Effects of Personality Traits on Pull Request Acceptance

Rahul N. Iyer, S. Alex Yun, Meiyappan Nagappan, and Jesse Hoey

Abstract—In this paper, we examine the influence of personality traits of developers on the pull request evaluation process in GitHub. We first replicate Tsay *et al.*'s work that examined the influence of social factors (e.g., 'social distance') and technical factors (e.g., test file inclusion) for evaluating contributions, and then extend it with personality based factors. In particular, we extract the Big Five personality traits (Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism) of developers from their online digital footprints, such as pull request comments. We analyze the personality traits of 16,935 active developers from 1,860 projects and compare their relative importance to other non-personality factors from past research, in the pull request evaluation process. We find that pull requests from authors (requesters) who are more open and conscientious, but less extroverted, have a higher chance of approval. Furthermore, pull requests that are closed by developers (closers) who are more conscientious, extroverted, and neurotic, have a higher likelihood of acceptance. The larger the difference in personality traits between the requester and the closer, the more positive effect it has on pull request acceptance. Finally, although the effect of personality traits is significant and comparable to technical factors, we find that social factors are still more influential on the likelihood of pull request acceptance.

Index Terms—Pull request, GitHub, online collaborative environments, open source systems, personality, Big Five, five-factor model

1 Introduction

GitHub is an online collaborative environment that uses pull request based development where developers work collectively to improve projects. In this process, the developers (requesters) 'fork' the project (i.e., make a personal copy of the project) and make changes to their personal copies of the code to add functionality. Developers can request project managers (closers) to merge the changes they have made in their personal copies to the main branch in the form of a pull request. The project manager then evaluates this pull request on different parameters and decides to either merge (accept) or close (reject) the pull request.

Previous studies on open source systems have consistently shown that the closer's decision to accept a contribution not only depends on the technical quality of the contribution, but also on underlying social impressions [1]. Project managers use social signals, such as 'social distance' (e.g., a requester who follows the closer has a social distance of one), while evaluating the pull request [2]. These social signals were equally as important as technical factors, such as test file inclusion, lines of code modified, and files changed. In addition to social signals, some impressions are directly visible to other developers in the form of comments, displaying different sentiments and behaviours that form the user's underlying personality [3], [4]. Another study suggests that closers look at person-based factors like previous interactions with the developer when they are uncertain about the value of the contribution [5]. Impolite or argumentative comments suggest negative 'personality' and such developers are deemed difficult to work with. Recently, a top Linux developer's behaviour was found to have negative effects on the open source community, which may have pushed a lot of volunteer contributors away from the open source development [6].

Given that non-technical factors, such as interpersonal skills, are important, it is reasonable to believe that understanding developers' personalities can provide valuable information about group dynamics and the success of a project. Personality, by definition, seeks to make predictions about how individuals will behave in a given situation [7]. This has already been widely applied in the field of organizational behaviour to understand job performance and leadership in the workplace. In addition to examining individual personality traits, it is useful to examine how one's personality traits may differ from other members in the group. The question "Do birds of a feather flock together?" is a longstanding question in social psychology [8] and have been studied extensively across many different fields. In artificial intelligence and robotics, it was found that users in rehabilitation preferred assistive robots that matched their own personalities, which in turn could increase therapeutic goals [9]. Conversely, individuals found robots that have complimentary personalities as more intelligent and socially present [10]. A study on pair programming found that variability in personality did not have a great impact on performance [11]. These contradictory findings suggest that personality alone is often not sufficient to predict behaviours, but the combination of personality and context gives a rich understanding of how one will behave in certain situations.

In this regard, it is sensible to investigate the same set of ideas in an online setting like GitHub. The diversity of GitHub

- *The authors are with the David R. Cheriton School of Computer Science, University of Waterloo, Waterloo, ON N2L 3G1, Canada. E-mail: {r3iyer, alex.yun, mei.nagappan}@uwaterloo.ca, jhoey@cs.uwaterloo.ca.*

Manuscript received 19 June 2019; revised 25 Nov. 2019; accepted 29 Nov. 2019. Date of publication 17 Dec. 2019; date of current version 12 Nov. 2021.
(Corresponding author: Rahul N. Iyer.)
Recommended for acceptance by A. Sarma.
Digital Object Identifier no. 10.1109/TSE.2019.2960357

will be reflected by developers exhibiting different behaviours, characterized by varying personality traits. We can observe how these developers have behaved in a given project, extract their personality traits, and examine their influence on the pull request acceptance. Some developers may have strict and high standards of coding, while others may have relaxed standards for acceptance. Similarly, some may be encouraging and readily available to assist outside contributors, while others reject pull requests immediately. With the rich data available on GitHub, understanding the behaviours of developers has become easier.

In this paper, we replicate and extend the work of Tsay *et al.* [2] to examine the role of developers' personality traits. These are derived from the Five-Factor Model (FFM) or the 'Big Five' [12], which consists of five traits: Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism. The main question we address is: "Does the personality of a developer affect the pull request acceptance?".

1.1 Research Questions and Hypotheses

The main research questions are:

- *(RQ0)* Can Tsay *et al.*'s [2] results be replicated on a more recent dataset?
- *(RQ1)* Does the personality of a requester affect the likelihood of the pull request being accepted?
- *(RQ2)* Does the personality of a closer affect the likelihood of the pull request being accepted?
- *(RQ3)* Does the difference in personality between the requester and the closer affect the likelihood of the pull request being accepted?

Based on the personality psychology literature, we postulated that:

- *Openness will have a positive effect on the pull request acceptance.*
- *Conscientiousness will have a positive effect on the pull request acceptance.*
- *Extraversion will have a positive effect on the pull request acceptance.*
- *Agreeableness will have a positive effect on the pull request acceptance.*
- *Neuroticism will have a negative effect on the pull request acceptance.*
- *Variability in personality between the requester and the closer will have a positive effect on the pull request acceptance.*

In sum, the primary contributions of this paper are:

1) Replication of Tsay *et al.*'s [2] work to show generalizability.
2) Empirical evidence showing the effects of requesters' and closers' personality traits, and the difference in personality traits between the requester and the closer on pull request acceptance.
3) Supporting and explaining the empirical evidence from the literature in personality psychology.

2 Background and Related Works

In this section, we first define the term 'personality' according to theories in personality psychology and describe two popular models used to describe one's personality. Further, we summarize related works on the study of personality in software engineering and finally review past research on the pull request evaluation process.

2.1 Theories of Personality

The study of personality has a rich history with many perspectives, but at its core, it is concerned with human nature and patterns of behaviour [13]. It seeks to explain why there exist individual differences in behaviour and to predict how one would behave in a given situation. Personality psychologists have been interested in developing a unified theory of personality and are concerned with the validity of its measurements [14].

Psychologists naturally gravitated towards categorizing people into 'types'. In particular, the Jungian personality types received a lot of attention and are thought to be measured by the Myers–Briggs Type Indicator (MBTI) questionnaire [15]. Despite its popularity, the validity of MBTI has met with many criticisms [16], [17].

One of the most influential personality models is the Five-Factor Model (FFM) or the 'Big Five' [12]. It is comprised of several traits that collectively explain one's disposition and behaviour. The Big Five is empirically well validated, showing good reliability [18], validity [19] and consistency across cultures [14]. We adopt the definitions of traits from the 10 Aspects scale [20]:

- *Openness to Experience:* A measure of intellect and openness. Individuals who are high on Openness to Experience enjoy solving complex problems and show aesthetic appreciation. Hereafter, Openness to Experience will be referred to as Openness.
- *Conscientiousness:* A measure of industriousness and orderliness. Individuals who are high on Conscientiousness are detail-oriented and reliable workers.
- *Extraversion:* A measure of enthusiasm and assertiveness. Individuals who are high on Extraversion are gregarious and take charge in social situations.
- *Agreeableness:* A measure of compassion and politeness. Individuals who are high on Agreeableness are nurturing and trusting of others, and uncomfortable confronting others.
- *Neuroticism:* A measure of volatility and withdrawal. Individuals who are high on Neuroticism are prone to experiencing negative emotions and doubt themselves often.

2.2 Study of Personality in Software Engineering

Past research has focused on evaluating the developers' personality traits using the MBTI. Karn and Cowling [21] examined the association between the personality profile of teams and their performances. The results showed teams that were more heterogeneous in personality worked better together than homogeneous teams. They also reported some evidence of personality clashes among the members that produced intense debates, resulting in no progress. Capretz and Ahmed [22] mapped software engineering tasks to specific MBTI traits using skills required for the tasks and determined which traits would be useful for a team. Although there has been plenty of research on modelling personalities in the software domain using MBTI, Varona *et al.* [23] found many inconsistencies among different works.

Researchers have started to utilize the Big Five measures to study the effects of personality in the software domain. Acuna *et al.* [24] examined the relationships of personality traits with team processes, task characteristics, product quality, and team satisfaction. They found that teams with high job satisfaction tend to have members who scored highly on Agreeableness and Conscientiousness. Martinez *et al.* [25] conducted an empirical case study which suggested relationships between certain Big Five personality traits and software engineering roles. They found Extraversion to be high for the architect and the tester, but low for the developer, while Conscientiousness was found to be low only for the developer. Bell *et al.* [26] examined the relation between personalities of students and their approach to using software design patterns and learning achievements. The results showed no correlations between the personality and individual performance. Salleh *et al.* [27], on the other hand, found a strong correlation between Openness and pair programming performance. In an academic setting, pairs comprised of students with high Openness performed much better than those with low Openness. Kost *et al.* [28] identified two personality groups among students—intense and moderate. The intense personality was characterized by high Extraversion and Openness. When comparing the personalities of software testers and developers in a group, Kanij *et al.* [29] found that testers are higher on Conscientiousness than other developers. Finally, Smith *et al.* [30] found that mangers have high Conscientiousness and Extraversion, while agile developers have high Neuroticism and Extraversion.

With advancements in natural language processing (NLP) and the availability of different tools, researchers have started deriving personality traits from raw text using different psycho-linguistic tools instead of interviews and surveys. Rigby and Hassan [31] studied the personality of developers in the Apache httpd server mailing list using the Linguistic Inquiry and Word Count (LIWC) dictionary. Their results suggested that two top developers responsible for the major Apache releases had similar personalities that were different from the rest of the developers. Bazelli *et al.* [32] conducted analyses on the personality traits of Stack Overflow users using the LIWC dictionary. They retrieved the personality traits from the raw text which included questions asked and answers given by the users. Their results revealed that top reputed authors in Stack Overflow tend to be more extroverted than other users. Licorish *et al.* [33] profiled personalities of developers across the globe for the IBM Jazz repository. The results showed that top contributors tend to score high on Openness, practitioners involved in usability tasks tend to score high on Extraversion, and coders tend to score high on Neuroticism. Rastogi and Nagappan [34] also used the LIWC dictionary to find the personalities of GitHub developers. Their results showed that the top contributor was dramatically different in that they were significantly more neurotic than other contributors.

Newly developed methods are taking advantage of machine learning and are outperforming rule-based methods that use the LIWC dictionary [35]. A recent work by Paruma *et al.* [36] used the *IBM Watson Personality Insights* (same method used in this paper) to retrieve personality traits for clustering developers together. They found a relationship between the personality traits of committers and their social and technical activities on a project. Calefato *et al.* [37] analyzed the developers' personality in Apache projects using the *IBM Watson Personality Insights*. They observed that developers became more conscientious, agreeable, and neurotic over time and found no significant evidence of contributor's team membership affecting their personalities. They also noted that developers who were more open and agreeable have a better chance of becoming project contributors. Our work bares some resemblance to Calefato *et al.*'s work in that we also examine the effects of personality traits on the likelihood of becoming project contributors. However, it is different in that we try to perform the experiment on the data from GitHub pull requests and at a much larger scale. We also position the personality factors with other social and technical factors that help us better gauge the relative importance of personality traits on pull request. Lastly, we examine the closer's personality traits, as well as the difference in personality traits between the requester and the closer.

2.3 Pull Request Evaluation

The first empirical study in this area was done by Tsay *et al.* [2]. They showed that project managers not only use technical factors, but also use social clues while evaluating pull requests. Prior interactions inside the project and 'social distance' were important to the pull request acceptance process. Gousios *et al.* [38] conducted a large-scale analysis of factors affecting pull request acceptance on 1.9 million pull requests. The results reaffirmed the existence of non-technical factors involved in pull request evaluation process. Soares *et al.* [39] showed that the speed of the pull request evaluation process is significant in pull request acceptance. Yu *et al.* [40] studied different factors that explained the latency of a pull request evaluation. They recognized continuous integration and first human response time as important factors.

3 Methodology

In this section, we describe our data selection, data extraction, and data modelling process.

3.1 Data Selection

We considered valid projects from two different sources. The first set of projects is the 11,000 projects used by Tsay *et al.* [2]. The second source was taken from a set of more than one million projects curated by Muniah *et al.* [41]. The dataset contains project meta data including a binary label on project maintenance. The dataset is publicly available on their project website *RepoReaper* [42]. We filtered the *RepoReaper* dataset down to 15,000 projects and kept projects that:

- Were properly maintained which excludes defunct projects, such as student course projects and/or casual projects uploaded by developers.
- Had more than ten issues. This removes projects that have not received a good amount of attention. Issues are usually raised when one notices bugs or requests a feature to be implemented.

- Had at least three contributors. We want projects to have received some contributions from volunteer developers.
- Were not forks. Duplicate copies of the project were not considered.

From a pool of 26,000 projects, we filtered it to 1,860 projects by only considering those projects that have at least 250 closed or merged pull requests to ensure good amount of discussions in the project for the *IBM Watson Personality Insights* to extract the personality traits. We then extracted the pull request data of 1,860 projects using GHTorrent [43], GitHub API [44], and a modified version of the open source code provided by Gousios *et al.* [38]. The data and the model building script can be found here.[1]

3.2 Obtaining Personality Traits

We used the *IBM Watson Personality Insights* [45] to retrieve personalities of the developers from raw text. In addition to the Big Five personality traits, this service can extract other information such as developers' needs and underlying values. The service uses Arnoux *et al.*'s [35] open vocabulary approach that applies Gaussian Processes on GloVe Word Embeddings to infer the personality [46]. The authors were able to achieve similar or better accuracy with 8 times less data. This method outperforms the previous rule-based methods that used LIWC dictionary [45]. Recent works have utilized the *IBM Watson Personality Insights* to compute personality traits [36], [37], [47]. We extracted all the comments made by the developers from 1,860 projects using the GitHub API. To get a reasonable estimate of personality profile, the *IBM Watson Personality Insights* recommends to have at least 600 words and states that accuracy increases as the number of words increases. Hence, we only considered those profiles who have left comments on GitHub and have more than 1,000 words across those comments. We applied the following preprocessing steps before we sent the text to the *IBM Watson Personality Insights* that includes:

- Removing all profiles whose communication language was not English. We used polyglot [2] library's language detector function to detect the dominant language in the text. Manually analyzing the comments revealed that developers sometimes used their native language to communicate with other fellow developers from the same country. Although the service provides personality scores for a few other languages apart from English, we wanted our scores to be consistent, and thus chose only English comments.
- We converted the comments available in the markdown format to html and removed the tags corresponding to the code. Comments made by the developers often have code included in the text. Since we only wanted natural language in the comments, we tried removing code from the text as much as possible.
- Lower-casing all the characters.
- Removing all the special characters except punctuation.

Once we send the raw text to the *IBM Watson Personality Insights*, this service returns a JSON object containing the percentile value between 0 and 1 for each personality trait that represents where the user stands relative to others. We gathered the personality traits of 29,396 developers. Additionally, we considered the pull requests only if we had both closer's and requester's extracted personality traits. This meant some of the developers' personalities were never considered because of not having a corresponding closer's or requester's personality information. Our final dataset included a total of 501,327 pull requests from 1,860 projects and had 16,935 developers. We describe the data modeling process in the next section.

3.3 Overview of the Data Modelling Process

3.3.1 Feature Selection

We used technical-, social-, and personality-related factors as independent variables and pull request acceptance as the binary dependent variable. Table 1 shows the different features used in the modelling process and also the descriptive statistics of pre-transformation values. The statistics include feature values at 5th percentile, 50th percentile (median), 95th percentile, and the mean.

We used social and technical factors (Table 1) in all research questions: RQ0–RQ3. We used the Big Five personality traits of the requester for RQ1, closer's personality traits for RQ2, and absolute difference in the personality traits between requester and closer for RQ3.

3.3.2 Data Preparation

We examined the distribution of the data and found that several features were skewed. Thus, they were normalized via a log transformation. Since the range of the each feature varies drastically, we scaled the data using the default `scale()` function provided by *R* [48]. This function transforms the features using the z-score transformation which subtracts the mean and divides by the standard deviation so that each feature has a unit standard deviation. Furthermore, to remove correlated features, we used variable clustering on the features and used $|\rho| = 0.6$ as the threshold. Variable clustering analysis uses hierarchical clustering on the correlation values derived from either a Pearson or a Spearman correlation test, to group features together. We used Pearson correlation after normalising the data. We found that `file_changes` was grouped together with `total_churn` and thus decided to remove `file_changes`. After removing, we computed the clusters again but did not find any more correlations. We further analyzed the case of multicollinearity in the data by using Variance Inflation Factor (VIF) and setting the threshold as 5 (as used by Tantithamthavorn *et al.*'s study [49]) to identify them.

3.3.3 Model Construction

We used a mixed effects logistic regression model (the same model used by Tsay *et al.*) provided by `lme4` library in *R* [50]. Pull requests grouped into a hierarchy of repository names, requester and closer, and represent data over time. By using a mixed effects model, we explicitly model the correlation among the hierarchy. We used the project name, requester and closer as random effects, and used features described in Table 1 as fixed effects. We report the odds ratio of the mixed effects model and also the confidence interval of the odds ratio using bootstrapping. Bootstrapping is a method to

1. https://bit.ly/2sJbmgD
2. https://pypi.org/project/polyglot/

TABLE 1
Different Features Used with Descriptive Statistics

Category	Variables	Description	5%	Mean	Median	95%	SD
Social Factors (Tsay et al.)	*social_distance*	A binary variable that tells if the pull request author follows the closer.	-	-	-	-	-
	prior_interaction	Indicates previous interaction of the requester in the project.	5	1840	640	7423	3545
	followers	Total number of followers of the requester at the time of data collection. This is constant for all appearances of a specific requester in the data.	0	114	21	370	583
Technical Factors (Tsay et al.)	*test_file_present*	A binary variable that represents if the pull request contains test files.	-	-	-	-	-
	total_churn	Represents the total number of lines changed in a pull request.	0	1101	27	1959	20824
	files_changed	Total number of file changed in a pull request	0	9.5	2	32	44
	num_comments	Total number of comments pertaining to the pull request	0	4	1	18	11
	main_team_member	Indicates if the requester is a core team member or not.	-	-	-	-	-
	team_size	Size of the core team of the project at the time of data collection.This is constant for all the pull requests of a particular project.	13	88	102	238	105
	stars	Indicates the number of stars in a project at the time of data collection.	3	3122	614	14135	6860
	project_age	Age of project in days at the time of data collection	1126	1998	2019	2893	530
Personality Factors (Ours)	*openness*	Represents the Openness of a requester/closer.	0.84	0.95	0.98	0.99	0.11
	conscientiousness	Represents the Conscientiousness of a requester/closer.	0.11	0.36	0.34	0.67	0.17
	extraversion	Represents the Extraversion of a requester/closer.	0	0.06	0.04	0.21	0.08
	agreeableness	Represents the Agreeableness of a requester/closer.	0	0.02	0.0	0.7	0.05
	neuroticism	Represents the Neuroticism of a requester/closer.	0.54	0.76	0.78	0.92	0.13
	diff_X_abs	Absolute difference between closer's and requester's personality.Where X is different personality traits.	-	-	-	-	-

estimate the empirical distribution by generating data multiple times with replacement. A bootstrap sample is a resample of the same size, but with replacements. If u is a true distribution of the statistic, we aim to find $u*$, which is an approximation of the the true distribution. We retrieved 50 bootstrap samples from the dataset and separately modelled each of them. Later, we extracted the odds ratio from each of the 50 models and computed confidence interval on them. For each research question, we report the influence of different features as an increase or a decrease in odds ratio on pull request acceptance. If the odds ratio is greater than 1, this represents a positive relationship of the independent variable on the dependent variable, whereas a value less than 1 suggests negative relationship. We also report the marginal and the conditional R-Squared values for each model. Marginal R-squared value represents the variation explained by the fixed-factors alone, while conditional R-Squared represents the variation explained by both fixed and random effects in the model. We used `MuMIn` R package to compute the marginal and the conditional R-Squared values [51]. Lastly, we consider the odds ratio to be significant only if the p-value is less than 0.001.

4 Results

For each research question, we will address our motivation, approach, and results. We will follow up the results with explanations from the psychology literature in the Discussion.

4.1 (RQ0) Can Tsay *et al.*'s Results be Replicated on a More Recent Dataset?

Motivation. We replicated Tsay *et al.*'s [2] findings on a more recently mined dataset to determine if the results still hold. GitHub has experienced a tremendous growth in the past five years with the number of repositories increasing from 4.6 million in 2013 to more than 96 million in 2018 [52]. During the past five years, new and advanced tools have been developed to facilitate better project management and to track code changes precisely. The replication process provides an important insight on the generalizability of Tsay *et al.*'s results to a dataset extracted at a different point in time. Additionally, creating a baseline model with only technical and social factors assist in comparing other personality models with social, technical, and personality factors.

Approach. We used the same modelling technique as suggested by Tsay *et al.* [2] on a new set of data. We used a mixed effects logistic regression using only the features used in their work.

Results. We replicated the effects of social and technical factors on pull request acceptance. All factors had similar overall influences except for the follower count, for which we did not find significant effects. We found that the effects of having test file included in a pull request decreased from 17 percent (reported by Tsay *et al.* [2]) to 8 percent in our model. Social distance and prior interaction were the most important factors that influence pull request acceptance positively. Moreover, number of comments in a pull request and the number of stars of a project were the most important factors that influence pull request acceptance negatively. Although there were fluctuations in the odds ratio of all the features, we believe Tsay *et al.*'s overall results still stand. The change in some of the factors may be due to the selection of projects. Since we only included projects that have more than 250 closed or merged pull requests, it may have resulted in selection of projects that have been active for a longer duration. Table 2 provides a comparison between our results and Tsay *et al.*'s results.

4.2 (RQ1) Does the Personality of a Requester Affect the Likelihood of the Pull Request Being Accepted?

Motivation. A requester can either be a core team developer or an outside contributor. They often make useful suggestions

TABLE 2
Odds Ratio of the Model for RQ0

Variables	Ours	Tsay *et al.*'s
(Intercept)	2.81	1.106
test_file	1.08 (***)	1.171 (***)
total_churn	0.9 (***)	0.738 (***)
files_changed	-	0.927 (***)
social_distance	2.35 (***)	2.870 (***)
num_comments	0.68 (***)	0.454 (***)
prior_interaction	1.53 (***)	1.356 (***)
followers_current	1.	1.181 (***)
main_team_member	1.16 (***)	1.636 (***)
age_current	0.91 (***)	0.820 (***)
team_size	0.99	0.954 (**)
stars_current	0.53 (***)	0.648 (***)
test_file:num_comments	1.12 (***)	1.106 (***)
total_churn:num_comments	1.06 (***)	1.169 (***)
file_changed:num_comments	-	1.035 (***)
social_connection:num_comments	0.92 (***)	0.796 (***)
num_comments:prior_interaction	1.05 (***)	1.142 (***)
AIC	**394718**	-

*$p < .05$. **$p < .01$. ***$p < .001$.
$R^2 M = 0.11$ $R^2 C = 0.58$

in the form of pull request that can have a positive impact on the project. Previous work by Marlow *et al.* [5] suggests that project owners use personality clues derived from the communication activities of a developer to get an idea of what the person is like to work with. By examining the requesters, we aimed to understand whether having specific personality traits lead to a higher likelihood of pull request acceptance.

Approach. We modelled the personality traits of the requester—Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism—along with the features used by Tsay *et al.* [2]. Similar to RQ0, we used a mixed effects logistic regression to model the data.

Results. We observed the following effects (Table 3).

Openness: Openness of a requester was positively associated with pull request acceptance, increasing acceptance likelihood by 7 percent per unit increase.

Conscientiousness: Conscientiousness of a requester was positively associated with pull request acceptance, increasing acceptance likelihood by 6 percent per unit increase.

Extraversion: Extraversion of a requester was negatively associated with pull request acceptance, decreasing acceptance likelihood by 6 percent per unit increase.

Agreeableness: Agreeableness of a requester was positively associated with pull request acceptance, increasing acceptance likelihood by 2 percent per unit increase but the result was not significant.

Neuroticism: Neuroticism of a requester was negatively associated with pull request acceptance, increasing the acceptance likelihood by 3 percent per unit increase but the result was not significant.

TABLE 3
Odds Ratio of the Mixed Effect Model with Requester's Personality

Variables	Single Run Odds Ratio	Bootstrapping 95% CI
(Intercept)	2.87 (***)	-
test_file	1.08 (***)	[1.05,1.11]
total_churn	0.9 (***)	[0.89,0.9]
social_distance	2.35 (***)	[2.59,2.8]
num_comments	0.68 (***)	[0.65,0.69]
prior_interaction	1.52 (***)	[1.51,1.57]
followers_current	0.99	[0.95,0.98]
main_team_member	1.15 (***)	[1.12,1.2]
age_current	0.91	[0.83,0.93]
team_size	0.99	[0.85,0.98]
stars_current	0.54 (***)	[0.45,0.49]
openness	1.07 (***)	[1.05,1.08]
conscientiousness	1.05 (***)	[1.03,1.07]
extraversion	0.94 (***)	[0.93,0.95]
agreeableness	1.01	[1.,1.02]
neuroticism	0.97	[0.94,0.98]
test_file x num_comments	1.12 (***)	[1.11,1.15]
total_churn x num_comments	1.06 (***)	[1.06,1.08]
social_connection x num_comments	0.92 (***)	[0.9,0.96]
num_comments x prior_interaction	1.05 (***)	[1.06,1.08]
AIC	**394635**	

*$p < .05$. **$p < .01$. ***$p < .001$.
$R^2 M = 0.11$ $R^2 C = 0.55$

Openness and Conscientiousness have positive effects on the pull request acceptance, while Extraversion has a negative effect.

4.3 (RQ2) Does the Personality of a Closer Affect the Likelihood of the Pull Request Being Accepted?

Motivation. A closer, unlike a requester, is always part of the core team. Their responsibilities include reviewing pull requests, suggesting changes, prioritizing tasks, and engaging in secondary functions (e.g., attracting new developers). They are essential part of the pull request process as their decisions directly impact the functioning of the project and can move the project forward. Wang [53] conducted an empirical study on 116 projects and found positive evidence of project manager's personality affecting project success. While some closer may choose to reject an imperfect pull request, others may be more supportive and try to address the imperfection with the requester. By analyzing the closers, we aimed to understand whether specific personality traits affect the likelihood of the pull request getting accepted.

Approach. We replaced the requester's personality traits with the closer's personality traits and modelled them along with the factors used in RQ0.

TABLE 4
Odds Ratio of the Mixed Effect Model with Closer's Personality

Variables	Single Run Odds Ratio	Bootstrapping 95% CI
Intercept	2.87	-
test_file	1.08 (***)	[1.05,1.11]
total_churn	0.90 (***)	[0.89, 0.91]
social_distance	2.35 (***)	[2.35,2.83]
num_comments	0.68 (***)	[0.65,0.68]
prior_interaction	1.49 (***)	[1.52,1.56]
followers_current	0.98	[0.94,0.99]
main_team_member	1.16 (***)	[1.12,1.21]
age_current	0.92 (***)	[0.86,0.95]
team_size	0.97	[0.88,1.0]
stars_current	0.54 (***)	[0.44,0.51]
openness	1.05 (**)	[1.02,1.10]
conscientiousness	1.12 (***)	[1.11,1.18]
extraversion	1.06 (***)	[1.06,1.13]
agreeableness	1.01	[0.99,1.04]
neuroticism	1.08 (***)	[1.06,1.14]
test_file x num_comments	1.12 (***)	[1.08,1.16]
total_churn x num_comments	1.06 (***)	[1.05,1.08]
social_connection x num_comments	0.92 (***)	[0.88,0.95]
num_comments x prior_interaction	1.05 (***)	[1.05,1.08]
AIC	**394548**	

*$p < .05$. **$p < .01$. ***$p < .001$.
$R^2M = 0.12$ $R^2C = 0.57$

Results. We observed the following effects (Table 4).

Openness: Openness of a closer was positively associated with pull request acceptance, increasing acceptance likelihood by 5 percent per unit increase but the result was not significant

Conscientiousness: Conscientiousness of a closer is positively associated with pull request acceptance, increasing acceptance likelihood by 12 percent per unit increase.

Extraversion: Extraversion of a closer was positively associated with pull request acceptance, increasing acceptance likelihood by 6 percent per unit increase.

Agreeableness: Agreeableness of a closer was positively associated with pull request acceptance, increasing acceptance likelihood by 1 percent per unit increase but the result was not significant.

Neuroticism: Neuroticism of a closer was positively associated with pull request acceptance, increasing acceptance likelihood by 8 percent per unit increase.

> Conscientiousness, Extraversion, and Neuroticism have positive effects on the pull request acceptance.

4.4 (RQ3) Does a Difference in Personality Between Requester and Closer Affect the Likelihood of the Pull Request Being Accepted?

Motivation. There have been numerous contradictory studies by software engineering community on the effects of personality differences in the offline setting. Baddoo and Hall [54] showed that having differences in personalities sometimes lead to personality clashes which demotivate and frustrate developers. Similarly, Karn *et al.* [55] observed occasional personality clashes in one of the software engineering teams they studied. Conversely, Carpretz and Ahmed [22] showed that having diverse personality is an essential part to solve different problems in a software development environment. A study by Chen *et al.* [56] looked at Wikipedia, an another online collaborative environment and found diversity in tenure and interests among the people had a positive effect on productivity. There has been no known study that examines the effect of personality difference in an online software engineering collaborative environment and hence motivated us to examine the effect of personality differences on pull request acceptance likelihood. Analyzing the differences in the personality traits assist us to understand whether a higher difference leads to personality clashes or prompts more discussions, and in turn a higher pull request acceptance.

TABLE 5
Odds Ratio of the Mixed Effect Model with Personality Differences

Variables	Single Run Odds Ratio	Bootstrapping 95% CI
(Intercept)	3.34 (***)	-
test_file	1.09 (***)	[1.05,1.12]
total_churn	0.92 (***)	[0.9,0.93]
social_distance	1.81 (***)	[1.86, 2.03]
num_comments	0.66 (***)	[0.65,0.67]
prior_interaction	1.66 (***)	[1.63,1.69]
followers_current	1.07 (***)	[1.06,1.11]
main_team_member	1.27 (***)	[1.23,1.31]
age_current	0.92 (***)	[0.87,0.93]
team_size	0.96	[0.89,1]
stars_current	0.55 (***)	[0.44,0.50]
diff_openness_abs	1.01	[1.01,1.04]
diff_conscientiousness_abs	1.29 (***)	[1.29,1.35]
diff_extraversion_abs	1.12 (***)	[1.11,1.16]
diff_agreeableness_abs	1.02 (***)	[1.0,1.04]
diff_neuroticism_abs	1.22 (***)	[1.21,1.27]
test_file x num_comments	1.11 (***)	[1.09,1.15]
total_churn x num_comments	1.06 (***)	[1.05,1.07]
social_connection x num_comments	0.93 (***)	[0.89,0.97]
num_comments x prior_interaction	1.05 (***)	[1.05,1.08]
AIC	**390495**	

*$p < .05$. **$p < .01$. ***$p < .001$.
$R^2M = 0.13$ $R^2C = 0.56$

Approach. We considered the effects of personality differences in the model by adding the absolute differences between personality traits of requester and closer along with the features used in RQ0.

Results. We observed the following effects (Table 5).

Openness: The difference between requester's Openness and the closer's Openness was positively associated with

TABLE 6
Effect of Personality Traits on Pull Request Acceptance for Different Research Questions

Research Question	Openness	Conscientiousness	Extraversion	Agreeableness	Neuroticism
RQ1 (Requester's personality)	7%	5%	-6%	NS	NS
RQ2 (Closer's personality)	NS	12%	6%	NS	8%
RQ3 (Personality difference)	NS	29%	12%	2%	22%

NS implies not significant. Each cell represents whether the likelihood of pull request acceptance increases or decreases.

pull request acceptance, increasing the acceptance likelihood by 1 percent per unit increase but the result was not significant.

Conscientiousness: The difference between requester's Conscientiousness and the closer's Conscientiousness was positively associated with pull request acceptance, increasing the acceptance likelihood by 29 percent per unit increase.

Extraversion: The difference between requester's Extraversion and closer's Extraversion was positively associated with pull request acceptance, increasing the acceptance likelihood by 12 percent per unit increase.

Agreeableness: The difference between requester's Agreeableness in closer's Agreeableness was associated with pull request acceptance, increasing the acceptance likelihood by 2 percent per unit increase. Although significant, the effect size is extremely small for the factor to be considered important.

Neuroticism: The difference between requester's Neuroticism and the closer's Neuroticism was positively associated with pull request acceptance, increasing the acceptance likelihood by 22 percent per unit increase.

The absolute differences in Openness, Conscientiousness, Extraversion, and Neuroticism affect the pull request acceptance positively.

A summary of results for different research questions is presented in Table 6. We also saw a decrease in the Akaike Information Criterion (AIC) statistic for the models used in RQ1–RQ3 from the model used in RQ0. This suggests that models with personality traits have a better fit compared to the model without personality traits. In RQ3, the AIC decreased from 394,718 to 390,495 (Tables 5 and 2).

5 Discussion

We now present additional insights gained by applying the literature in personality psychology.

Openness is considered the creativity domain and individuals with high Openness welcome new experiences and are intellectually curious [57]. A previous study has shown that Openness is positively correlated with creative behaviour in an office setting [58]. In the context of open-source software development, we believe submitting novel and interesting pull requests would be considered a form of creative behaviour. As such, we hypothesized that both requester's and closer's Openness would have a positive association with pull request acceptance. Our result indeed showed that pull requests were more likely to be accepted when the requester was high on Openness. In addition to their creative ideas, highly open individuals tend to articulate their thoughts and ideas more cogently [20]; thus, these requesters are more persuasive to the project members. Similarly, when closers were high on Openness, pull requests were more likely to be accepted, but this result was not significant.

Conscientiousness being the measure of industriousness, it is typically associated with high job performance [59]. We hypothesized that the effect of Conscientiousness on pull request acceptance would be positive regardless of the requester's identity (i.e., being a core team member versus outside contributor). It is worth noting that the Conscientiousness of the requester was not as significant as other personality traits; nevertheless, the results were as expected. Given that highly conscientious individuals are well-organized and have strong work ethics [60], such requesters are detail-oriented are likely to make fewer errors in their work, thereby increasing their likelihood of acceptance.

Previous work noted that individuals with high Conscientiousness make effective managers and are likely to occupy leadership roles [61]. As such, we suspect that highly conscientious closers assume the leadership role, motivating others to produce good work. This is reflected in a sample comment where a highly conscientious developer said: "Thanks for reaching out. I want to loop in a few people..." (Table 7). As a result, if a higher quality of work is achieved, then it is also more likely to be accepted by the closers.

Extraversion as a measure of one's social engagement, individuals with high Exraversion are gregarious and very active in social settings. We hypothesized that both requester's and closer's Extraversion would have a positive association with pull request acceptance. This was not true in the case of the requester; in particular, pull requests were less likely to be accepted when the requester was high on Extraversion. This finding may be explained by the assertiveness characteristic of highly extraverted individuals [62], which may be interpreted as aggressive by fellow developers. For the closer, pull requests were more likely to be accepted when they were high on Extraversion, as expected. Mehrabian, for example, found a positive correlation between Extraversion and dominance [63]. This may suggest that when the closer is more extraverted, their assertiveness is more likely to be interpreted as being a good leader. Our result is also consistent with Wang and Li's study that found a positive correlation between Extraversion and leadership performance of the project managers [53].

Agreeableness as a factor in interpersonal interaction, highly agreeable individuals are thought to be compassionate and polite. In the context of GitHub, agreeable developers would be polite and respectful in their communications in the form of a pull request. As such, we hypothesized a positive effect on pull request acceptance regardless of the developers' identity. A study by Calefato *et al.* [64] found a positive association of closer's Agreeableness with pull request acceptance. However, we did not find any significant

TABLE 7
Sample Comments from Profiles with High Scores on Each Trait

Personality	Sample Comments
Openness	That's possible, though it wouldn't be my approach. I would go for a context_id integer within the context structure and set it via mbedtls_ssl_init() with an unique id. But in the end, it's you who will read the logfile when I have something to report. :) So, i'm okay to try this method.
	Ok better than the previous commit but still not perfect :). Not fixed yet then, lets continue
Conscientiousness	I'll only merge if all the checks are passed.
	Thanks for reaching out. I want to loop in a few people. Can you let me know if this is consistent for all web views or only initial ones? I want to figure out if this is because the screenshot is happening before the view loads or if it's the way the screenshot is taken.Thanks!
Extraversion	Its okay if it didn't work, thanks for attempting though. Keep the contributions coming in!!!
	Cheers, this seems like a big contribution. Thanks for the pull request. This will improve the app's functionality drastically.
Agreeableness	I apologize for troubling you. I don't intend to add this package against your will.
	Thank you for pointing out it. I have amended the pull request
Neuroticism	I'm not an accountant but I have worked more than you in this field.I'm pretty sure that US based businesses need to charge VAT for EU customers.Digital goods are actually considered services. Do it like WooCommerce to solve this tax nightmare.
	Thats a bummer, anyway if its not fixed soon i'll revert the master to stable.

effects of Agreeableness associated with requesters or with closers. The discrepancy may be due to the differences in our sample data and modelling method.

Given that Neuroticism measures one's sensitivity to negative emotions, high Neuroticism has been implicated in negative aspects in the workplace, including higher turnover rate [59], less work engagement [65], and lower job performance [66]. These negative trends led us to hypothesize the effect of Neuroticism would be negative. Our results showed no significant effect of the requester's Neuroticism on the likelihood of pull request acceptance. Surprisingly for the closer, there was a positive effect such that pull requests were more likely to be accepted when the closer was high on Neuroticism.

One of the reasons for this association may be due to the fact that closers express more negative sentiments in the communication thread to help the requesters fix their pull requests. If a requester, for example, suggests new features, the closer with high Neuroticism may share their opinions as to how the new feature might not work. While being more critical, the closers give better feedback and hence get the pull request to a better condition. Further studies are required to understand the effects of high Neuroticism on pull request acceptance.

The study of personality differences of group members and its relation to group performance is somewhat limited. We based our hypothesis based on a study that found team personality diversity was positively related to team performance [67]. We indeed found that pull requests were more likely to be accepted when the difference was greater. This greater difference or the diversity argument has been explored in various fields including artificial intelligence and organizational behaviour. When forming multi-agent teams, a greater strategic diversity, even with weak agents, outperformed strong but less diverse teams [68]. In addition, a comprehensive review noted that diversity of personality in the workplace led to positive performance [69]. While there is a limited number of studies that examine the differences in personality traits within the software engineering context, variability in personality was found to increase the amount of communication-intensive collaboration in pair programming [70]. Following this evidence, we believe that the diverse pool of developers in a given project is associated with higher likelihood of pull request acceptance.

While the effects of personality traits on pull request acceptance are present and significant, we note that non-personality factors have more effects on the pull request acceptance. In all the research questions, we found social distance to have the highest importance. For RQ1, the most important personality factor was Openness (1.05 percent) compared to social distance (1.25 percent). Neuroticism had the highest importance (1.12 percent) among the personality factors, but social distance was the most important overall (2.22 percent). Finally for RQ3, difference in Conscientiousness (1.3 percent) came out as the most important but prior interactions (1.66 percent) was the most important overall. It is evident that social factors are the most influential factors when it comes to predicting pull request acceptance. In addition to social factors, both personality factors and technical factors are important in the pull request acceptance predictions. Our results specifically show that personality factors were as important as the technical factors. In the model with the requester's personality, we have Openness that has similar odds ratio to test file inclusion and number of line changed. Further, in the model with closer's personality, we have Conscientiousness and Neuroticism that are better predictors than technical factors.

5.1 Implications for Practice

In spite of our novel findings, one has to be cautious to suggest that there is one definitive personality profile that can guarantee the success of an open-source project. Personality explains one's stable disposition and many patterns of behaviours. Theoretically, this entails that one cannot simply change their personality at their own will. Furthermore, we saw the importance of absolute difference in personality traits between the requester and the closer from our findings, as well as the positive influence of diversity from the literature. We also note that while we mainly examined the likelihood of one's pull request being accepted, there are many different metrics that could be used to infer the success of a given project.

That said, we believe having good communication skills is crucial for developers. This would be characterized by being civil and polite, while expressing their ideas cogently. On the other hand, we would encourage project managers to seek out fellow developers who are highly open as they are likely to bring creative solutions to the team. Conscientious developers will not only work diligently, but will be detail-oriented and deliver high quality standard of work. We are unable to make specific suggestions regarding the traits Agreeableness and Neuroticism, as our results were non-significant for Agreeableness and opposite from the trends seen in the literature for Neuroticism. Thus, further investigations are warranted.

6 Threats to Validity

Recall that the personality traits of developers were obtained using the *IBM Watson Personality Insights*. It is possible that the personality model generated by this service is not actually representative of the developers' true personality. Even if this is true, we claim that this is not an issue as we only care about the digital footprint or perceived personality—that is, how developers are perceived by other developers on GitHub. Moreover, there is evidence that self-reported personality measures are not the only valid way to measure one's personality. Self-reports show strong correlations with observer ratings [71]. Future works could explore indirect communication media, such as Gitter,[3] Discord,[4] or Slack channels,[5] to extract personality traits and remap them to GitHub usernames. These communication channels are known to be more informal: developers often express sentiments more casually and use emojis in the conversations, making it easier to extract their digital personalities.

We retrieved the personalities of the developers from the comments on the pull request discussion thread. We agree there may be some concerns with the performance of the service on software engineering-specific text content, but we mitigate this issue by using different regular expressions to remove code from the text. We first converted the markdown into html and then removed all the code in the `<pre>` and `<code>` html tags, but some code still ended up not getting filtered. A manual inspection of 50 comments with code suggested that 42 of them were properly formatted in the markdown. Since we only considered developers who have communicated at least 1,000 words as comments in a GitHub project, some developers were ignored. We believe the 16,935 sample of highly active users is large enough to perform an empirical analysis.

There may be some concerns with the data not being representative of the true population. We tackled this concern by taking project samples from 2 different sources: projects derived from the original Tsay *et al.*'s work [2] and a sample extracted from the set of valid projects from RepoRepear. We analyzed a total of 501,327 pull requests with 16,935 users, which in itself is a good representation of the population. Further, we followed the same process as used by Tsay *et al.*, and additionally perform bootstrapping and report the confidence intervals to make our results more reliable.

Finally, in this study we have only considered the requester and closer while evaluating the effects of personality factors. There are other developers who might express their opinions in the pull request thread who might have valuable feedback for the requester, but we believe the final decision is still made by the closer and thus chose to solely focus on closer's personality rather than personality of the group.

3. https://gitter.im/
4. https://discordapp.com
5. https://slack.com/

7 Conclusion

We presented an empirical study showing the effects of developers' personality traits on pull request acceptance of GitHub. We first replicated Tsay *et al.*'s work [2], noting the importance of social and technical factors. Our results showed that the likelihood of pull request acceptance is significantly influenced by personality traits of developers. We also noted the absolute difference in personality traits between the requester and the closer results in positive effects, suggesting that diversity in personality is beneficial in open source projects. It is important to note that the personality traits had a significant effect comparable to the technical factors, but not to the extent of social factors. In sum, we observe requesters who are high on Openness, Conscientiousness, and low on Extraversion have a higher likelihood of getting the pull request accepted. Similarly, a closer who are high on Openness, Conscientiousness, Extraversion, and Neuroticism accepts more pull requests.

Our work provides a stepping-stone for researchers to conduct further experiments, observing social and group dynamics in online collaborative work environments. We believe that expensive qualitative studies such as detailed developer interviews, can now be conducted and can shed light on the underlying mechanisms for online collaboration.

Acknowledgments

This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) and the Social Sciences and Humanities Research Council of Canada (SSHRC) through the Trans-Atlantic Platform's Digging into Data Program.

References

[1] L. Dabbish, C. Stuart, J. Tsay, and J. Herbsleb, "Social coding in GitHub: Transparency and collaboration in an open software repository," in *Proc. ACM Conf. Comput. Supported Cooperative Work*, 2012, pp. 1277–1286.
[2] J. Tsay, L. Dabbish, and J. Herbsleb, "Influence of social and technical factors for evaluating contribution in GitHub," in *Proc. 36th Int. Conf. Softw. Eng.*, 2014, pp. 356–366.
[3] A. Murgia, P. Tourani, B. Adams, and M. Ortu, "Do developers feel emotions? An exploratory analysis of emotions in software artifacts," in *Proc. 11th Work. Conf. Mining Softw. Repositories*, 2014, pp. 262–271.
[4] E. Guzman, D. Azócar, and Y. Li, "Sentiment analysis of commit comments in GitHub: An empirical study," in *Proc. 11th Work. Conf. Mining Softw. Repositories*, 2014, pp. 352–355.

[5] J. Marlow, L. Dabbish, and J. Herbsleb, “Impression formation in online peer production: Activity traces and personal profiles in GitHub,” in *Proc. Conf. Comput. Supported Cooperative Work*, 2013, pp. 117–128.

[6] J. Wakefield, “Linus Torvalds: ‘I’ll never be cuddly but I can be more polite.’” Accessed 2018. [Online]. Available: https://www.bbc.com/news/technology-45664640

[7] R. B. Cattell, “Personality: A systematic theoretical and factual study,” 1st ed., New York, NY, USA: McGraw-Hill. vol. vii, p. 689, 1950. [Online]. Available: https://doi.org/10.1037/10773-000

[8] D. Byrne and W. Griffitt, “Similarity and awareness of similarity of personality characteristics as determinants of attraction,” *J. Exp. Res. Pers.*, vol. 3, pp. 179–186, 1969.

[9] A. Tapus and M. J. Mataric, “Socially assistive robots: The link between personality, empathy, physiological signals, and task performance,” in *Proc. AAAI Spring Symp.: Emotion Pers. Soc. Behav.*, 2008, pp. 133–140.

[10] K. M. Lee, W. Peng, S.-A. Jin, and C. Yan, “Can robots manifest personality?: An empirical test of personality recognition, social responses, and social presence in human–robot interaction,” *J. Commun.*, vol. 56, no. 4, pp. 754–772, 2006.

[11] J. E. Hannay, E. Arisholm, H. Engvik, and D. I. Sjoberg, “Effects of personality on pair programming,” *IEEE Trans. Softw. Eng.*, vol. 36, no. 1, pp. 61–80, Jan./Feb. 2010.

[12] P. Costa, R. McCrae, and G. Kay, “Persons, places, and personality: Career assessment using the revised NEO personality inventory,” *J. Career Assessment*, vol. 3, pp. 123–139, 1995. doi: 10.1177/106907279500300202.

[13] R. Hogan, “What is personality psychology?” *Psychol. Inq.*, vol. 9, no. 2, pp. 152–153, 1998.

[14] J. M. Digman, “Personality structure: Emergence of the five-factor model,” *Annu. Rev. Psychol.*, vol. 41, no. 1, pp. 417–440, 1990.

[15] I. B. Myers, M. H. McCaulley, and R. Most, *Manual: A Guide to the Development and Use of the Myers-Briggs Type Indicator*. Palo Alto, CA, USA: Consulting Psychologists Press, 1985.

[16] R. R. McCrae and P. T. Costa Jr, “Reinterpreting the myers-briggs type indicator from the perspective of the five-factor model of personality,” *J. Pers.*, vol. 57, no. 1, pp. 17–40, 1989.

[17] D. J. Pittenger, “Cautionary comments regarding the myers-briggs type indicator,” *Consulting Psychol. J.: Pract. Res.*, vol. 57, no. 3, 2005, Art. no. 210.

[18] C. Viswesvaran and D. S. Ones, “Measurement error in big five factors personality assessment: Reliability generalization across studies and measures,” *Educ. Psychol. Meas.*, vol. 60, no. 2, pp. 224–235, 2000.

[19] R. R. McCrae and P. T. Costa, “Validation of the five-factor model of personality across instruments and observers,” *J. Pers. Soc. Psychol.*, vol. 52, no. 1, 1987, Art. no. 81.

[20] C. G. DeYoung, L. C. Quilty, and J. B. Peterson, “Between facets and domains: 10 aspects of the big five,” *J. Pers. Soc. Psychol.*, vol. 93, no. 5, 2007, Art. no. 880.

[21] J. Karn and T. Cowling, “A follow up study of the effect of personality on the performance of software engineering teams,” in *Proc. ACM/IEEE Int. Symp. Empir. Softw. Eng.*, 2006, pp. 232–241.

[22] L. F. Capretz and F. Ahmed, “Making sense of software development and personality types,” *IT Prof.*, vol. 12, no. 1, pp. 6–13, Jan./Feb. 2010.

[23] D. Varona, L. F. Capretz, Y. Piñero, and A. Raza, “Evolution of software engineers’ personality profile,” *ACM SIGSOFT Softw. Eng. Notes*, vol. 37, no. 1, pp. 1–5, 2012.

[24] S. T. Acuña, M. Gómez, and N. Juristo, “How do personality, team processes and task characteristics relate to job satisfaction and software quality?” *Inf. Softw. Technol.*, vol. 51, no. 3, pp. 627–639, 2009.

[25] L. G. Martínez, A. Rodríguez-Díaz, G. Licea, and J. R. Castro, “Big five patterns for software engineering roles using an ANFIS learning approach with RAMSET,” in *Proc. Mex. Int. Conf. Artif. Intell.*, 2010, pp. 428–439.

[26] D. Bell, T. Hall, J. E. Hannay, D. Pfahl, and S. T. Acuna, “Software engineering group work: Personality, patterns and performance,” in *Proc. 48th Annu. Conf. Comput. Personnel Res. ACM SIGMIS*, Ser. SIGMIS-CPR 10, 2010, pp. 43–47. [Online]. Available: https://doi.org/10.1145/1796900.1796921

[27] N. Salleh, E. Mendes, and J. Grundy, “Investigating the effects of personality traits on pair programming in a higher education setting through a family of experiments,” *Empir. Softw. Eng.*, vol. 19, no. 3, pp. 714–752, Jun. 2014. [Online]. Available: https://doi.org/10.1007/s10664-012-9238-4

[28] M. V. Kosti, R. Feldt, and L. Angelis, “Personality, emotional intelligence and work preferences in software engineering: An empirical study,” *Inf. Softw. Technol.*, vol. 56, no. 8, pp. 973–990, 2014. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0950584914000639

[29] T. Kanij, R. Merkel, and J. Grundy, “An empirical investigation of personality traits of software testers,” in *Proc. IEEE/ACM 8th Int. Workshop Cooperative Human Aspects Softw. Eng.*, May 2015, pp. 1–7.

[30] E. K. Smith, C. Bird, and T. Zimmermann, “Beliefs, practices, and personalities of software engineers: A survey in a large software company,” in *Proc. 9th Int. Workshop Cooperative Human Aspects Softw. Eng.*, 2016, pp. 15–18.

[31] P. C. Rigby and A. E. Hassan, “What can OSS mailing lists tell us? A preliminary psychometric text analysis of the apache developer mailing list,” in *Proc. 4th Int. Workshop Mining Softw. Repositories*, 2007, Art. no. 23.

[32] B. Bazelli, A. Hindle, and E. Stroulia, “On the personality traits of stackoverflow users,” in *Proc. 29th IEEE Int. Conf. Softw. Maintenance*, 2013, pp. 460–463.

[33] S. A. Licorish and S. G. MacDonell , “Personality profiles of global software developers,” in *Proc. 18th Int. Conf. Eval. Assessment Softw. Eng.*, 2014, Art. no. 45.

[34] A. Rastogi and N. Nagappan, “On the personality traits of GitHub contributors,” in *Proc. IEEE 27th Int. Symp. Softw. Rel. Eng.*, 2016, pp. 77–86.

[35] P.-H. Arnoux, A. Xu, N. Boyette, J. Mahmud, R. Akkiraju, and V. Sinha, “25 tweets to know you: A new model to predict personality with social media,” 2017. [Online]. Available: https://aaai.org/ocs/index.php/ICWSM/ICWSM17/paper/view/15681

[36] O. H. Paruma-Pabón, F. A. González, J. Aponte, J. E. Camargo, and F. Restrepo-Calle, “Finding relationships between socio-technical aspects and personality traits by mining developer e-mails,” in *Proc. 9th Int. Workshop Cooperative Human Aspects Softw. Eng.*, 2016, pp. 8–14.

[37] F. Calefato, G. Iaffaldano, F. Lanubile, and B. Vasilescu, “On developers’ personality in large-scale distributed projects: The case of the apache ecosystem,” in *Proc. 13th Int. Conf. Global Softw. Eng.*, 2018, pp. 92–101. [Online]. Available: http://doi.acm.org/10.1145/3196369.3196372

[38] G. Gousios, M. Pinzger, and A. V. Deursen, “An exploratory study of the pull-based software development model,” in *Proc. 36th Int. Conf. Softw. Eng.*, 2014, pp. 345–355.

[39] D. M. Soares, M. L. de Lima Júnior, L. Murta, and A. Plastino, “Acceptance factors of pull requests in open-source projects,” in *Proc. 30th Annu. ACM Symp. Appl. Comput.*, 2015, pp. 1541–1546.

[40] Y. Yu, H. Wang, V. Filkov, P. Devanbu, and B. Vasilescu, “Wait for it: Determinants of pull request evaluation latency on GitHub,” in *Proc. IEEE/ACM 12th Work. Conf. Mining Softw. Repositories*, 2015, pp. 367–371.

[41] N. Munaiah, S. Kroh, C. Cabrey, and M. Nagappan, “Curating GitHub for engineered software projects,” *Empir. Softw. Eng.*, vol. 22, no. 6, pp. 3219–3253, 2017.

[42] Reporeaper. 2018. Accessed: Aug. 28, 2018. [Online]. Available: http://reporeapers.github.io/results/1.html

[43] G. Gousios, “The GHTorrent dataset and tool suite,” in *Proc. 10th Work. Conf. Mining Softw. Repositories*, 2013, pp. 233–236. [Online]. Available: http://dl.acm.org/citation.cfm?id=2487085.2487132

[44] GitHub API. 2018. Accessed: Aug. 30, 2018. [Online]. Available: https://developer.github.com/v3/

[45] IBM personality insights service. 2018. Accessed: Aug. 28, 2018. [Online]. Available: https://console.bluemix.net/docs/ services/personality-insights/science.html#science

[46] J. Pennington, R. Socher, and C. Manning, “GloVe: Global vectors for word representation,” in *Proc. Conf. Empir. Methods Natural Lang. Process.*, 2014, pp. 1532–1543.

[47] R. A. P. Junior and D. Inkpen, “Using cognitive computing to get insights on personality traits from Twitter messages,” in *Proc. Can. Conf. Artif. Intell.*, 2017, pp. 278–283.

[48] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2017. [Online]. Available: https://www.R-project.org/

[49] C. Tantithamthavorn and A. E. Hassan, “An experience report on defect modelling in practice: Pitfalls and challenges,” in *Proc. 40th Int. Conf. Softw. Eng.: Softw. Eng. Practice*, 2018, pp. 286–295.

[50] D. Bates, M. Mächler, B. Bolker, and S. Walker, “Fitting linear mixed-effects models using lme4,” *J. Statist. Softw.*, vol. 67, no. 1, pp. 1–48, 2015.

[51] K. Barto, *MuMIn: Multi-Model Inference*, R package version 1.43.6, 2019. [Online]. Available: https://CRAN.R-project.org/package=MuMIn

[52] Octoverse 2018. [Online]. Available: https://octoverse.github.com/

[53] Y. Wang, "Building the linkage between project managers' personality and success of software projects," in *Proc. 3rd Int. Symp. Empir. Softw. Eng. Meas.*, 2009, pp. 410–413.

[54] N. Baddoo and T. Hall, "De-motivators for software process improvement: An analysis of practitioners views," *J. Syst. Softw.*, vol. 66, no. 1, pp. 23–33, 2003.

[55] J. S. Karn, S. Syed-Abdullah, A. J. Cowling, and M. Holcombe, "A study into the effects of personality type and methodology on cohesion in software engineering teams," *Behav. Inf. Technol.*, vol. 26, no. 2, pp. 99–111, 2007.

[56] J. Chen, Y. Ren, and J. Riedl, "The effects of diversity on group productivity and member withdrawal in online volunteer groups," in *Proc. SIGCHI Conf. Human Factors Comput. Syst.*, 2010, pp. 821–830.

[57] R. McCrae, "Creativity, divergent thinking, and openness to experience," *J. Pers. Social Psychol.*, vol. 52, no. 6, pp. 1258–1265, 1987.

[58] J. M. George and J. Zhou, "When openness to experience and conscientiousness are related to creative behavior: An interactional approach," *J. Appl. Psychol.*, vol. 86, no. 3, 2001, Art. no. 513.

[59] J. F. Salgado, "Predicting job performance using FFM and non-FFM personality measures," *J. Occupational Organizational Psychol.*, vol. 76, no. 3, pp. 323–346, 2003.

[60] P. T. Costa Jr, R. R. McCrae, and D. A. Dye, "Facet scales for agreeableness and conscientiousness: A revision of the NEO personality inventory," *Pers. Indiv. Differ.*, vol. 12, no. 9, pp. 887–898, 1991.

[61] M. R. Barrick and M. K. Mount, "The big five personality dimensions and job performance: A meta-analysis," *Pers. Psychol.*, vol. 44, no. 1, pp. 1–26, 1991.

[62] R. R. McCrae and P. T. Costa, "Updating Norman's "adequacy taxonomy": Intelligence and personality dimensions in natural language and in questionnaires," *J. Pers. Soc. Psychol.*, vol. 49, no. 3, 1985, Art. no. 710.

[63] A. Mehrabian, "Analysis of the big-five personality factors in terms of the pad temperament model," *Aust. J. Psychol.*, vol. 48, no. 2, pp. 86–92, 1996.

[64] F. Calefato, F. Lanubile, and N. Novielli, "A preliminary analysis on the effects of propensity to trust in distributed software development," in *Proc. IEEE 12th Int. Conf. Global Softw. Eng.*, 2017, pp. 56–60.

[65] R. Akhtar, L. Boustani, D. Tsivrikos, and T. Chamorro-Premuzic, "The engageable personality: Personality and trait EI as predictors of work engagement," *Pers. Indiv. Differ.*, vol. 73, pp. 44–49, 2015.

[66] T. A. Judge, J. E. Bono, R. Ilies, and M. W. Gerhardt, "Personality and leadership: A qualitative and quantitative review," *J. Appl. Psychol.*, vol. 87, no. 4, 2002, Art. no. 765.

[67] G. A. Neuman, S. H. Wagner, and N. D. Christiansen, "The relationship between work-team personality composition and the job performance of teams," *Group Org. Manage.*, vol. 24, no. 1, pp. 28–45, 1999.

[68] L. S. Marcolino, A. X. Jiang, and M. Tambe, "Multi-agent team formation: Diversity beats strength?" in *Proc. 23rd Int. Joint Conf. Artif. Intell.*, 2013, pp. 279–285.

[69] E. Mannix and M. A. Neale, "What differences make a difference? The promise and reality of diverse teams in organizations," *Psychol. Sci. Public Interest*, vol. 6, no. 2, pp. 31–55, 2005.

[70] T. Walle and J. E. Hannay, "Personality and the nature of collaboration in pair programming," in *Proc. 3rd Int. Symp. Empir. Softw. Eng. Meas.*, 2009, pp. 203–213.

[71] G. J. Boyle, G. Matthews, and D. H. Saklofske, (Eds.) *SAGE Handbook of Personality Theory and Assessment*. vol. 1, Personality theories and models. Thousand Oaks, CA, USA: Sage, 2008.

Rahul N. Iyer received the MMath degree in computer science from the University of Waterloo, Waterloo, Canada. He is currently working as a Software Engineer.

S. Alex Yun received the HBSc degree in psychology and human biology from the University of Toronto, Toronto, Canada. He is currently working toward the MMath degree in computer science at the University of Waterloo, Waterloo, Canada under the supervision of Dr. Hoey. His work focuses on computational social science and affective computing.

Meiyappan Nagappan received the PhD degree in computer science from North Carolina State University, Raleigh, North Carolina. He is an assistant professor David R. Cheriton School of Computer Science, University of Waterloo. His research is centered around the use of large-scale software engineering (SE) data to address the concerns of the various stakeholders (e.g., developers, operators, and managers). He has published in various top SE venues such as the *IEEE Transactions on Software Engineering*, ICSE, FSE, the *Empirical Software Engineering*, the *Journal of Systems and Software*, ICSME and the *IEEE Software*. He is an associate editor of the *IEEE Transactions on Software Engineering*, the *Empirical Software Engineering*, and the *Journal of Systems and Software* and has served on the PC of several conferences like ICSE, MSR, and ICSME. You can find more at mei-nagappan.com.

Jesse Hoey received the PhD degree in computer science from the University of British Columbia, Vancouver, Canada, in 2004. He is an associate professor with the David R. Cheriton School of Computer Science, University of Waterloo, where he leads the Health Informatics Group and the Computational Health Informatics Laboratory (CHIL). He is also an adjunct scientist with the Toronto Rehabilitation Institute in Toronto, Canada. He has published more than 100 peer reviewed scientific papers in highly visible journals and conferences. He has received awards from Microsoft, the American Sociological Association, and the International Association for Pattern Recognition. He is an associate editor of the *IEEE Transactions on Affective Computing*. He is an area chair for the International Joint Conferences on Artificial Intelligence (IJCAI 2019), and was the program chair for the 10th EAI International Conference on Pervasive Computing Technologies for Healthcare, in 2016.