Personality and Project Success: Insights from a Large-Scale Study with Professionals

Xin Xia[*][†], David Lo[‡], Lingfeng Bao[*][√], Abhishek Sharma[‡], and Shanping Li[*]

[*]Zhejiang University, China; [†]University of British Columbia, Canada; [‡]Singapore Management University, Singapore;

{xxia, baolingfeng, shan}@zju.edu.cn, {davidlo, abhisheksh.2014}@smu.edu.sg

Abstract—A software project is typically completed as a result of a collective effort done by individuals of different personalities. Personality reflects differences among people in behaviour patterns, communication, cognition and emotion. It often impacts relationships and collaborative work, and software engineering teamwork is no exception. Some personalities are more likely to click while others to clash. A number of studies have investigated the relationship between personality and collaborative work success. However, most of them are done in a laboratory setting, do not involve professionals, or consider non software engineering tasks. Additionally, they only answer a limited set of questions, and many other questions remain open.

To enrich the existing body of work, we study professionals working on real software projects, answering a new set of research questions that assess linkages between project manager personality and team personality composition and project success. In particular, our study investigates 28 recently completed software projects, which contain a total of 346 professionals, in 2 large IT companies. We asked project members to do a DISC (Dominance, Influence, Steadiness, and Compliant) personality test, and correlated the test outcomes with project success scores measured in six different dimensions. The scores were given by managers of three office as part of their regular day-to-day work. Our results show that project teams with dominant managers, along with those with more influential members and less dominant members, have higher success scores. This work provides new insights to construct a personality matching strategy that can contribute to building an effective project team.

Keywords-Personality, Software Project, Project Success, Team Formation

I. INTRODUCTION

Personality is a set of individual differences including values, attitudes, personal memories, social relationships, habits and skills, which can be affected by the socio-cultural development of an individual [1]–[3]. A previous study shows that the compatibility of individual personalities in a team can affect the efficacy of collaboration significantly [4]. Different people have different personalities, and a team is likely to perform less effectively if personalities clash or are incompatible [2], [5]. On the other hand, when a team meshes well (e.g., a team formed with complementary personalities), its members can communicate with one another efficiently, enjoy the work environment, exhibit strong levels of commitment, and are more likely to achieve common goals (e.g., project success) [6].

A number of past studies have investigated the impact of personality to team work. Acuna et al. studied 105 students to investigate the impact of personality, team processes, task characteristics to software quality and satisfaction in software development projects [7]. Gorla and Lam ran Myers Briggs Type Indicator (MBTI) test on 92 IT professionals across 20 small project teams (of 3 to 7 members) to study the relationship between personality and self-perceived team performance [8]. Karn and Cowling also employed MBTI to study the interactions among members of student teams and found that teams which communicate or interact less among themselves seem to have lower quality of work [9]. Wang and Li employed Five Factor model (FFM) and Mohan Thite's model on managers of 50 software project teams, and found project managers' personalities (openness, neuroticism, agreeableness, conscientiousness, and extraversion) affect the success of software development projects [10].

Very recently, Lykourentzou et al. studied 14 teams, of 5 members each, which were asked to complete a simple collaborative advertisement creation task [11]. They employed DISC (Dominant, Influence[1], Steadiness, and Compliance) personality test and found that balanced teams (i.e., teams with one leader) produce better work product. DISC test is well founded on existing team building theories [12], and has not been used to assess software engineers before. Personality test models used in previous software engineering studies treat people as individuals in isolation. Moreover, compared to other tests (e.g., Five Factor Model), DISC has these advantages: frequently used by businesses [13], easy to interpret [14], [15], shown to be a predictor of success for employee retention, job success, sales management, and persuading patients to accept treatment plans [16], and proven to be reliable and consistent [17].

Our study is inspired and extends the above mentioned body of work. In particular, we want to: (1) investigate professionals instead of students, (2) analyze software projects instead of other collaborative work, (3) consider teams of various sizes instead of only those of small sizes, (4) use a more objective measure of team performance instead of self assessment, (5) employ a suitable personality test well-founded on latest theories on team work, (6) include substantially more participants than prior work, and (7) investigate new research questions to gain deeper insights on personality and team success. To achieve these goals, we conducted an empirical study in two large sister IT companies named Hengtian [18]

√Corresponding author.

[1]It is also referred to as Inducement

and Insigma Global Service [19].

We invited a total of 346 participants across 28 recently completed project teams (of size 7 to 25 members) to join our study. All the participants were asked to complete a DISC personality test which is widely used as an HR assessment tool when hiring new employees [13], [20]. DISC personality test outputs four distinct personality types, i.e., dominance, influence, steadiness, and compliance. The dominant and influential types are leader types, and the steady and compliant types are non-leader types. In these two companies, when a project gets completed, managers in 3 different offices (i.e., account management office (AMO), project management office (PMO), and quality assurance office (QAO)) would be required to evaluate the success of the project based on 6 dimensions: schedule, effort, risk, issue, quality, and customer satisfaction. Based on the personality and project success data, we investigate several research questions that assess linkages between project manager personality and team personality composition and project success. The answers to these research questions can help in understanding how to form an effective software project team, and provide guidelines on how to select team members by considering both their technical skills and personalities.

The main contributions of the paper are as follows:

1) We conducted a large-scale empirical study on 346 participants across 28 project teams to study the relationship between personality and software project success. Our study differentiates itself from existing work by its focus on professionals working on real software development projects, analysis of diverse teams of various sizes, consideration of a more objective assessment of project success, deployment of a personality test well-founded on group work theory, and inclusion of a substantially larger number of participants.

2) Our study sheds new insights on the relationship between personality and project success. Project teams with a D-type (i.e., dominant) project manager, with less D-type members, or with more I-type (i.e., influential) members achieve significantly higher project success scores. This highlights the potential of employing a simple personality matching strategy to help build a more effective project team.

Paper organization. The remainder of this paper is organized as follows. Section II elaborates the case study setup. Section III presents our case study results. Section IV discusses the implications and the threats to validity. Section V briefly reviews related work. Section VI draws the conclusions and mentions future work.

II. Case Study Setup

In this section, we first present the personality test used in our study. Next, we describe project team and participant selection process. Then, we elaborate on how project success is assessed. Finally, we present open questions that we asked participants at the end of the study.

A. Personality Test

Since our study focuses on investigating relationship between personality and project/team success, an ideal personality test tool should be: (1) easy to deploy online, (2) allow the extraction of individual team member personality, (3) provide information on the effect of personality on teamwork effectiveness, instead of only providing individual personality assessment in isolation. There are a number of personality tests, e.g., Costa and McCrae's NEO-PI-R five dimension analysis (i.e., Five-Factor Model) [21], Holland's 6 personality types [22], and Eysenck's supertraits [23]. However, these personality tests only treat people as individuals in isolation (rather than characterizing their interactions and roles in a group), and are not well-founded on team formation theories.

In this paper, following the previous study by Lykourentzou et al. [11], we use DISC personality test which explicitly describes how individual personality interacts at the team level and the roles that they would play inside the team. Notice that the DISC personality test outputs only four distinct team member types, while other personality tests (e.g., Costa and McCrae's NEO-PI-R five dimension analysis) output more; this can help reduce the complexity of our analysis, and help us to get statistically significant results with lower number of teams [11]. Also, the four personalities outputted by the DISC test can cover personality traits that matter for team performance – which is not the case for some other tests [11]. Moreover, the DISC personality test is widely used as an HR assessment tool when hiring new employees [13], [20]. DISC centers on four different behavioral types named dominance, influence, steadiness, and compliance. The characteristics of these four types are summarized below:

- D-type individuals (Leader type): people of this type exhibit high dominance, are task-oriented, and place emphasis on accomplishing results and the bottom line.
- I-type individuals (Leader type): people of this type exhibit high influence, are socio-emotionally oriented, and place emphasis on persuading others, openness, and relationships.
- S-type individuals (Non-leader type): people of this type exhibit high steadiness, are socio-emotionally oriented, and place emphasis on cooperation, sincerity, and dependability.
- C-type individuals (Non-leader type): people of this type exhibit high compliance, are task-oriented, and place emphasis on quality, accuracy, expertise, and competency.

In this paper, we use a 40-item DISC personality test from a popular Chinese psychology test website[2], which is widely used in China.

B. Project Team and Participant Selection

We selected projects from two large sister IT companies in China, namely Insigma Global Service (IGS), and Hengtian. IGS is an outsourcing company which has more than 500 employees, and it mainly does outsourcing projects for Chinese vendors. Hengtian is also an outsourcing company which has more than 2,000 employees, and it mainly does

[2]http://www.weiceyan.com/s/64.htm

outsourcing projects for US and European corporations. Since some developers may leave the company after the completion of their projects, to reduce this risk, we carefully selected projects which were completed between March 1, 2016 and May 1, 2016 – close to the start of our study.

We contacted the managers of the project management office (PMO) in the two companies, and they provided us a list of projects which were completed in the last two months prior to our study. For every project, they also provided us the contact information of the project members. The list contains a total of 35 projects and 448 professionals work on these projects. Next, we dropped emails to each of the 448 professionals to invite them to take our personality test. A total of 362 professionals took the test. Since our study requires that all members of a project team complete the personality test, and not all members participated for some teams, we removed 16 respondents and their teams from our study. We were left with 346 professionals who were members of 28 teams.

The 28 projects are diverse in nature. Among them, 11, 8, 5, 2, and 2 projects used Java, .Net, C/C++, Python, and Matlab as the main program language respectively. Seven projects developed Android apps, 10 projects developed desktop applications, and 11 projects developed web applications. Moreover, 15 projects were related to financial systems, 8 projects were related to e-commerce systems, 2 projects were related to flash memory systems, 2 projects were related to cloud computing systems, and 1 project was related to programming language translation (i.e., translate COBOL programs to Java programs). The size of these 28 projects vary from 0.2M LOC to 15M LOC, and these projects have existed for one to five years. The diversity of these 28 projects can help to generalize our findings. The average number of developers is 11.35, with a minimum and maximum number of developers of 7 and 25 respectively. On average across the 28 project teams, the the number of developers of D, I, S, and C types are 2.28, 2.14, 4.25, and 2.68, respectively.

C. Project Success Evaluation

In IGS and Hengtian, whenever a project has been completed, the companies would ask managers in 3 different offices to evaluate a number of dimensions that the companies use to characterize project success. The 3 offices are the account management office (AMO), project management office (PMO), and quality assurance office (QAO). The managers are required to complete a survey form to evaluate a project considering 6 dimensions, i.e., schedule, effort, risk, issue, quality, and customer satisfaction. Notice that the managers in AMO, PMO, and QAO would complete the survey form independently, and at the end of this step, 3 pieces of survey responses are recorded. Next, the managers in these 3 offices would participate in a roundtable meeting to discuss the detail problems met during the project development and maintenance and learnt experience. In the roundtable meeting, a final score is collectively decided for each project success dimension. In the following paragraphs, we elaborate the three offices and success dimensions considered.

1) Three Offices: The 3 offices evaluate the success of a project from their own views and different emphasis points (i.e., AMO mainly focuses on cost control, PMO mainly focuses on in-time delivery, and QAO mainly focuses on software quality):

1) **AMO:** The main responsibility of AMO is to control the cost for software development and maintenance. A manager in AMO gets payments from clients, and disburses salaries of project members. Thus, the account manager is very sensitive to project cost. One account manager typically takes care of several project teams simultaneously. He/she communicates with project members frequently, and reads status reports carefully to ensure that everything proceeds smoothly.

2) **PMO:** The main responsibility of PMO is to monitor software development and maintenance process to avoid software release delay. The project manager of a team needs to report to PMO every week, and all the team members are required to complete and submit monthly reports. PMO has the right to check all software artifacts produced by a project team.

3) **QAO:** The main responsibility of QAO is to ensure the quality of software delivery. The QA leader in a project team would report to the QAO every week, and QAO typically checks for the number of bugs, the number of bugs which are fixed, and the number of crashing bugs in a project every month.

Note that managers in AMO, PMO and QAO are different from managers of project teams. To avoid potential judgement bias, managers of project teams are not involved in project success evaluation process.

2) Success Dimensions: The two sister companies ask managers in AMO, PMO, and QAO to evaluate the degree a project is successful considering each of the 6 dimensions. For each dimension, the managers are required to assign a score from 1 to 3 – 3 means that there is no problem in the dimension (normal), 2 means that there are some minor problems but they are under control (warning), and 1 means that there are serious/critical problems (dangerous). If a score of 2 or 1 is given, a manager needs to give a justification. The definitions of the 6 dimensions are as follows:

Schedule: In this dimension, a manager is asked to judge whether a project's schedule is still under control. Notice that various factors affect project schedule. For example, unclear or frequently changed requirements can adversely affect project schedule. If actual project development schedule diverges widely from the initial plan, the project has a high chance to fail. The manager would check whether there are some delays in the project releases[3] and if such a delay happens, the manager needs to investigate why the project is delayed.

Effort: Some project teams may need to work overtime to ensure that projects can be released on time. However, working

---

[3]A project may have several intermediary releases and a final release. The managers in these 3 offices would consider all of the intermediary releases and the final release.

for too many extra hours may mean low efficiency, poor project planning and execution, and even translate to low project quality. For the effort dimension, the managers in the 3 offices analyze the effective working hours of a project team, and evaluate whether the team suffers from serious overtime work problem. In IGC and Hengtian, developers are required to report their effective working hours once a week. Normally, a developer is expected to work 40 hours per week. If the average working time of a developer is more than 50 hours per week, the managers would give a score of 2 (warning), and if the average working time of a developer is more than 60 hours per week, the managers would give a score of 1 (dangerous).

External Risk: In this dimension, the managers in the 3 offices mainly analyze the external risks that a project team meets during the development process. Some risks can be under control, while other risks are out of control, or hard to control. It is very important for the managers to identify the uncontrolled risks. If a risk cannot be controlled, causes the loss of a customer, or causes a disruption in the normal functioning of a team, the managers would give a score of 1. And if the risk is still under control, but causes some bad impact to the project (e.g., some delays in project delivery), the managers would give a score of 2.

Internal Issue: Different from external risk, internal issue refers to problems appearing inside the project team. For example, team members may not communicate with one another effectively. Similar to the external risk dimension, the managers assign scores to this dimension based on the impact and controllability of the internal issues.

Software Quality: The quality of a project is an important project success dimension. If many bugs exist when a product is shipped, it would cause bad user experience and even make critical functionalities unavailable to users. For this dimension, the managers would analyze a project's bug tracking system (e.g., Bugzilla or JIRA), and also the code quality report outputted by Sonar [24]. Specially, they would check for the total number of bugs, the number of crash and blocking bugs, the number of post-release bugs, bug distribution, code coverage, unit test success rate, etc.

Customer Satisfaction: Customer satisfaction is important for a sustainable business. If customers are not satisfied with the end product or service provided by a project team, they may engage the service of another company, which will cause a loss of income. In IGS and Hengtian, when a project gets completed, the companies would collect feedback from customers [25]. Specially, customers are asked to rate whether they are satisfied with the project team in various aspects, e.g., work attitude, communication quality, business knowledge, technical skills, cost-effectiveness, etc. Customers are also asked to provide reasons why they are satisfied or unsatisfied. For the customer satisfaction dimension, managers in the 3 offices would analyze the customer satisfaction feedback report as well as the monthly reports from the project team to assign a suitable score.

TABLE I
WEIGHTS OF THE 6 DIMENSIONS.

Weight	Value (mean±std)	Weight	Value (mean±std)
Schedule	19.3% ± 1.8%	Internal Issue	9.4% ± 2.1%
Effort	17.6% ± 1.2%	Software Quality	21.4% ± 1.3%
External Risk	11.3% ± 1.1%	Customer Satisfaction	21.0% ± 2.2%

3) Overall Project Success Score: After the managers in the 3 offices completed their own independent assessments of the project and recorded their assessments through the survey form, they would participate in a roundtable meeting to discuss the final scores for each of the 6 dimensions. We denote the final score for each of the 6 dimensions as S_{sche}, S_{effort}, S_{risk}, S_{issue}, S_{qua}, and S_{cust}, respectively.

In this study, we also compute an overall project success score F which aggregates all the 6 dimensions using the following equation:

$$F = \alpha_1 \times S_{sche} + \alpha_2 \times S_{effort} + \alpha_3 \times S_{risk} +$$
$$\alpha_4 \times S_{issue} + \alpha_5 \times S_{qua} + \alpha_6 \times S_{cust} \quad (1)$$

In the above equation, α_1 to α_6 denote the weights assigned to each dimension, and $\alpha_1 + \alpha_2 + \cdots + \alpha_6 = 1$. To determine reasonable α_1 to α_6 values, we dropped emails to managers in the 3 offices. In total, we sent 72 emails, and we received 32 responses. We assigned the weights of the 6 dimensions by averaging the responses that we received from the managers. Table I presents the average weights of the 6 dimensions. Notice that the weights of external risks and internal issues are around 10%. For the other four dimensions, the weights are around 20%. On average across the 28 projects, the average project success score is 2.75, which means that most of the projects are successful.

D. End-of-Study Survey

At the end of our study, we also dropped emails to the 346 participants and the managers in the 3 offices to ask them some open questions. The questions were listed below:

1) Do you think personalities contribute towards the success of software projects? And why?
2) Can you describe the personality of your project manager? Is he/she easy to get along with, or has a strong opinion? Does his/her personality play a role in influencing the software development process?
3) Can you describe the personalities of some team members you are impressed with or disliked? Do their personalities play a role in influencing the software development process?

The participants can speak freely on these 3 questions as we promised to keep their data confidential and only publish anonymized responses. In total, we received 264 responses. Then, we analyzed these responses to find the reasons why some personalities influence project success.

III. CASE STUDY RESULTS

In this paper, we formalize our study in the following research questions:

RQ1: Does project manager personality influence project success?

Motivation. In a typical project team, the project manager takes the overall responsibility for the successful planning and execution of a project, and he/she needs to design a good project plan, organize the team structure, monitor the execution of the project plan, and control the risks and schedule to keep the project on track [26]. Considering the importance of the project manager, his/her personality may affect the success of the project. In this research question, we plan to investigate which personality type is more correlated to project success.

Results. In our collected data, among the 28 projects, 10, 3, 10, and 5 project managers are of the personality type D, I, S, and C, respectively. We first investigate whether the project manager personality would affect the project success score. Since the distribution of project success score is not normally distributed as shown by the results of the Shapiro-Wilk test [27] (i.e., p-value is less than 0.05), we apply Kruskal-Wallis test [28]. Kruskal-Wallis test is a non-parameter and distribution free test, which assesses for significant differences on a continuous dependent variable (in our case, project success score) by a categorical independent variable (in our case, the personality of project managers). The results show that the p-value of the Kruskal-Wallis test is 0.03, which indicates that **the project manager personality affects the project success score.**

Next, we investigate whether projects of which their project managers are of a specific personality type would achieve a statistically significantly and substantially higher mean project success score than projects of which the project managers are of the other personality types. Specifically, for each personality type, we divide the 28 projects into two groups, one group for projects of which the managers are of a specific personality (e.g., D), and the other group for projects whose managers are of the other personalities (e.g., I, S, and C). We apply the Wilcoxon Rank Sum test [29] to measure whether the improvement is significant, and we compute Cliff's delta [30][4], which is a non-parametric effect size measure that quantifies the amount of difference between the two groups. Since we run the test multiple times (i.e., we compare different pairs of groups), we use Bonferroni correction [31] to counteract the results of multiple comparisons. Here, we choose Wilcoxon Rank Sum test and Cliff's delta effect size since the distribution of project success score is not normally distributed.

Table II presents the adjusted p-values and Cliff's Deltas for the comparison of different project manager's personalities and their corresponding project success scores. The p-values are adjusted by Bonferroni correction. **We find that projects with managers of D-type personality achieve higher project success scores, and the difference is statistically significant at the confidence level of 95%. Moreover, the Cliff's Delta is 0.75, which corresponds to a large effect size.** For projects with managers of each of the 3 other personality types, there is no statistically significant difference between their success scores and the success scores of other projects.

---

[4]Cliff defines a delta of less than 0.147, between 0.147 to 0.33, between 0.33 and 0.474, and above 0.474 as negligible, small, medium, and large effect size respectively.

TABLE II
ADJUSTED P-VALUES AND CLIFF'S DELTAS FOR THE COMPARISON OF DIFFERENT PROJECT MANAGER'S PERSONALITIES AND THEIR CORRESPONDING PROJECT SUCCESS SCORES.

Comparison	p-value	Cliff's Delta
D vs. I, S, C	**0.0020**	**0.75 (Large)**
I vs. D, S, C	1.0000	-0.61 (Negligible)
S vs. D, I, C	1.0000	-0.32 (Medium)
C vs. D, I, S	1.0000	0.30 (Small)

To delve further as to the reason of this observation, we also analyze the responses that we collected from members of projects with D-type managers. The following are some of them:

👍 *"My project manager is a bit push, and sometimes we are afraid of him. He would check our status every two or three days, and criticize us if we do not complete the tasks. But anyhow we can always release the project on time, and our team won a lot of awards from the company. Although I do not like him personally, I should admit he is a good project manager."*

👍 *"I have been in my PM's team for a long time, and release at least 5 projects until now. She is task-oriented, and has strong opinion. We all believe her since she has the magic to make a successful project."*

👍 *"Sometimes the customers would propose some unreasonable requirements, my project manager would try to persuade the customers but not transfer the work to us directly. My project manager knows clearly what is in the scope of the project, and what is out of the scope. And he would control the things in scope strictly, but he would also help to fight for our rights."*

👎 *"Everyday my PM would push us, which makes me feel nervous. I am not happy to work with my PM."*

👎 *"Our project manager is not in good temper. If the schedule is delayed even a little, he will rant to us. He is overly strict."*

The following are some comments that we collected from project teams with managers of other personality types:

👎 **S-type:** *"Our project manager is our best friend, and he always listens to our suggestions patiently. However, it seems that he does not have many opinions, and he wants to make everyone happy, but it may make the things worse. As a result, our project contains a lot of bugs when we release it to the production environment."*

👎 **I-type:** *"It seems our project manager is too optimistic when designing project plans, and he always considers things in the best case, which makes us very passive especially when the deadline is approaching."*

👎 **C-type:** *"To be honest, my project manager think too much on the detailed things, and he always feels pessimistic to the success of our project, which affects our confidence to complete our project."*

👍 **I-type:** *"I like my PM not because she is a good leader, but she is our best friend. We work together and discuss together, and she always encourages us to solve difficult problems."*

👍 **S-type:** *"Our PM can always find some important problems which we may ignore in a project. He is more like a technical leader than a PM, and we all ask him for help when we meet technical problems."*

From the above comments, we notice a good project manager is typically task-oriented, has a strong opinion, knows what is the objective of the project clearly, and can control the various steps in a software development process strictly. These are characteristics of people with D-type personality. However, D-type managers may also destroy the harmonious relationship in a team. We can also note that project managers who are very nice to everyone, feel too optimistic or pessimistic to a project, may negatively impact project success, however they may build good relationships inside a team.

RQ2: Does team member personality influence project success?

Motivation. Lykourentzou et al. find that teams with only one D-type and one I-type member produce better work product than a team which has more than one D-type or I-type team members [32]. In this research question, we would like to investigate whether the increase of people with leader type personalities would impact project success score. The answer of this research question would shed light on how to organize

TABLE III
SPEARMAN'S RHO AND ADJUSTED P-VALUES FOR CORRELATIONS
BETWEEN PERCENTAGE OF MEMBERS OF DIFFERENT PERSONALITIES AND
PROJECT SUCCESS SCORE.

Type(%)	Spearman's rho	p-value
D%	-0.61 (High)	0.0005
I%	0.74 (High)	$6.809e^{-6}$
S%	0.26	0.1875
C%	-0.35	0.0684
L%	0.22	0.2535
NL%	-0.22	0.2535

TABLE IV
ADJUSTED P-VALUES AND CLIFF' DELTAS FOR DIFFERENT PERCENTAGES
OF MEMBERS OF D-TYPE AND I-TYPE PERSONALITY.

D%	P-value	Cliff's Delta	I%	P-value	Cliff's Delta
5%	0.8436	0.67 (large)	5%	1.0000	-0.16 (small)
10%	0.8424	0.68 (negligible)	10%	0.0042	-0.77 (large)
15%	0.0024	0.70 (large)	15%	0.0060	-0.68 (large)
20%	0.0002	0.86 (large)	20%	0.0003	-0.86 (large)
25%	0.0378	0.58 (large)	25%	0.0060	-0.73 (large)
30%	0.0036	0.84 (large)	30%	0.0218	-0.72 (large)

a better project team, and how to select team members based on their personalities.

Results. To answer this research question, for each project team, we compute the percentages of members of different personalities in the team. Here, we exclude the project manager. In total, we consider 6 independent variables, i.e., percentages of D-type (D%), I-type (I%), S-type (S%), C-type (C%), leader-type (L% = D%+I%), and non-leader type (NL%=S%+C%) personalities, respectively. Here, we choose Spearman's rho [33] to measure the statistical significance and correlation level of the project success score to each of the 6 variables since (1) the distributions of project success score and the percentages of various personality types are not normally distributed as shown by the results of the Shapiro-Wilk test [27], and (2) both the dependent variable (in our case, project success score) and independent variable (in our case, the percentages of various personality types) are continuous. The Spearman's rho ranges from -1 to 1, where -1 and 1 correspond to a perfect negative and positive monotonic correlation respectively, and 0 means that the variables are independent[5]. Similar to RQ1, we also use Bonferroni correction to adjust the p-values outputted by Spearman's rho test.

Table III presents the Spearman's rho and adjusted p-values for correlations between different personalities and project success score. We notice among the 6 variables, only the percentages of D-type personality and I-type personality are highly correlated to the project success score, and the correlation is statistically significant at the confidence level of 99%. **We notice that the percentage of D-type members is negatively correlated to the project success score, i.e., the more D-type members are in a project team, the lower is the project success score. On the other hand, the percentage of I-type members is positively correlated to the project score, i.e., the more I-type members are in a project team, the higher is the project success score.**

Next, to understand what is a good percentage of D-type and I-type members in a project team, we also investigate the following two sets of hypotheses:

Hypothesis Set 1. *Project teams with percentage of D-type members of at most p% achieve a higher mean project success score than that of other teams.*

Hypothesis Set 2. *Project teams with percentage of I-type members of at least p% achieve a higher mean project success score than that of other teams.*

---

[5]Spearman's rho of < 0, [0.0, 0.1], (0.1, 0.3], (0.3, 0.5], (0.5, 0.7], (0.7, 0.9], (0.9, 1] is considered as no, small, moderate, high, very high, and perfect correlation, respectively.

To evaluate each hypothesis in the two sets, we divide the project teams into two groups: one group contains teams with percentage of D-type (I-type) members of at most (at least) p%, and the other group contains the other teams. We use Wilcoxon Rank Sum test to check whether the difference in the mean project success score is significant, and Cliff's delta to assess if the difference is substantial. We investigate the following values of p: 5, 10, 15, 20, 25, and 30.

Table IV presents the p-values and Cliff's Delta for the different values of p. The p-values are adjusted by Bonferroni correction. We notice when p is between 15 to 30, the first hypothesis can be accepted. Moreover, we also notice when p is larger than 10, the second hypothesis is accepted. **Thus, in practice, we recommend that a project team should have at most 15% members of D-type personality and at least 10% members of I-type personality.**

We assign a team into one of the following two groups: D-type dominant team and I-type dominant team. The assignment is made according to the proportion of D-type and I-type personality team members in a team. Teams who have more D-type than I-type personality members are considered D-type dominant, while those who have more I-type than D-type personality members are considered I-type dominant. In our study, we did not find a team who has the same number of D-type and I-type personality members. The following are some of the comments related to team member personality that we have received:

☞ **I-type dominant:** *"I like our team since we have many enthusiastic developers. They often listen to other people's suggestions, and always help to create good solutions to any problem we meet during the project development."*

☞ **D-type dominant:** *"A good project team should have members of different personalities, but, one key point is, not many dominant team members."*

☞ **D-type dominant:** *"Software development is a teamwork, and we need to make a trade off in a lot of things. However, if a team has many dominant developers, it is hard to make the final decision. Too many dominant developers would result in many internal fights."*

☞ **D-type dominant:** *"It seems that our project team has some developers who do not listen to others, and they have a strong opinion. Our project team has a big communication problem. Every time a customer proposes a new requirement, we would implement it in our own way. Finally, the customer is not satisfied with our project."*

☞ **D-type dominant:** *"Three developers in our project team are too aggressive, and they always issue orders to other developers, which make us not happy to work with them."*

☞ **D-type dominant:** *"Our project team has too many "leaders". Every "leader" wants to persuade others, and all of them are strong-willed, and hard to accept other developers' suggestions."*

RQ3: Does project manager personality influence a project success dimension?

Motivation. In the earlier RQs, we consider the overall project success score which is aggregated over the 6 dimensions. In this research question, we investigate whether project manager personality influences *each* project success dimension.

Results. Similar to RQ1, for each personality type, we divide

the 28 projects into two groups, one group for projects of which the managers are of a specific personality (e.g., D), and the other group for projects whose managers are of the other personalities (e.g., I, S, and C). We apply the Wilcoxon Rank Sum test [29] to assess whether the difference is statistically significant, and we compute Cliff's Delta to measure whether the difference is substantially large. Since we run the Wilcoxon Rank Sum test multiple times, we use Bonferroni correction [31] to counteract the effect of multiple comparisons. Here, we choose Wilcoxon Rank Sum test and Cliffs delta effect size, since the distributions of the scores in each dimension are not normally distributed as shown by the results of the Shapiro-Wilk test.

Table V presents the adjusted p-values and Cliff's Deltas for the comparison of different project manager's personalities and their corresponding 6 project success dimension scores. **We find that projects with managers of D-type personality achieve higher scores in the following dimensions: internal issue, software quality, and customer satisfaction.** The differences are statistically significant at the confidence level of 95%. The differences are also substantial, with the effect sizes varying from medium (for software quality and customer satisfaction dimensions) to large (for internal issue dimension). The results for the other personalities are not statistically significant. The results for D-type managers may be attributed to the fact that they are often firm, forceful, task-oriented, and have high standard in doing things. The following are some comments that we received from members of projects with D-type managers:

☞ *"Our PM has the magic to know every potential risk that would happen in advance. He would talk to everyone once a week to ask whether we meet some communication or other internal problems during software development."*
☞ *"It is interesting that our project manager will read every bug report, and analyze the impact of it carefully. Before the release, he seems to know the status of all the bugs in Bugzilla, and pushes us to fix most high-impact bugs."*
☞ *"My PM is an expert in requirement analysis, and he will check with the customers a lot of times to confirm the requirements. Thus, the customers are always satisfied with what we do finally since we strictly follow the requirements."*

RQ4: Does team member personality influence a project success dimension?

Motivation. In this research question, we augment our finding for RQ2, by investigating the relationship between team member personality and each of the six project success dimensions. In particular, we investigate whether the proportion of team members of a particular personality is correlated to each of the team's dimension scores. The answer to this research question can help us better understand the advantage or disadvantage of having more or fewer members of a specific personality to a dimension of success. Answer to this research question augments findings of RQ2 which only consider the overall success score.

Results. Similar to RQ2, in this research question, we also consider 6 independent variables, i.e., D%, I%, S%, C%, L%, and NL%. Next, following RQ3, since the distributions of the scores in each dimension are not normally distributed, we compute Spearman's rho [33] to measure the correlation level of each of the 6 variables and project success measured in

terms of one of the six dimensions. We also use Bonferroni correction to adjust the p-values outputted by Spearman's rho test. Table VI presents the adjusted p-values and Spearman's rhos for the correlations between each of the 6 dimension scores and the 6 variables. We describe the findings for each dimension below:

Schedule. The percentage of C-type members is negatively correlated to the schedule score, and the correlation level is moderate. This suggests that **increasing the percentage of C-type members is likely to decrease the schedule score**. On the other hand, **the percentage of members with leader type personality is positively correlated to the schedule score**, and the correlation level is moderate. Also, notice that the percentage of members with non-leader type personality is negatively correlated to the schedule score. This may be attributed to the fact that C-type members may sometimes consider too many things and are indecisive, which may delay project schedule. On the other hand, people with leader-type personality tend to be self-motivated, follow project plan strictly, and either force or influence other team members to execute the project plan. The following are some relevant comments that we collected:

☞ *"I believe some of our team members consider too many things and are indecisive, even for a simple requirement, which causes the delay of our project."*
☞ *"One thing that causes the delay of our project is that many team members are a bit lazy and not self-motivated, they need to be pushed to do things. "*

Effort. **There is no statistically significant correlation between the variables and the effort score**. This result suggests that differences in team member personality play a non-major role in determining the amount of effort invested in a project. Other factors such as tight project schedule and frequent requirement change may play a more major role.

External Risk. **There is no statistically significant correlation between the variables and the external risk score**. Not all of the 28 project may have encountered significant external risk and thus the data may not have been sufficient to draw conclusion.

Internal Issue. The percentage of I-type members is positively correlated to the internal issue score, and the correlation level is moderate. This indicates **an increase in the percentage of I-type members is likely to increase the internal issue score** (i.e., internal issues are more controlled and have less impact). On the other hand, **the percentages of D-type and C-type members are negatively correlated to the internal issue score**, and the correlation levels are moderate. From our collected data, most of the internal issues are related to communication problems. I-type people are good communicators and they are outgoing and optimistic. However, D-type people can be aggressive sometimes, and they are less inclined to listen to others. C-type people are typically reserved and may not communicate effectively with others. The following are some relevant comments that we collected:

☞ *"We have a good development environment, and everyone tend to communicate with others on the problem he/she meets. Thus, we do not have any communication problem."*
☞ *"Some developers seem to be silent, and they do not express their opinion. They try to solve problems in a quiet way, and no one knows what they are*

TABLE V
ADJUSTED P-VALUES AND CLIFF'S DELTA FOR THE COMPARISON OF DIFFERENT PROJECT MANAGER'S PERSONALITIES AND THEIR CORRESPONDING 6
DIMENSION SCORES. STATISTICALLY SIGNIFICANT RESULTS AT THE CONFIDENCE LEVEL OF 95% ARE HIGHLIGHTED IN BOLD. MEDI. = MEDIUM

Type	Schedule		Effort		External Risk		Internal Issue		Software Quality		Customer Satisfaction	
	P-value	Delta	P-value	Delta	P-value	Delta	P-value	Delta	P-value	Delta	P-value	Delta
D vs. Other	1.0000	0.18	0.1904	0.33	1.0000	0.13	**0.0450**	**0.5 (large)**	**0.0275**	**0.39 (medi.)**	**0.0467**	**0.33 (medi.)**
I vs. Other	1.0000	-0.13	0.2756	-0.48	1.0000	0.32	1.0000	0.00	1.0000	-0.09	0.2132	-0.51
S vs. Other	1.0000	-0.17	0.7256	-0.41	1.0000	0.22	1.0000	-0.15	1.0000	-0.11	0.3076	-0.5
C vs. Other	1.0000	0.02	0.8928	-0.26	1.0000	-0.14	1.0000	-0.16	0.2200	-0.43	1.0000	0.02

TABLE VI
ADJUSTED P-VALUES AND SPEARMAN'S RHO FOR CORRELATIONS BETWEEN PERCENTAGE OF MEMBERS OF DIFFERENT PERSONALITIES AND EACH OF
THE SIX DIMENSION SCORES. STATISTICALLY SIGNIFICANT RESULTS AT THE CONFIDENCE LEVEL OF 95% ARE HIGHLIGHTED IN BOLD. MO.=MODERATE

Type(%)	Schedule		Effort		External Risk		Internal Issue		Software Quality		Customer Satis.	
	P-value	rho	P-value	rho	P-value	rho	P-value	rho	P-value	rho	P-value	rho
D%	0.6814	-0.08	0.3270	-0.19	0.3413	-0.19	**0.0088**	**-0.49 (mo.)**	**0.0017**	**-0.56 (high)**	0.0919	-0.32
I%	0.1177	0.30	0.3444	0.19	0.4105	0.16	**0.0037**	**0.53 (high)**	**0.0053**	**0.51 (high)**	**0.0255**	**0.42 (mo.)**
S%	0.0924	0.32	0.8682	0.03	0.6194	0.10	0.1212	0.30	0.6410	0.09	0.5109	0.13
C%	**0.0033**	**-0.53 (high)**	0.7597	-0.06	0.5179	-0.13	**0.0500**	**-0.37 (mo.)**	0.5685	-0.11	0.8486	-0.04
L%	**0.0188**	**0.44 (mo.)**	0.6067	0.10	0.9006	0.02	0.7342	0.07	0.8348	-0.04	0.9781	0.01
NL%	**0.0188**	**-0.44 (mo.)**	0.6067	-0.10	0.9006	-0.02	0.7342	-0.07	0.8348	0.04	0.9781	-0.01

experienced during software development. "

☞ *"I do not like some of our team members since they are not friendly, and they are not interested in listening to what others say. It makes none of us want to talk to others, and our team suffers from a big communication problem."*

Software Quality. The percentage of D-type members is negatively correlated to the software quality score, and the correlation level is high. This indicates that **an increase in the percentage of D-type members is likely to decrease the software quality score**. On the other hand, **the percentage of I-type members is positively correlated to the software quality score**, and the correlation level is high. Bug fixing is a tedious and time-consuming task, and D-type developers may not have the necessary patience to carefully fix bugs. Moreover, for some bugs, there are different solutions, and if a project team has too many D-type members, each may have their own conflicting opinion which may delay or even hamper the bug fixing effort. Different from D-type people, I-type people are more inclined to consider different opinions. The following are some relevant comments that we collected:

☞ *"At least five of developers have very strong opinions. When the testers report a bug, their will argue with the testers that it is not a bug. Even for a real bug, they just fix it in their own way, and refuse to adopt other solutions."*

☞ *"Some developers in our team are open-minded, and we always discuss solutions on some difficult bugs. Moreover, they are so nice and they always encourage us to be not afraid of bugs."*

Customer Satisfaction. The percentage of I-type members is positively correlated to the customer satisfaction score, and the correlation level is moderate. This indicates that **an increase in the proportion of I-type members is likely to increase the customer satisfaction score**. As mentioned earlier, I-type people tend to be good at communication, and they are outgoing and easy to get along with. Thus, they are more likely to make the customer satisfied. One respondent pointed out: *"We have many emails with the customers every day. I notice some developers in our team have good communication skills when they reply emails to customers. They are patient and polite, and customers always praise them."*

RQ3 vs. RQ4. Comparing results of RQ3 and RQ4, we find that **D-type managers perform well in internal issue and software quality dimensions, while an increase in the proportion of D-type members is likely to decrease the scores in these two dimensions**. This may be attributed to the fact that managers and team members play different roles in project teams. The main responsibilities of managers are to design the project schedule, manage the whole team, and monitor the process to ensure everything is smooth. On the other hand, the main responsibilities of team members include development, maintenance, testing, debugging, and bug fixing, and most of these are technical. The difference in responsibilities seems to make D-type people contribute differently to the two success dimensions when they are in different roles.

IV. DISCUSSION

A. Interaction Effect Analysis

In this paper, we consider the impact of multiple independent variables (i.e., project manager personality, percentages of D-type (D%), I-type (I%), S-type (S%), C-type (C%), leader-type (L% = D%+I%), and non-leader type (NL%=S%+C%) personalities) to project success score. Here, we would like to investigate the interaction effects of these independent variables to the project success score. As shown in the previous RQs, the distribution of project success score is not normally distributed, and thus we cannot apply statistical tests such as two-way ANOVA [34] to evaluate the interaction effects. Instead, we apply the Scheirer-Ray-Hare test [35], which is a non-parametric test that has similar purpose as two-way ANOVA test. For any two independent variables A and B, we use Scheirer-Ray-Hare test to check for the interaction effects of them to the project success score. In total, we ran the Scheirer-Ray-Hare tests 42 times, and use Bonferroni correction to adjust the p-values. However, among the 42 tests, we do not find any statistically significance result. Thus, the interactions of these 7 independent variables have no statistically significant effect to the project success score. Similarly, we also investigate the impact of multiple independent variables to the scores in the 6 project success dimensions. The results of Scheirer-Ray-Hare test show that the interactions of these 7 independent variables have no statistically significant effect to the scores of the 6 project success dimensions.

To conclude, the independent variables proposed in our study affect the project success score independently, and there is no statistically significant interactions among them.

B. Implications

For Practitioners. Our results suggest that a simple personality matching strategy can help to build a more effective software project team. Developers in a company can take a DISC personality test, and the results of the test, along with the developer expertise, can be used to select suitable team members for a new project team. Alternatively, the test results can be used to estimate how likely an existing project team will be successful, and proper mitigation actions can be taken accordingly. Notice the DISC personality test only takes 20 minutes on average, and thus the cost is low.

For a new project team, our study suggests guidelines on how to select suitable team members. To achieve a higher project success score, our study suggests that the project manager should have a D-type personality, and there should be less than 15% D-type team members, and more than 10% I-type team members. Moreover, for existing project teams, we can analyze the personality distribution of the teams. For a team, if we notice there are too many D-type team members (e.g., more than 15%), it may be good to adopt mitigation strategies such as moving some of the D-type team members to other project teams, or if cost allows, include additional non D-type team members. Similarly, if a team has few I-type members (e.g., less than 10%), a team may want to include additional I-type personality members of suitable expertise from the pool of developers that a company has.

From the results of RQ4, we observe that teams with larger proportions of developers of different personalities perform differently for the six success dimensions. For example, when teams have more developers of types D and I, they are better at executing schedule. On the other hand, teams with more developers of C-type personality are less able to execute schedule well. Also, teams with more D-type members perform worse on software quality dimension. Management can use these findings as heuristics to optimize team's performance for a particular success dimension. For example, if management considers software quality as the most important factor to project success, our study suggests that they should avoid including too many D-type members to a team. And if a project team considers schedule is the most important factor, our study suggests to include more D-type or I-type members, and less C-type members.

For Researchers. In this paper, we analyze the relationship between project success and personality, and our results show that project manager personality and team personality composition do affect the success of software projects. Our study opens doors to a number of new research questions which we would like to encourage colleagues to investigate in the future.

Our study does a post-mortem analysis of projects with varying degrees of success. It would be interesting to complement our study with a field study where one monitors a set of developers in their day-to-day activities along with the artifacts that they create (either by manual observation or semi-automated analysis) for an extended period of time. Such a study can allow one to more comprehensively answer ques-

tions such as these: How D-type project managers perform in a team? What specific actions and patterns of interactions make D-type project managers good to team success? Why and how D-type and I-type team members contribute to lower software quality? Why and how I-type team members contribute to customer satisfaction? We encourage future studies to analyze and answer these research questions and many others that naturally arise from the results of this study.

We also encourage more applications of psychoanalysis tests to get a deeper insight into developers and their activities. Our literature survey finds that only a relatively low number of *recent* software engineering studies perform such analysis, with most related work done more than half a decade ago. Our study demonstrates that link between personality and overall project success exists. With a surge of open data available for analysis and increasing trend of practitioner participation in software engineering research, it would be a pity not to further drill down and investigate relationships on personality and software engineering activities that contribute to project success. For example, it would be interesting to explore whether developers of different personalities show different defect patterns, what personality types are better for code reviews, what personality types are better for requirement engineering tasks, etc.

C. Threats to Validity

Internal Validity. Threats to internal validity refers to errors in our collected data and experimental bias. We have double checked our collected data, however, there may exist some errors that we did not notice. Another threat is related to the selection of the personality test used. In this paper, we use DISC test which is often used by HR when hiring new employees [13], [20], is well founded on existing team formation theories [12], and was also used by a very recent work by Lykourentzou et al. [32].

External Validity. To improve the generalizability of our findings, we have conducted our study on data collected from a total of 346 professionals working on 28 projects from two IT outsouring companies. The 28 projects build a wide range of systems from Android apps to e-commerce portals, and all projects that we analyze are outsourced projects. In this study, all projects that we analyze are outsourced projects. Additionally, among the 28 projects, 15 are in the domain of finance; this may impact the generalizability of our findings. Moreover, all participants in our study are Chinese; it is not clear whether our results can be generalized to non-Chinese developers. In the future, to further reduce the threat to external validity, we plan to invite more developers from more different types of project teams in more IT companies to join our study.

Another threat is related to the 6 project success dimensions used in our study. Our two industry partners (Hengtian and IGS) use the success dimensions for more than 8 years and find them useful in practice. Also, we invite 32 managers to provide the weights of the 6 dimensions, and we average the responses from these managers. Notice that the standard deviation is low, which means the managers have similar opinions about

the weights of the 6 dimensions. Moreover, we ask them to revalidate the average weights, and they have a consensus that the average weights are reasonable.

V. RELATED WORK

Personality and Software Project. There have been some studies that investigate the link between personality types and team performance for project success and various software engineering tasks. We have described the work by Acuna et al. [7], Gorla and Lam [8], Karn and Cowling [9], and Wang and Li [10] along with differences with this work in Section I.

Aside from the above mentioned studies, Cunha et al. employed MBTI on 64 second-year undergraduate students to study the relationship between code review effectiveness and the personality of the reviewer [36]. They found that people who are of intuition type performed significantly better on code-review tasks than people of sensing type. Recently, Smith et al. employed Five-Factor Model on 797 participants to understand the beliefs, practices and personalities of software developers [37]. They found that there are no significant differences between developers and testers, and managers are more conscientious and more extraverted. Different from these studies, our study focuses on a different problem: the relationship between project manager personality and team personality composition and project success.

Our findings are also different from the previous studies. Wang and Li found Agreeableness is positively related to the team success. Agreeableness is one of the characteristics of I-type individuals. And our study shows that I-type project manager does not statistically significantly influence team success. Lykourentzou et al., who studied non-developers performing non-software engineering tasks, found that a team should have balanced personality (i.e., only one leader-type team member). However, our study suggests that a software team should have more members with I-type personality, and less members with D-type personality. Note that both I-type and D-type personalities are leader-type personalities.

Personality and General Team Work. The relationship between personality and how it affects the work performance of a team has been investigated in a number of past studies. Halfhill et al. found that teamwork is affected a lot by personality traits, e.g., agreeableness, openness and emotional stability, of the various team members [38]. They used the mean and variance of such traits to represent the group's personality composition (GPC) and studied its effect on team outcome. Team building strategies and team diversity were studied by Gilley et al. [6] and Muller et al. [39], and they found that more diverse teams collaborate better and produce comparatively better results.

How teams collaborate in virtual and crowd-sourced environments has also been of interest to researchers lately. Furumo et al. [40] and Horton and Chilton [41] studied how individual personality affects the trust among members of virtual and face-to-face teams, and found that in virtual teams, there is an accentuation of the effect on trust due to personality differences. In crowdsourced teams, the members drastically

differ in their agendas, values, beliefs and interpersonal communication styles [42]. These factors can have a major impact on team productivity as found by Gilley [43]. Irani et al. found that in crowdsourced work and teams, the members may not have a strong sense of sense of belonging and commitment, as the team members are treated as short term members and can be replaced anytime [44]. Lykourentzou et al., in a very recent work, found that balanced teams (i.e., teams with one leader) produce better work product [32].

Different from the above studies which consider general team work, our study focuses on software engineering team work. Findings for general team work may not translate to software development team work. Software development and maintenance is a complex and intelligence-intensive activity which involves a lot of communications and collaborations among team members. Additionally, the size of software teams are normally larger than teams considered in the above mentioned studies. Moreover, we consider a comprehensive set of success criteria which includes schedule, effort, external risk, internal issue, software quality, and customer satisfaction – these are not considered in many of the existing studies on general team work. The closest work to ours is by Lykourentzou et al.'s, which also used DISC test and investigated relationship between personality and project success. Our study identifies new findings not reported by Lykourentzou et al., i.e., software project teams with dominant managers, along with those with more influential members and less dominant members, have higher success scores.

VI. CONCLUSION AND FUTURE WORK

In this work, we investigate the relationship between developers' personality and project success. We conducted our study in 2 IT companies, inviting a total of 346 professionals who worked on 28 projects. We first asked all the participants to complete a DISC personality test. We then measured relationships between project manager personality and team personality composition and project success measured in various dimensions. We find that project teams with dominant managers, along with those with more influential members and less dominant members, have higher success scores. This work provides new insights to develop a personality matching strategy that can contribute in building effective project teams.

In the future, we plan to invite more people from more project teams to join our study to further validate our findings. We also plan to further investigate how large impact that of the personality of project manager and the team members can have on the project success while controlling for other confounding factors (e.g., the complexity of the system, the system size, and the team sizesystem complexity, team size, etc.) following methods employed in prior studies, e.g., [45]–[48].

REFERENCES

[1] W. Michel, Y. Shoda, and R. Smith, "Introduction to personality: Toward an integration," 2004.

[2] R. Bolton and D. G. Bolton, *People styles at work: Making bad relationships good and good relationships better*. Amacom, 1996.

[3] W. Mischel, *Introduction to personality*. London, 1993, vol. 250.

[4] J. Kidd, "Management teamslwhy they succeed or fail," *Journal of the Operational Research Society*, vol. 33, no. 4, pp. 392–393, 1982.

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

[6] J. W. Gilley, M. L. Morris, A. M. Waite, T. Coates, and A. Veliquette, "Integrated theoretical model for building effective teams," *Advances in Developing Human Resources*, vol. 12, no. 1, pp. 7–28, 2010.

[7] S. T. Acuña, M. Gómez, and N. Juristo, "How do personality, team processes and task characteristics relate to job satisfaction and software quality?" *Information and Software Technology*, vol. 51, no. 3, pp. 627–639, 2009.

[8] N. Gorla and Y. W. Lam, "Who should work with whom?: building effective software project teams," *Communications of the ACM*, vol. 47, no. 6, pp. 79–82, 2004.

[9] J. Karn and T. Cowling, "A follow up study of the effect of personality on the performance of software engineering teams," in *Proceedings of the 2006 ACM/IEEE international symposium on Empirical software engineering*. ACM, 2006, pp. 232–241.

[10] Y. Wang and F. Li, "How does project managers' personality matter?: building the linkage between project managers' personality and the success of software development projects," in *Proceedings of the 24th ACM SIGPLAN conference companion on Object oriented programming systems languages and applications*. ACM, 2009, pp. 867–874.

[11] I. Lykourentzou, A. Antoniou, Y. Naudet, and S. P. Dow, "Personality matters: Balancing for personality types leads to better outcomes for crowd teams," pp. 260–273, 2016.

[12] J. W. Gilley, M. L. Morris, A. M. Waite, T. Coates, and A. Veliquette, "Integrated theoretical model for building effective teams," *Advances in Developing Human Resources*, vol. 12, no. 1, pp. 7–28, 2010.

[13] J. H. Reynierse, D. Ackerman, A. A. Fink, and J. B. Harker, "The effects of personality and management role on perceived values in business settings," *International Journal of Value-Based Management*, vol. 13, no. 1, pp. 1–13, 2000.

[14] M. K. Slowikowski, "Using the disc behavioral instrument to guide leadership and communication," *AORN journal*, vol. 82, no. 5, pp. 835–843, 2005.

[15] R. A. Spies and B. S. Plake, *The sixteenth mental measurement yearbook*. Lincoln, NE: The University of Nebraska Press, 2005.

[16] D. Deviney, L. H. Mills, and R. N. Gerlich, "Environmental impacts on gpa for accelerated schools: A values and behavioral approach," *Journal of Instructional Pedagogies*, vol. 3, p. 1, 2010.

[17] "Disc validation research report," *https://www.manager-tools.com/docs/DiSC_Validation_Research_Report*.

[18] "Hengtian," *http://www.hengtiansoft.com/*.

[19] "Insigma Global Service," *http://www.insigmaservice.com/*.

[20] J. Diekmann, C. J. König, I. Nikolaou, and J. Oostrom, "Personality testing in personnel selection: Love it? leave it? understand it," 2015.

[21] P. Costa and R. McCrae, "Revised neo personality (neo-pi-r) and neo five factor inventory (neo-ffi) professional journal manual," *Odessa Florida: Psychological Assessment Resources*, 1992.

[22] J. L. Holland, *Making vocational choices: A theory of vocational personalities and work environments*. Psychological Assessment Resources, 1997.

[23] H. J. Eysenck, "The inequality of man," 1973.

[24] "Sonar," *http://www.sonarqube.org/*.

[25] X. Xia, D. Lo, J. Tang, and S. Li, "Customer satisfaction feedback in an it outsourcing company: a case study on the insigma hengtian company," in *Proceedings of the 19th International Conference on Evaluation and Assessment in Software Engineering*. ACM, 2015, p. 34.

[26] S. El-Sabaa, "The skills and career path of an effective project manager," *International journal of project management*, vol. 19, no. 1, pp. 1–7, 2001.

[27] S. S. Shapiro and M. B. Wilk, "An analysis of variance test for normality (complete samples)," *Biometrika*, vol. 52, no. 3-4, pp. 591–611, 1965.

[28] P. E. McKight and J. Najab, "Kruskal-wallis test," *Corsini Encyclopedia of Psychology*, 2010.

[29] F. Wilcoxon and R. A. Wilcox, *Some rapid approximate statistical procedures*. Lederle Laboratories, 1964.

[30] N. Cliff, *Ordinal methods for behavioral data analysis*. Psychology Press, 2014.

[31] H. Abdi, "Bonferroni and šidák corrections for multiple comparisons(http://www. utdallas. edu/˜ herve/abdi-bonferroni2007-pretty. pdf). in nj salkind (ed.). encyclopedia of measurement and statistics," *Encyclopedia of measurement and statistics*, 2007.

[32] I. Lykourentzou, A. Antoniou, Y. Naudet, and S. P. Dow, "Personality matters: Balancing for personality types leads to better outcomes for crowd teams," in *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*. ACM, 2016, pp. 260–273.

[33] W. G. Hopkins, *A new view of statistics*. Will G. Hopkins, 1997.

[34] S. Boslaugh, *Statistics in a nutshell*. " O'Reilly Media, Inc.", 2012.

[35] D. Holmes, P. Moody, and D. Dine, *Research methods for the biosciences*. Oxford university press, 2011.

[36] A. D. Da Cunha and D. Greathead, "Does personality matter?: an analysis of code-review ability," *Communications of the ACM*, vol. 50, no. 5, pp. 109–112, 2007.

[37] E. K. Smith, C. Bird, and T. Zimmermann, "Beliefs, practices, and personalities of software engineers: a survey in a large software company," in *Proceedings of the 9th International Workshop on Cooperative and Human Aspects of Software Engineering*. ACM, 2016, pp. 15–18.

[38] T. Halfhill, E. Sundstrom, J. Lahner, W. Calderone, and T. M. Nielsen, "Group personality composition and group effectiveness an integrative review of empirical research," *Small group research*, vol. 36, no. 1, pp. 83–105, 2005.

[39] M. Muller, K. Ehrlich, T. Matthews, A. Perer, I. Ronen, and I. Guy, "Diversity among enterprise online communities: collaborating, teaming, and innovating through social media," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2012, pp. 2815–2824.

[40] K. Furumo, E. de Pillis, and D. Green, "Personality influences trust differently in virtual and face-to-face teams," *International Journal of Human Resources Development and Management*, vol. 9, no. 1, pp. 36–58, 2008.

[41] J. J. Horton and L. B. Chilton, "The labor economics of paid crowdsourcing," in *Proceedings of the 11th ACM conference on Electronic commerce*. ACM, 2010, pp. 209–218.

[42] N. Salehi, L. C. Irani, M. S. Bernstein, A. Alkhatib, E. Ogbe, K. Milland *et al.*, "We are dynamo: Overcoming stalling and friction in collective action for crowd workers," in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, 2015, pp. 1621–1630.

[43] A. M. Gilley, *The manager as change leader*. Greenwood Publishing Group, 2005.

[44] L. C. Irani and M. Silberman, "Turkopticon: interrupting worker invisibility in amazon mechanical turk," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2013, pp. 611–620.

[45] D. A. da Costa, S. L. Abebe, S. McIntosh, U. Kulesza, and A. E. Hassan, "An empirical study of delays in the integration of addressed issues," in *Software Maintenance and Evolution (ICSME), 2014 IEEE International Conference on*. IEEE, 2014, pp. 281–290.

[46] P. Thongtanunam, S. McIntosh, A. E. Hassan, and H. Iida, "Review participation in modern code review," *Empirical Software Engineering*, vol. 22, no. 2, pp. 768–817, 2017.

[47] C. Bird, N. Nagappan, B. Murphy, H. Gall, and P. Devanbu, "Don't touch my code!: examining the effects of ownership on software quality," in *Proceedings of the 19th ACM SIGSOFT symposium and the 13th European conference on Foundations of software engineering*. ACM, 2011, pp. 4–14.

[48] X. Xia, L. Bao, D. Lo, Z. Xing, A. Hassan, and S. Li, "Measuring program comprehension: A large-scale field study with professionals," *IEEE Transactions on Software Engineering, to appear*, 2017.