model). However, as people are simply hard to predict (Agrawal and Agrawal, 2017), in many psychology studies, the coefficient of determination is reportedly less than 50 %. In such cases, the studies reporting even a small value of the coefficient of determination could have potential implications (Vishnubhotla et al., 2020).

On the other hand, the cross-validation procedure employed in relation to the regression models showed satisfactory prediction accuracy in all the cases (see Table 10). The absolute error (MAR) in most of the models ranged from 0.3 to 0.7. Considering the DMS scores ranging from 1 to 5, an error between 0.3 and 0.7 can be regarded as small. Among the six regression models that were developed in this study, the accuracy was observed to be relatively high for the model predicting IPTC, i.e., the overall team climate (MMRE: 9.24 % - 12.22 %, MAR: 0.32 to 0.47 and Pred(25) indicated that 86.6 % of the prediction error is below 25 %).

The comparison of the median model and regression-based model using paired T-test did not show any statistical difference between them. This means that the predictions could be obtained by either using the regression model or simply the median model. This observation warrants for further investigation in this regard, and it could be facilitated by means of recruiting a large and diverse sample.

Despite the hardest efforts to conduct exact replications, there is still a chance of encountering conflicting results (A. Santos et al., 2021). Under such circumstances, information such as first-hand knowledge of study execution and participant characteristics play a central role. In the event of the availability of such information, it is relatively easy to hypothesize on the variables that may be behind the divergent results. This could also aid in hypothesizing moderators that could be influencing the results.

In light of the small values of the coefficient of determination observed from the regression analyses performed in the original study and the current study, we firmly believe the inclusion of other independent variables could possibly improve the explanatory power of regression models and aid in improving their accuracy. Since role allocation, which is dependent on the capabilities of an individual, is a key factor in software team climate composition (Soomro et al., 2016), perhaps a relevant direction for future research would be to include capability measures as additional variables to predict agile team climate.

Given factors like our meta-analysis including data from only three samples, performing replications within divisions of a single telecom company, and other validity threats listed in Section 5, the results from this study cannot yet be confidently and fully generalized to other organizations. So, before adopting our results, it would be worthwhile to inspect whether and to what extent the relationships uncovered in our study apply to one's organization. This could, in turn, contribute to improving the external validity of our study.

6.1. Validation of results

To discuss our findings and receive feedback from company A, like the original study, we organized seminars with product owners after analyzing responses from each sample. The seminars were used to explain to the managers how the analysis was conducted. Further, the characteristics associated with each of the personality traits and team climate factors were shared, and they were informed how to interpret each of the variables in the study.

The results from the current study were presented together with the findings from previous iteration(s). The managers scrutinized the results, and they further brainstormed with us about the possible reasons for variations in observed results among teams and compared them with observations from the previous iteration(s). Thus, the results were vetted by the people in charge of managing teams, and we are confident about the validity of the findings from both the samples. Overall, our study received positive feedback from company A, and our findings were further disseminated to a wider audience.

7. Conclusions

This study emphasizes on exploring the relationship between personality traits and agile team climate within the context of a telecom company. By performing two iterations of replication at geographically distant divisions of a telecom company, we surveyed members from 19 teams (12 teams in the first iteration, followed by seven teams). By means of a correlation analysis and meta-analysis, the survey data was used towards uncovering associations between personality traits and team climate factors. The data was further used in a regression analysis for studying which personality traits serve as significant predictors of team climate factors.

The correlation analysis in this study did not replicate the significant relationships observed in the original study. However, the analysis unveiled new significant relationships between neuroticism and team climate factors (team vision (r = 0.24), task orientation (r = 0.25), support for innovation ($r = 0.27$) and participative safety ($r = 0.25$ and $r = 0.34$ in different samples)). All the correlations observed in relation to neuroticism were negative and significant at 0.05 level.

Our meta-analysis of correlations identified a significant medium-size negative effect ($r = 0.292$ and $p < 0.01$) between neuroticism and participative safety variables. Further, a significant small-size positive effect ($r = 0.227$ and $p < 0.05$) was observed between agreeableness and team vision variables, which is a relationship that was not observed to be significant among the individual samples we gathered.

The main findings from the two replications detailed in the paper, based upon correlations and meta-analysis, showed significant and negative relationships between neuroticism and team climate factors. Neuroticism is a personality trait characterized by emotionally unstable behaviour and by being more anxious and less calm. Such feelings can adversely affect one's perceived team climate. One of the critical outcomes from such findings is the need for organizations to support employees, and specifically teams, in ways that can increase their mutual sharing and trust (climate) and, in this way, mitigate possible issues relating to individuals who present higher levels of neuroticism. It is also known in the literature that neuroticism is significantly and negatively associated with personality traits like conscientiousness; thus, under circumstances where higher levels of neuroticism are observed in a Scrum team, we recommend managers who are in charge of assembling the Scrum team to counteract neuroticism levels by undertaking elevation of conscientiousness levels. We further recommend that organizations provide support/training so conscientiousness is elevated/increased throughout. Elevation in conscientiousness is expected to promote perseverance towards a team's goal completion and commitment to tasks. This can, in turn, indirectly mitigate possible issues relating to individuals who present higher levels of neuroticism.

The independent personality variable within each of our regression models, built in relation to the significant correlations observed from individual samples, could explain only around 10 % (or less) of variance in team climate factors. The accuracy of the regression models was observed to be good. However, a paired T-test indicated that there was no advantage of using a predicted model over the corresponding median model. The prediction models developed in our study could perhaps be improved by taking into account additional variables that could fit as significant predictors of team climate and by recruiting a larger and more diverse sample of agile practitioners.

CRediT authorship contribution statement

Sai Datta Vishnubhotla: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Writing – original draft. **Emilia Mendes:** Supervision, Validation, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

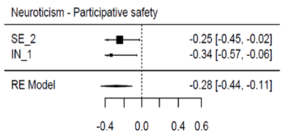
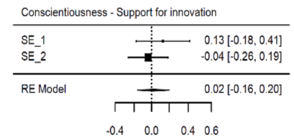
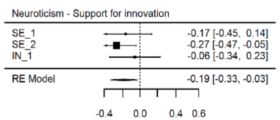
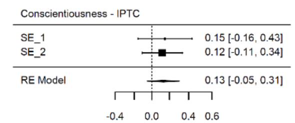
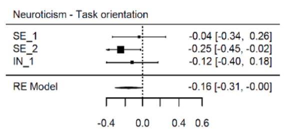
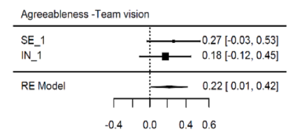
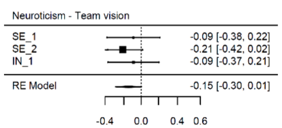
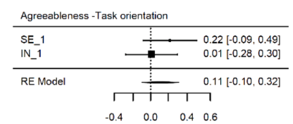
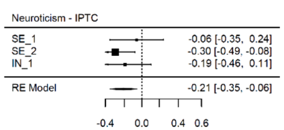
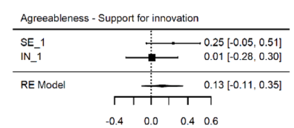
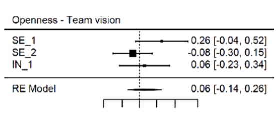
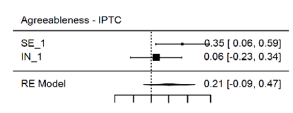
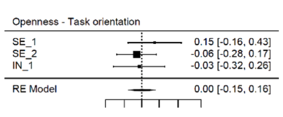
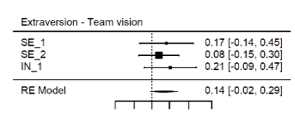
The authors do not have permission to share data.

Acknowledgment

We want to thank all those who participated in the surveys that were detailed in this study. This research was partially funded by the KK-stiftelsen-HÖG AGILE research project.

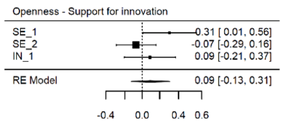
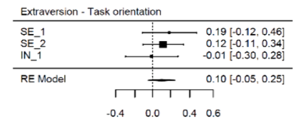
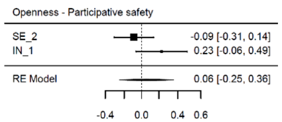
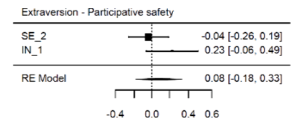
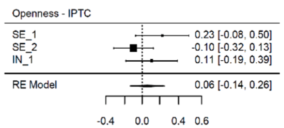
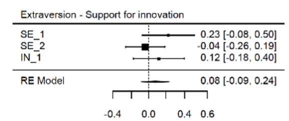
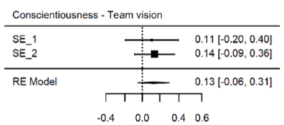
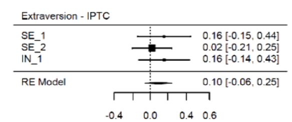
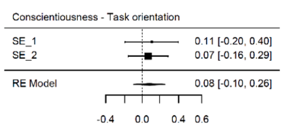
Appendix A

Results of meta-analysis, where p-value: '**'0.01 and '*' 0.05.

Forest plot	Estimate		Heterogeneity			Forest plot	Estimate		Heterogeneity		
	Corr	p	Q	p	I^2		Corr	p	Q	p	I^2
	−0.292**	0.0017	0.262	0.608	0.0 %		0.021	0.824	0.749	0.386	0.0 %
	−0.189*	0.0183	1.282	0.526	0.0 %		0.131	0.164	0.024	0.876	0.0 %
	−0.162*	0.0432	1.297	0.522	0.0 %		0.227*	0.038	0.186	0.665	0.0 %
	−0.147	0.066	0.582	0.747	0.0 %		0.113	0.303	0.946	0.330	0.0 %
	−0.212**	0.008	1.624	0.443	0.0 %		0.129	0.292	1.248	0.263	19.88 %
	0.061	0.548	3.091	0.213	36.09 %		0.209	0.169	1.932	0.164	48.25 %
	0.002	0.972	1.211	0.545	0.0 %		0.140	0.079	0.528	0.768	0.0 %

(*continued on next page*)

(*continued*)

Forest plot	Estimate		Heterogeneity			Forest plot	Estimate		Heterogeneity		
	Corr	p	Q	p	I^2		Corr	p	Q	p	I^2
	0.095	0.407	3.937	0.139	49.26 %		0.102	0.200	0.890	0.640	0.0 %
	0.057	0.721	2.833	0.092	64.71 %		0.080	0.556	2.024	0.154	50.60 %
	0.059	0.564	3.136	0.208	37.92 %		0.078	0.356	2.055	0.357	10.26 %
	0.130	0.168	0.023	0.877	0.0 %		0.095	0.233	0.770	0.680	0.0 %
	0.084	0.371	0.041	0.837	0.0 %						

Appendix B

An excerpt from the findings of our original study's correlation analyses (Vishnubhotla et al., 2020), where sample SE_1 was analyzed, is presented in this appendix. The Pearson correlation matrix for personality traits is presented in Table A. Next, the Pearson correlation matrix for team climate factors is presented in Table B. Note: In Table B, except for the scores of participative safety, the scores for the other three team climate variables were observed to be normally distributed. This is why the matrix in Table B presents correlations among three variables. Finally, the Pearson correlation matrix for correlations between personality traits and team climate factors is presented in Table C.

Table A
Pearson correlation matrix for personality traits (Sample: SE_1).

	Extraversion	Agreeableness	Conscientiousness	Neuroticism
Agreeableness	0.15			
Conscientiousness	0.43**	0.39**		
Neuroticism	−0.52***	−0.23	−0.70****	
Openness	0.40**	0.41**	0.10	−0.20

****$p < 0.0001$.
***$p < 0.001$ and.
**$p < 0.01$.

Table B
Pearson correlation matrix for team climate factors (Sample: SE_1).

	Team vision	Task orientation
Task orientation	0.47**	
Support for innovation	0.43**	0.62****

****$p < 0.0001$ and.
**$p < 0.01$.

Table C
Pearson correlation matrix for personality traits and team climate factors (Sample: SE_1).

	Team vision	Task orientation	Support for innovation	IPTC
Extraversion	0.17	0.19	0.23	0.16
Agreeableness	0.27	0.22	0.25	0.35*
Conscientiousness	0.11	0.11	0.13	0.15
Neuroticism	−0.09	−0.04	−0.17	−0.06
Openness	0.26	0.15	0.31*	0.23

*$p < 0.05$.

References

Acuña, S.T., Gómez, M., Juristo, N., 2008. Towards understanding the relationship between team climate and software quality a quasi-experimental study. Empir. Softw. Eng. 13, 401. https://doi.org/10.1007/s10664-008-9074-8.
Acuña, S.T., Gómez, M.N., Hannay, J.E., Juristo, N., Pfahl, D., 2015. Are team personality and climate related to satisfaction and software quality? aggregating results from a twice replicated experiment. Inf. Softw. Technol. 57, 141–156. https://doi.org/10.1016/j.infsof.2014.09.002.
Agile development at scale: the next frontier | IEEE journals & magazine | IEEE Xplore, (n.d.). https://ieeexplore.ieee.org/abstract/document/8648272 (accessed October 16, 2023).
Agrawal, V.K., Agrawal, V.K., Requirements of commercial-off-the-shelf software a comparison between manufacturing and service Sectors, in: 2017.
Allemand, M., Flückiger, C., 2017. Changing personality traits: some considerations from psychotherapy process-outcome research for intervention efforts on intentional personality change. J. Psychother. Integr. 27, 476–494. https://doi.org/10.1037/int0000094.
Ambler, S.W., 2008. Has agile peaked? DR Dobbs J. 33, 52–54.
Ampatzoglou, A., S. Bibi, P. Avgeriou, A. Chatzigeorgiou, Guidelines for Managing Threats to Validity of Secondary Studies in Software Engineering, in: M. Felderer, G. H. Travassos (Eds.), Contemporary Empirical Methods Software Engineering, Springer International Publishing, Cham, 2020: pp. 415–441. https://doi.org/10.1007/978-3-030-32489-6_15.
Anderson, N.R., West, M.A., 1998. Measuring climate for work group innovation: development and validation of the team climate inventory. J. Organ. Behav. 19, 235–258. https://doi.org/10.1002/(SICI)1099-1379(199805)19:3<235::AID-JOB837>3.0.CO;2-C.
Anderson, G., Keith, M.J., Francisco, J., Fox, S., The effect of software team personality composition on learning and performance: making the "dream" team, Hawaii International Conference On System Sciences (HICSS-51). (2018). https://aisel.aisnet.org/hicss-51/cl/social_and_psychological_perspectives/3.
Balijepally, V., Mahapatra, R., Nerur, S.P., 2006. Assessing personality profiles of software developers in agile development teams. Commun. Assoc. Inf. Syst. 18 https://doi.org/10.17705/1CAIS.01804.
Barrick, M.R., Mount, M.K., 1991. The big five personality dimensions and job performance: a meta-analysis. Pers. Psychol. 44, 1–26. https://doi.org/10.1111/j.1744-6570.1991.tb00688.x.
Barrick, M.R., Stewart, G.L., Neubert, M.J., Mount, M.K., 1998. Relating member ability and personality to work-team processes and team effectiveness. J. Appl. Psychol. 83, 377–391. https://doi.org/10.1037/0021-9010.83.3.377.
Barry, B., Stewart, G.L., 1997. Composition, process, and performance in self-managed groups: the role of personality. J. Appl. Psychol. 82, 62–78. https://doi.org/10.1037/0021-9010.82.1.62.
Berraies, S., Chouiref, A., 2022. Exploring the effect of team climate on knowledge management in teams through team work engagement: evidence from knowledge-intensive firms. J. Knowl. Manag. 27, 842–869. https://doi.org/10.1108/JKM-09-2021-0720.
Braver, S.L., Thoemmes, F.J., Rosenthal, R., 2014. Continuously cumulating meta-analysis and replicability. Perspect. Psychol. Sci. 9, 333–342. https://doi.org/10.1177/1745691614529796.
Calefato, F., Lanubile, F., 2022. Using Personality detection tools for software engineering research: how far can we go? ACM Trans. Softw. Eng. Methodol. 31 https://doi.org/10.1145/3491039, 42:1–42:48.
Capretz, L.F., Ahmed, F., da Silva, F.Q.B., 2017. Soft sides of software. Inf. Softw. Technol. 92, 92–94. https://doi.org/10.1016/j.infsof.2017.07.011.
Caulo, M., Francese, R., Scanniello, G., Tortora, G., 2021. Relationships between personality traits and productivity in a multi-platform development context. In: Proceedings of the 25th International Conference on Evaluation and Assessment in Software Engineering. Association for Computing Machinery, New York, NY, USA, pp. 70–79. https://doi.org/10.1145/3463274.3463327.
Chatzi, S., Nikolaou, I., Anderson, N., 2022. Team personality composition and team innovation implementation: the mediating role of team climate for innovation. Appl. Psychol. https://doi.org/10.1111/apps.12408 n/a.

(Chad) Chiu, C.Y., Lin, H.C., Ostroff, C., 2021. Fostering team learning orientation magnitude and strength: roles of transformational leadership, team personality heterogeneity, and behavioural integration. J. Occup. Organ. Psychol. 94, 187–216. https://doi.org/10.1111/joop.12333.

Coordination challenges in large-scale software development: a case study of planning misalignment in hybrid settings | IEEE journals & magazine | IEEE Xplore, (n.d.). https://ieeexplore.ieee.org/abstract/document/7990187?casa_token=vuTUAeudRn8AAAAA:q3M9O462rgk-E0PPDehL4kp-CK_l-2ggdnG6cfOPgCx0t-UkLUUacclRmWw6s0iZPxdl0HiYTzo (accessed October 16, 2023).

Costa Jr., P.T., McCrae, R.R., The revised NEO personality inventory (NEO-PI-R), in: SAGE Handb. Personal. Theory Assess. Vol 2 Personal. Meas. Test., Sage Publications, Inc, Thousand Oaks, CA, US, 2008: pp. 179–198. https://doi.org/10.4135/9781849200479.n9.

Costa, P.T., McCrae, R.R., Pi-R, Neo, 1992. Psychological Assessment Resources.

Cruz, S.S.J.O., da Silva, F.Q.B., Monteiro, C.V.F., Santos, P., Rossilei, I., dos Santos, M.T., Personality in software engineering: preliminary findings from a systematic literature review, in: 15th Annual Conference on Evaluation & Assessment in Software Engineering. EASE 2011, 2011: pp. 1–10. https://doi.org/10.1049/ic.2011.0001.

Cruz, S., da Silva, F.Q.B., Capretz, L.F., 2015. Forty years of research on personality in software engineering: a mapping study. Comput. Hum. Behav. 46, 94–113. https://doi.org/10.1016/j.chb.2014.12.008.

da Silva, F.Q.B., Suassuna, M., França, A.C.C., Grubb, A.M., Gouveia, T.B., Monteiro, C.V.F., dos Santos, I.E., 2014. Replication of empirical studies in software engineering research: a systematic mapping study. Empir. Softw. Eng. 19, 501–557. https://doi.org/10.1007/s10664-012-9227-7.

Digital.ai, 15th state of agile report, (2021). https://digital.ai/resource-center/analyst-reports/state-of-agile-report.

Domino, M.A., Collins, R.W., Hevner, A.R., Cohen, C.F., 2003. Conflict in collaborative software development. In: Proceedings of the 2003 SIGMIS conference on Computer personnel research: Freedom in Philadelphia leveraging differences and diversity in the IT workforce. Association for Computing Machinery, New York, NY, USA, pp. 44–51. https://doi.org/10.1145/761849.761856.

Donnellan, M.B., Oswald, F.L., Baird, B.M., Lucas, R.E., 2006. The Mini-IPIP scales: tiny-yet-effective measures of the big five factors of personality. Psychol. Assess. 18, 192–203. https://doi.org/10.1037/1040-3590.18.2.192.

Dutra, E., Santos, G., 2020. Organisational climate assessments of Agile teams a qualitative multiple case study. IET Softw. 14, 861–870. https://doi.org/10.1049/iet-sen.2020.0048.

Dutra, E., Lima, P., Santos, G., 2020. An Instrument to assess the organizational climate of agile teams a preliminary study. In: 19th Brazilian Symposium Software Quality. Association for Computing Machinery, New York, NY, USA, pp. 1–10. https://doi.org/10.1145/3439961.3439968.

Dybå, T., Sjøberg, D.I.K., Cruzes, D.S., 2012. What works for whom, where, when, and why? on the role of context in empirical software engineering. In: Proceedings of the 2012 ACM-IEEE International Symposium on Empirical Software Engineering and Measurement. Association for Computing Machinery, New York, NY, USA, pp. 19–28. https://doi.org/10.1145/2372251.2372256.

Fay, D., Lührmann, H., Kohl, C., 2004. Proactive climate in a post-reorganization setting: when staff compensate managers' weakness. Eur. J. Work Organ. Psychol. 13, 241–267. https://doi.org/10.1080/13594320444000083.

Feldt, R., Angelis, L., Torkar, R., Samuelsson, M., 2010. Links between the personalities, views and attitudes of software engineers. Inf. Softw. Technol. 52, 611–624. https://doi.org/10.1016/j.infsof.2010.01.001.

Fernando Capretz, L., 2014. Bringing the human factor to software engineering. IEEE Softw 31, 104. https://doi.org/10.1109/MS.2014.30.

Fowler, M., Highsmith, J., 2001. The agile manifesto. Softw. Dev. 9, 28–35.

Francese, R., Milione, V., Scanniello, G., Tortora, G., A preliminary investigation on the relationships between personality traits and team climate in a smart-working development context, in: L. Ardito, A. Jedlitschka, M. Morisio, M. Torchiano (Eds.), International Conference on Product-Focused Software Process Improvement, Springer International Publishing, Cham, 2021: pp. 167–182. https://doi.org/10.1007/978-3-030-91452-3_11.

Garousi, V., Giray, G., Tüzün, E., Catal, C., Felderer, M., 2019. Aligning software engineering education with industrial needs: a meta-analysis. J. Syst. Softw. 156, 65–83. https://doi.org/10.1016/j.jss.2019.06.044.

Gila, A.R., Jaafa, J., Omar, M., Tunio, M.Z., Impact of personality and gender diversity on software development teams' performance, in: 2014 International Conference on Computer, Communications, and Control Technology (I4CT), 2014: pp. 261–265. https://doi.org/10.1109/I4CT.2014.6914186.

Gilal, A.R., Jaafar, J., Capretz, L.F., Omar, M., Basri, S., Aziz, I.A., 2018. Finding an effective classification technique to develop a software team composition model. J. Softw. Evol. Process. 30, e1920. https://doi.org/10.1002/smr.1920.

Goldberg, L.R., Johnson, J.A., Eber, H.W., Hogan, R., Ashton, M.C., Cloninger, C.R., Gough, H.G., 2006. The international personality item pool and the future of public-domain personality measures. J. Res. Personal. 40, 84–96. https://doi.org/10.1016/j.jrp.2005.08.007.

Goldberg, L.R., A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models, in: 1999. https://www.semanticscholar.org/paper/A-broad-bandwidth%2C-public-domain%2C-personality-the-Goldberg/af3fabb831eeffbef6af39588c489cc5d838a404 (accessed October 10, 2023).

Hogan, R., Johnson, J.A., Briggs, S.R., 1997. Handbook of Personality Psychology. Elsevier. https://doi.org/10.1016/B978-0-12-134645-4.X5000-8.

Jørgensen, M., Dybå, T., Liestøl, K., Sjøberg, D.I.K., 2016. Incorrect results in software engineering experiments: how to improve research practices. J. Syst. Softw. 116, 133–145. https://doi.org/10.1016/j.jss.2015.03.065.

Johnson, J.A., IPIP-NEO narrative report, (n.d.). https://drj.virtualave.net/IPIP/ipipneo120.htm.

Jalote, P., Natarajan, P., 2019. The growth and evolution of India's software industry. Commun. ACM. 62, 64–69. https://doi.org/10.1145/3347863.

Johnson, J.A., 2014. Measuring thirty facets of the five factor model with a 120-item public domain inventory: development of the IPIP-NEO-120. J. Res. Personal. 51, 78–89. https://doi.org/10.1016/j.jrp.2014.05.003.

Johnson, J.A., Interpreting individual IPIP scale scores, (n.d.). https://ipip.ori.org/InterpretingIndividualIPIPScaleScores.htm (accessed August 7, 2022).

Kang, H.R., Yang, H.D., Rowley, C., 2006. Factors in team effectiveness: cognitive and demographic similarities of software development team members. Hum. Relat. 59, 1681–1710. https://doi.org/10.1177/0018726706072891.

Klebanov, B.B., Using people and WordNet to measure semantic relatedness, (2006).

Kollmann, T., Häsel, M., Breugst, N., 2009. Competence of IT professionals in e-business venture teams: the effect of experience and expertise on preference structure. J. Manag. Inf. Syst. 25, 51–79.

Kosti, M.V., Feldt, R., Angelis, L., 2014. Personality, emotional intelligence and work preferences in software engineering: an empirical study. Inf. Softw. Technol. 56, 973–990. https://doi.org/10.1016/j.infsof.2014.03.004.

Lee, K., Ashton, M.C., HEXACO model of personality structure, the, in: V. Zeigler-Hill, T. K. Shackelford (Eds.), Encycl. Personal. Individ. Differ., Springer International Publishing, Cham, 2020: pp. 1932–1936. https://doi.org/10.1007/978-3-319-24612-3_1227.

Lee, J.C., Chen, C.Y., 2020. Exploring the team dynamic learning process in software process tailoring performance: a theoretical perspective. J. Enterp. Inf. Manag. 33, 502–518. https://doi.org/10.1108/JEIM-07-2019-0202.

Lee, J.M., Shneiderman, B., 1978. Personality and programming: time-sharing vs. batch preference. In: Proceedings of the 1978 annual conference Vol. 2. Association for Computing Machinery, New York, NY, USA, pp. 561–569. https://doi.org/10.1145/800178.810092.

Lindsjørn, Y., Sjøberg, D.I.K., Dingsøyr, T., Bergersen, G.R., Dybå, T., 2016. Teamwork quality and project success in software development: a survey of agile development teams. J. Syst. Softw. 122, 274–286. https://doi.org/10.1016/j.jss.2016.09.028.

Maier, M., Lakens, D., 2022. Justify your alpha: a primer on two practical approaches. Adv. Methods Pract. Psychol. Sci. 5 https://doi.org/10.1177/25152459221080396, 25152459221080396.

Mathisen, G.E., Einarsen, S., Jørstad, K., Brønnick, K.S., 2004. Climate for work group creativity and innovation: norwegian validation of the team climate inventory (TCI). Scand. J. Psychol. 45, 383–392. https://doi.org/10.1111/j.1467-9450.2004.00420.x.

Matthews, G., Deary, I.J., Whiteman, M.C., 2003. Personality Traits. Cambridge University Press.

Matturro, G., Fontán, C., Raschetti, F., Soft skills in scrum teams. a survey of the most valued to have by product owners and scrum masters, in: 2015. https://doi.org/10.18293/SEKE2015-026.

Matzler, K., Renzl, B., Müller, J., Herting, S., Mooradian, T.A., 2008. Personality traits and knowledge sharing. J. Econ. Psychol. 29, 301–313. https://doi.org/10.1016/j.joep.2007.06.004.

McCrae, R.R., Costa Jr., P.T., 1989. Reinterpreting the Myers-Briggs type indicator from the perspective of the five-factor model of personality. J. Pers. 57, 17–40. https://doi.org/10.1111/j.1467-6494.1989.tb00759.x.

McCrae, R.R., Terracciano, A., 2005. Personality Profiles of Cultures Project, Universal features of personality traits from the observer's perspective: data from 50 cultures. J. Pers. Soc. Psychol. 88, 547–561. https://doi.org/10.1037/0022-3514.88.3.547.

Mendes, F.F., Mendes, E., Salleh, N., 2019. The relationship between personality and decision-making: a Systematic literature review. Inf. Softw. Technol. 111, 50–71. https://doi.org/10.1016/j.infsof.2019.03.010.

Mendes, F., Mendes, E., Salleh, N., Oivo, M., 2021. Insights on the relationship between decision-making style and personality in software engineering. Inf. Softw. Technol. 136, 106586 https://doi.org/10.1016/j.infsof.2021.106586.

Molleman, E., Nauta, A., Jehn, K.A., 2004. Person-job fit applied to teamwork: a multilevel approach. Small Group Res. 35, 515–539. https://doi.org/10.1177/1046496404264361.

Neuman, G.A., Wright, J., 1999. Team effectiveness: beyond skills and cognitive ability. J. Appl. Psychol. 84, 376–389. https://doi.org/10.1037/0021-9010.84.3.376.

Neuman, G.A., Wagner, S.H., Christiansen, N.D., 1999. The relationship between work-team personality composition and the job performance of teams. Group Organ. Manag. 24, 28–45. https://doi.org/10.1177/1059601199241003.

O'Neill, T.A., Allen, N.J., 2011. Personality and the prediction of team performance. Eur. J. Personal. 25, 31–42. https://doi.org/10.1002/per.769.

Penzenstadler, B., Torkar, R., Montes, C.Martinez, 2022. Take a deep breath: benefits of neuroplasticity practices for software developers and computer workers in a family of experiments. Empir. Softw. Eng. 27, 98. https://doi.org/10.1007/s10664-022-10148-z.

Qamar, N., Malik, A.A., 2020. Determining the relative importance of personality traits in influencing software quality and team productivity. Comput. Inform. 39, 994–1021. https://doi.org/10.31577/cai_2020_5_994.

Baumgart, R., Hummel, M., Personality traits of scrum roles in agile software development teams - a qualitative analysis, (n.d.). https://core.ac.uk/outputs/301366940 (accessed August 16, 2022).

Ragazzoni, P., Baiardi, P., Zotti, A.M., Anderson, N., West, M., 2002. Research note: italian validation of the team climate inventory: a measure of team climate for innovation. J. Manag. Psychol. 17, 325–336. https://doi.org/10.1108/02683940210428128.

Rahman, H.U., Raza, M., Afsar, P., Khan, H.U., 2021. Empirical investigation of influencing factors regarding offshore outsourcing decision of application

maintenance. IEEE Access 9, 58589–58608. https://doi.org/10.1109/ACCESS.2021.3073315.
Rahmani, C., Khazanchi, D., A study on defect density of open source software, in: IEEE/ACIS International Conference on Computer and Information Science, 2010: pp. 679–683. https://doi.org/10.1109/ICIS.2010.11.
Reiter-Palmon, R., Wigert, B., de Vreede, T., 2012. Chapter 13 team creativity and innovation: the effect of group composition, social processes, and cognition. In: Mumford, M.D. (Ed.), Handbook of Organizational Creativity. Academic Press, San Diego, pp. 295–326. https://doi.org/10.1016/B978-0-12-374714-3.00013-6.
Rodríguez, P., Markkula, J., Oivo, M., Turula, K., 2012. Survey on Agile and Lean Usage in Finnish Software Industry. In: Proceeding ACM-IEEE International Symposium on Empirical Software Engineering and Measurement. ACM, New York, NY, USA, pp. 139–148. https://doi.org/10.1145/2372251.2372275.
Rothstein, M.G., Goffin, R.D., 2006. The use of personality measures in personnel selection: what does current research support? Hum. Resour. Manag. Rev. 16, 155–180. https://doi.org/10.1016/j.hrmr.2006.03.004.
Russo, D., Hanel, P.H., van Berkel, N., Understanding developers well-being and productivity: a longitudinal analysis of the covid-19 pandemic, ArXiv Prepr. ArXiv211110349. (2021).
Salleh, N., Mendes, E., Grundy, J., 2014. Investigating the effects of personality traits on pair programming in a higher education setting through a family of experiments. Empir. Softw. Eng. 19, 714–752. https://doi.org/10.1007/s10664-012-9238-4.
Salman, I., Turhan, B., 2018. Effect of time-pressure on perceived and actual performance in functional software testing. In: Proceedings of the 2018 International Conference on Software and System Process. Association for Computing Machinery, New York, NY, USA, pp. 130–139. https://doi.org/10.1145/3202710.3203148.
Santos, A., Vegas, S., Oivo, M., Juristo, N., 2021a. Comparing the results of replications in software engineering. Empir. Softw. Eng. 26, 13. https://doi.org/10.1007/s10664-020-09907-7.
Santos, A., Vegas, S., Oivo, M., Juristo, N., 2021b. A Procedure and guidelines for analyzing groups of software engineering replications. IEEE Trans. Softw. Eng. 47, 1742–1763. https://doi.org/10.1109/TSE.2019.2935720.
Shepperd, M., MacDonell, S., 2012. Evaluating prediction systems in software project estimation. Inf. Softw. Technol. 54, 820–827. https://doi.org/10.1016/j.infsof.2011.12.008.
Shepperd, M., Ajienka, N., Counsell, S., 2018. The role and value of replication in empirical software engineering results. Inf. Softw. Technol. 99, 120–132. https://doi.org/10.1016/j.infsof.2018.01.006.
Silva, D.S., Rabelo, R., Neto, P.S., Britto, R., Oliveira, P.A., A test case prioritization approach based on software component metrics, in: IEEE International Conference on Systems, Man and Cybernetics, 2019: pp. 2939–2945. https://doi.org/10.1109/SMC.2019.8914670.
Software project failure process definition | IEEE journals & magazine | IEEE Xplore, (n.d.). https://ieeexplore.ieee.org/abstract/document/9743906 (accessed October 16, 2023).
Soomro, A.B., Salleh, N., Nordin, A., How personality traits are interrelated with team climate and team performance in software engineering? A preliminary study, in: 2015 9th Malaysian Conference in Software Engineering (MySEC), 2015: pp. 259–265. https://doi.org/10.1109/MySEC.2015.7475230.
Soomro, A.B., Salleh, N., Mendes, E., Grundy, J., Burch, G., Nordin, A., 2016. The effect of software engineers' personality traits on team climate and performance: a Systematic Literature Review. Inf. Softw. Technol. 73, 52–65. https://doi.org/10.1016/j.infsof.2016.01.006.
St J. Burch, G., Anderson, N., 2004. Measuring person-team fit: development and validation of the team selection inventory. J. Manag. Psychol. 19, 406–426. https://doi.org/10.1108/02683940410537954.
Stieger, M., Flückiger, C., Rüegger, D., Kowatsch, T., Roberts, B.W., Allemand, M., 2021. Changing personality traits with the help of a digital personality change intervention. Proc. Natl. Acad. Sci. 118, e2017548118 https://doi.org/10.1073/pnas.2017548118.
Sturdee, M., Ivory, M., Ellis, D., Stacey, P., Ralph, P., 2022. Personality Traits in game development. In: International Conference on Evaluation and Assessment in Software Engineering 2022. ACM, Gothenburg Sweden, pp. 221–230. https://doi.org/10.1145/3530019.3530042.
Sumner, M., Molka-Danielsen, J., 2010. Global IT teams and project success. In: Proceedings of the 2010 Special Interest Group on Management Information System's 48th annual conference on Computer personnel research on Computer personnel research. Association for Computing Machinery, New York, NY, USA, pp. 34–42. https://doi.org/10.1145/1796900.1796920.
Trendowicz, A., Münch, J., 2009. Factors influencing software development productivity - state-of-the-art and industrial experiences. Adv. Comput. 77, 185–241. https://doi.org/10.1016/S0065-2458(09)01206-6.
Trochim, W.M.K., Donnelly, J.P., 2006. The Research Methods Knowledge Base, 3rd edition. 3rd Edition.
Truong, D., Jitbaipoon, T., 2016. How Can agile methodologies be used to enhance the success of information technology projects? Int. J. Inf. Technol. Proj. Manag. 7, 1–16. https://doi.org/10.4018/IJITPM.2016040101.
Usman, M., Britto, R., Damm, L.O., Börstler, J., 2018. Effort estimation in large-scale software development: an industrial case study. Inf. Softw. Technol. 99, 21–40. https://doi.org/10.1016/j.infsof.2018.02.009.
Vishnubhotla, S.D., Mendes, E., Lundberg, L., 2018. An insight into the capabilities of professionals and teams in agile software development: a systematic literature review. In: Proceedings of the 2018 7th International Conference on Software and Computer Applications. Association for Computing Machinery, New York, NY, USA, pp. 10–19. https://doi.org/10.1145/3185089.3185096.
Vishnubhotla, S.D., Mendes, E., Lundberg, L., 2020. Investigating the relationship between personalities and agile team climate of software professionals in a telecom company. Inf. Softw. Technol. 126, 106335 https://doi.org/10.1016/j.infsof.2020.106335.
Weinberg, G.M., 1971. The Psychology of Computer Programming. Van Nostrand Reinhold, New York.
Xu, X., Jiang, L., Wang, H.J., 2019. How to build your team for innovation? A cross-level mediation model of team personality, team climate for innovation, creativity, and job crafting. J. Occup. Organ. Psychol. 92, 848–872. https://doi.org/10.1111/joop.12277.
Yilmaz, M., O'Connor, R.V., Colomo-Palacios, R., Clarke, P., 2017. An examination of personality traits and how they impact on software development teams. Inf. Softw. Technol. 86, 101–122. https://doi.org/10.1016/j.infsof.2017.01.005.
Zolduoarrati, E., Licorish, S.A., Stanger, N., 2023. Secondary studies on human aspects in software engineering: a tertiary study. J. Syst. Softw. 200, 111654 https://doi.org/10.1016/j.jss.2023.111654.

Sai Datta Vishnubhotla holds a Licentiate degree in Computer Science and a Master's degree in Electrical Engineering with emphasis on Telecommunication Systems. He is pursuing a Ph.D.at the Blekinge Institute of Technology (Sweden) and is now working in the industry-academia project AGILE. His-current research interests are human aspects in software engineering, machine learning, cloud computing and virtualization.

Emilia Mendes is a Professor in Computer Science at the Blekinge Institute of Technology (Sweden). The core of the research she is involved in is applied and falls within Computer Science & Empirical Software Engineering. More specifically, she has focused upon: i) the use of statistical and machine-learning techniques to several areas such as Software/Web cost estimation, value-based decision making, software maintenance; ii) measurement and metrics; iii) evidence-based research; iv) software process improvement; and v) personality and its relation to learning and also to productivity.