

An Empirical Investigation of Personality Traits of Software Testers

Tanjila Kanij
Swinburne University of Technology
Hawthorn, Victoria, Australia
Email: tanjilakanij@gmail.com

Robert Merkel
Monash University
Clayton, Victoria, Australia
Email: robert.merkel@monash.edu

John Grundy
Swinburne University of Technology
Hawthorn, Victoria, Australia
Email: jgrundy@swin.edu.au

Abstract—Software testing is the process of an execution-based investigation of some aspects of the software’s quality. The efficiency of the process depends on the methods and technologies used, but crucially also on the human testers. Software testers typically attempt to anticipate and expose ways software may be defective, a fundamentally different task set to those of other software development practitioners. This raises the question of whether the personality of software testers may be different to other people involved in software development. To test this hypothesis, we collected personality profiles using the big five factor model of around 200 software development practitioners. Analysis of this data indicates that software testers are significantly higher on the conscientiousness factor than other software development practitioners, while other factors remain broadly consistent.

I. INTRODUCTION

A software tester is a person whose primary responsibility is to test software before release, helping to increase the reliability of a software product by reporting bugs so that they can be fixed. While software designers and programmers are largely *constructive*, in that they design and “build” something that meets customer requirements, a tester’s job is often in a sense fundamentally *destructive*, in that they attempt to “break” the software constructed by programmers. This fundamentally different task set, mindset and work approach of the testing profession raises an interesting research - and practical - question: might the effectiveness in the particular role of a tester be somehow related to their personality?

The majority of software testing research has been devoted to the enhancement of testing processes, test criteria, and to the development of new techniques and tools for different types of testing [6]. However, the limited available research [48], [5], [39], [22] supports the hypothesis that human factors have a strong influence on their effectiveness in this role. Some of these human factors are strongly connected with personality traits. If specific personality traits which have a significant influence on the effectiveness of software testing can be identified, this knowledge can potentially be helpful in recruiting, training, developing and possibly even in managing software testers. This information may also help beginning IT graduates to select an appropriate IT career path for themselves.

This research study was designed to investigate whether certain personality traits, as captured by the five-factor model, are over-represented among software testers. We collected the

personality profiles of a large group of software testers and a large group of people involved in other roles of software development in industry and compared those to determine if there are any significant commonalities or significant differences.

The rest of the article is organized as follows: Section II presents preliminary background information on personality, Section III discusses our review on the relevant literature, Section IV describes the details of this research study, Section V presents the results, Section VI lists the threats to validity of the research, Section VII presents our discussion on the findings and finally Section VIII concludes the article.

II. BACKGROUND

A. Personality

Personality encompasses non-intellectual, psychological characteristics that are most informative about an individual and that help to describe the differences between people. It is also thought to be organized, relatively enduring, an influence on the person’s interactions with others and influences their adaptation to their social environment.

B. Personality traits

The criteria by which people differ from each other are called psychological traits. Traits are representative factors to predict one’s behaviour patterns, feeling, thinking and related activities. Psychologists have derived five basic dimensions of personality that model a hierarchical organization of personality traits [38]. This model is well known as the “Big Five Factor” model of personality and is one of the most popular models in contemporary personality psychology research. The five factors are: Extraversion (E), Agreeableness (A), Conscientiousness (C), Neuroticism (N) and Openness to Experience (O).

Some popular personality assessment tests include the Myers-Briggs Type Indicator (MBTI) [1], NEO Personality Inventory [18], International Personality Inventory Pool [29] and so on.

III. RELATED WORK

A. Personality of software testers and debuggers

Most software testing research to date has focused on the technical side of the discipline, such as testing processes, techniques and tools [6]. There have been a few studies that

have examined the connection between ability in aspects of testing and various personality related factors.

Da Cunha and Greathead [20] examined the connection of MBTI personality types of students with their performance in a debugging task. They found that logical and ingenious people, as categorized by MBTI, were good at code review tasks. Almodaimeegh and Harrald [3] assessed the personality traits of programmers with locus of control. They also investigated the influence of social learning style on debugging skill measured with code comprehension, bug detection and bug repair testing. The study revealed that although there is no significant relation between locus of control and debugging skill, individual social learning style and experience influence debugging performance. Although debugging can be viewed as a part of testing, testing includes many other tasks besides just debugging. Studies focusing on debugging do not therefore give a full picture.

Shoab et al. [48], studied the effect of personality traits of students assessed with MBTI, on their effectiveness at exploratory testing and concluded that extrovert personality traits are positively correlated with effective exploratory testing. Exploratory testing is a specialized testing technique. Thus, based on the findings of this study, while it can be predicted that extraverts can be good exploratory testers, whether extraverts will be good testers in general remains an open question.

Capretz and Ahmed [14], [15] suggested that people assessed by the MBTI as “sensing and judging” will be good at software testing, based on the type of soft skills required for them. These two types of MBTI are roughly equivalent [42] to having high openness to experience and conscientiousness in the big five factor model. Rehman et al. [42] mapped the soft skills required for software testers to the “big five” personality traits and suggested that openness to experience and conscientiousness are important for the role of software testers. These suggestions were based on the experience and perception of the authors. No empirical data was provided.

Most of the studies described above employ the MBTI for personality assessment. While the MBTI is commonly used in this kind of research, there is considerable debate about the validity of the test [40], [9].

B. Other human factors of software testers

Beer and Ramler [5] explored the effect of experience on the effectiveness of software testing. In their field-based study they found that experienced testers used their domain knowledge to “fill in the gaps” in incomplete and ambiguous specifications. Itkonen et al. [31] studied 11 software testers to answer the question “How do testers do it?”. They found that software testers use a number of techniques and strategies for testing and they do not always rely on documents. Test execution techniques are strongly based on experience of the software tester and tests are run in non-systematic fashion. These studies emphasize the importance of experience in software testing.

Shah and Harrold [47] found that testing is considered a boring job and that junior and senior testers have different motivations and attitudes towards software testing. This finding

complements the outcome of an ethnographic study conducted by Rooksby et al. [43] where they found that testing is “a boring task” and most of the problems related to testing cannot be dealt with technical solutions, and instead require human and cooperative approaches.

Capretz et al. [2] analysed the soft skills listed in different software engineering job advertisements published across North America, Europe, Asia and Australia. They found that for software testers good communication skills were highly sought after. Good analytical and problem solving skills and organizational skills were secondarily sought in most job advertisements. The authors concluded that employers are not giving importance to the important soft skills needed for software testers.

Over all, there is limited evidence of research on different human attributes in software testing. The few studies mentioned above explain the attributes of a software tester such as coping with a monotonous job, and cooperativeness. These characteristics are plausibly related to certain personality traits. The investigation of such characteristics and relevant personality traits are, therefore, important.

C. Personality of programmers

There is a body of research associating personality types with the effectiveness of programmers. Until early 2000, most of the research of this category investigated whether a specific personality trait measured with MBTI was over-represented in the programming community. The studies of Sitton and Chmelir [40], Bush and Schkade [12], Lyons [36] and Chandler et al. [32] are all examples of this.

Later the research added to this by assessing the impact of the personality types on performance. Cegielski and Hall [16] explored the predictive power of personality along with theoretical value belief and cognitive ability on the performance in object oriented programming. They found that personality type was a predictor of the performance in object oriented programming. Darcy and Ma [21] used the five factor model of personality to find the influence of personality on the performance of students as programmers. They concluded that there were no significant differences in personality between the group that completed the given programming task and the group that did not.

D. Personality of software engineers

There is also some research analysing the personality traits of software engineers in general. Capretz [13] administered the MBTI to 100 software engineers ranging from student to professional level. They found that certain personality types were over-represented in the sample. In his further studies [14], [15] with Ahmed, Capretz associated different required soft skills of software engineers collected from job advertisements with MBTI personality types and suggested different personality types suitable for different roles of software engineering.

Sach et al. [44] analysed five research studies conducted from 1985 to 2010 using MBTI to determine the personality types of software engineers. The combined results indicates

thinking and judging type assessed with MBTI were over-represented among software engineers. A systematic literature review on personality research with software engineers conducted by Cruz et al. [19], found the majority of personality research examined pair programming and team effectiveness and the MBTI was used in most of the research.

Sodiya et al. [50] prepared a general test that assesses the five personality factors along with cognitive style with a standard questionnaire and suggests the software engineering role that best suits the test participant. Feldt et al. [24] studied the effect of personality of 47 professional software engineers, measured with the 50 item IPIP inventory on the attitudes of software engineers and found that there are different clusters of personalities among them and that each cluster has a significant correlation with attitude.

E. Expert views

Although the types and traits of effective software testers are yet to be established by research, there are some expert opinions about this. According to Armour [4], good testers have a “nose for testing”, an intuition that helps them to determine what and how to test. Pettichord [41] listed a number of distinct characteristics of programmers and software testers. He believes that software testers should tolerate tedium, be sceptical and be comfortable with conflicts while programmers should automate tedium, be believers and avoid conflicts.

Pol [37] and Black [7] suggested some specialized characteristics of software testers. While Pol suggests that software testers should be creative, accurate and strict in their methodical approach, Black thinks that software testers are professional pessimists who possess a curiosity for looking for faults. According to Burnstein [17], good communication skills, problem solving and team playing capability are important for software testers. The author also suggested software testers should be creative and open to new challenges.

IV. METHODOLOGY

In this research study, we collected the personality profiles of a large number of software engineers, based on the 50 item IPIP personality assessment test. We used a web-based survey as our research strategy since such a survey enables us to collect personality profiles of wide range of software engineers in a very short time. The survey was designed according to the six steps suggested by Kitchenham and Pfleeger [49] as discussed in the following subsections:

A. Setting the Objectives

The research study involved a group of software testers and a group of software developers who were involved in other roles of software development. The objective of the research study was to collect the personality profiles of the two groups and to conduct a comparative analysis to find out if there are any notable trends and significant differences among the two groups. The personality profiles of the participants were prepared based on the “Big five factor” model [38] of personality. We assume the following alternative hypothesis:

H_A: There is difference in mean on the five personality traits (Neuroticism, Extraversion, Openness to Experience, Agreeableness and Conscientiousness) between software testers and the non-testers.

In contrast to the alternative hypothesis we propose the following null hypothesis:

H_O: There is no difference in mean on the five personality traits (Neuroticism, Extraversion, Openness to Experience, Agreeableness and Conscientiousness) between software testers and the non-testers.

B. Survey Design

We conducted a web-based survey with a self-administered personality assessment questionnaire. One of the main benefits of web-based surveys is that the responses are collected in an automatic fashion and participants can complete the survey questionnaire at their convenience.

C. Development of Survey Instrument

We used the 50 item IPIP personality assessment test [29] for this research study. This test is designed based on the “Big five factor” model [38] of personality, which is the most commonly used contemporary model of personality psychology. The rationale behind selecting this personality test is that it can be used royalty-free and the test items as well as the scoring rules are available. The shorter version of the test could also be completed very quickly.

In order to have full administrative control over the collected data and to ensure the privacy and security of data to comply with human research ethics policy and standards, we built a data collection website for this survey, available at <http://www.testingsurveys.org/personality/>.

D. Questionnaire Design

The survey was divided into two sections: general demographic information and personality test items. The responses to each item could range from “Very inaccurate” to “Very accurate”. According to the scoring rule of the underlying test, a numeric value is associated with each possible response. To find the score on any factor of the five factors of personality, the numeric values of the respective items are summed. We used the numeric scores for analysis in this research. Based on the numeric score each of the five factors a personality assessment report is prepared and is presented to the participant. The report is also available in PDF format for download.

E. Evaluation of Survey Instrument

We conducted a pilot survey before the main survey. Based on the feedback obtained from the pilot survey minor spellings and a duplicate item were corrected. The responses to the pilot study are not used in the final analysis.

F. Obtaining Valid Data

We used cluster and purposive sampling to recruit participants for this survey. In cluster sampling, instead of selecting individuals from the population randomly, clusters of individuals are selected and within one cluster all individuals are included in the sample [23].

In this survey we requested permission from the 12 LinkedIn and 12 Yahoo! (software testing related) groups that gave us permission in our preliminary survey of software testers in previous work [39], [34]. Five Yahoo! and three LinkedIn groups approved us making the group response rate 41.67% and 25% for Yahoo! and LinkedIn, respectively.

In the purposive sampling process the sample is “hand picked” on the basis of relevance and knowledge [23]. As part of this process we attended a software testing related industry conference in Australia and posted to an industry conference email list related to software development. We also tweeted on the Twitter feed of the developer conference with the help of the organisers.

The participants could register for a draw of two \$100 Amazon.com gift vouchers by providing their email addresses. The email addresses were stored in a separate database table and were not associated with their responses. The individual response rate could not be determined since a participant can be member of more than one group. Nor is it known how many group members actually read the group emails.

G. Data Analysis

We report results with descriptive statistics. We used the Mann-Whitney U test to find the significance of difference of mean scores on five major personality traits. We used Hedge’s g to find the effect size to quantify the difference in mean. We also used power analysis to determine the statistical significance of our findings.

V. RESULTS

A. Demographic information

A total of 182 software engineers participated in our survey, and among them 45.1% were software testers. Among the rest of the participants 57% were programmers and 28% were managers. There were a small number of participants whose roles included business analysts, consultants, architects, software product designers. The gender and nationality of the participants are given in Table I.

The majority of our participants were male irrespective of their role. This is not surprising since the majority of practitioners in the IT field are male [46]. It is notable that there were a higher proportion of female participants among the software testers than in other respondents. We found that most of our participants were employed in IT companies. A small portion of the participants worked in non IT organizations and a very small portion of them were self employed. The majority of our participants had more than 5 years of experience.

B. Personality distribution

Figure 1 shows the percentages of participants with different levels of five major personality traits obtained in our study. The numerical score on the five major personality traits were categorized in three distinct levels: low, high and medium, suggested by Johnson [33] and applied by Norsaremah et al. [45] on a New Zealand based student sample. According to the scheme, if the score lies within lowest 30% boundary, the level is low, if the score lies within middle 40% the level is

TABLE I
GENDER AND NATIONALITY OF THE PARTICIPANTS

Criteria	Software testers (%)	Non testers (%)
Gender		
Female	16.5	12.1
Male	28.6	42.9
Nationality		
Australia	7.7	14.3
Bangladesh	8.2	8.2
Brazil	0	0.5
Canada	1.1	2.2
Croatia	0.5	0
Ecuador	0.5	0
Egypt	0.5	0
Germany	1.6	1.1
Hungary	0	0.5
India	4.9	2.7
Indonesia	0.5	0
Iran	0.5	0
Israel	0.5	0
Malaysia	0.5	0.5
Nepal	0.5	0
New Zealand	1.1	1.1
Pakistan	2.7	0
Peru	0.5	0
Philippines	0.5	0
Romania	1.1	1.6
Serbia	0	0.5
South Africa	0	0.5
Spain	0	1.1
Turkey	0	0.5
Ukraine	0.5	0
United Kingdom	2.7	2.7
United States of America	6	13.2
Not selected	1.6	3.3

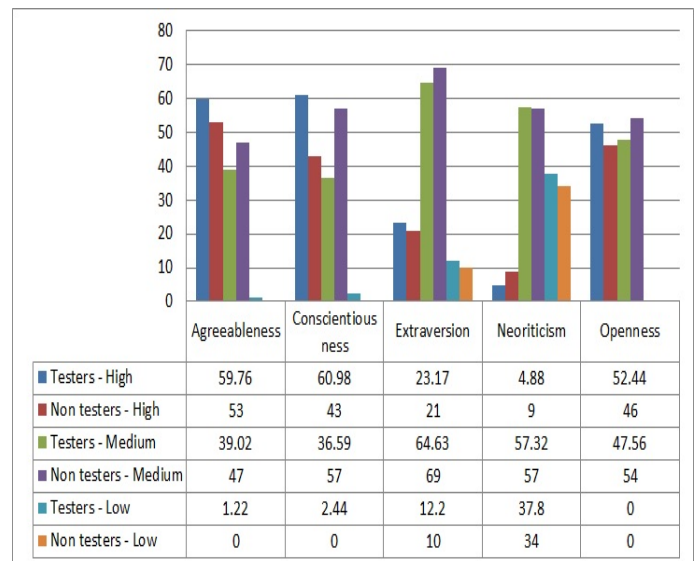


Fig. 1. Personality distribution

medium and if the score lies within the highest 30% then the level is high. From the distribution presented in Figure 1 we see there were many more non-testers with medium conscientiousness compared to testers who were highly conscientious. Both testers and non testers were agreeable, extravert and open to a close degree. However, we noticed a higher number of non-testers with high neuroticism.

C. Tests of normality

To test whether our sample was normally distributed or not, we have applied two well known tests of normality; the Kolomogorov-Smirnov and Shapiro Wilk tests [8], as shown in Table II. For these tests, if the significance value is less

than 0.05 then the distribution significantly deviates from the normal distribution.

From the obtained significance values via the Kolmogorov-Smirnov test, we see that, except for extraversion, the distribution of scores for other factors significantly deviated from the normal distribution. Using the Shapiro-Wilk test, we found the distribution of scores agreeableness, conscientiousness and neuroticism significantly deviated from normal distribution. From both test results we found only the distribution of extraversion did not significantly deviate from the normal distribution. The results of the tests of normality indicate that our sample was not normally distributed for each of the five factors.

TABLE II
TESTS OF NORMALITY

	Kolmogorov-Smirnova			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
A	.092	182	.001	.977	182	.005
C	.067	182	.045	.984	182	.036
E	.062	182	.081	.987	182	.096
N	.069	182	.036	.976	182	.003
O	.086	182	.002	.987	182	.090

a. Lilliefors Significance Correction, N= Neuroticism, E= Extraversion, O= Openness to experience, A= Agreeableness, C= Conscientiousness

D. Internal consistency

We have calculated the Cronbach's Alpha [28] to determine the internal consistency of the personality test items. Cronbach's Alpha is a measure of reliability of a test and can range from 0 to 1. The closer the Cronbach's Alpha is to 1 the greater the reliability is. The obtained Cronbach's Alpha of the items for agreeableness, conscientiousness, extraversion, neuroticism and openness to experience are 0.74, 0.81, 0.85, 0.83 and 0.67, respectively.

According to the proposed interpretation of George and Mallery [27], on our sample the reliability for conscientiousness, extraversion and neuroticism were good, reliability for agreeableness was acceptable and reliability for openness to experience was questionable.

E. Hypothesis testing

The mean and standard deviation of each of the five major personality traits measured by 50 item IPIP test on our sample is presented in Table III. The scores on each factor could range from 10 to 50 inclusive. We applied the Mann-Whitney U test to find if the mean score on each factor significantly varied between software testers and the non testers. This test is applied to compare the differences of mean between two independent groups when the normality of underlying distribution is questionable [8]. The results of the 2 tailed Mann Whitney U test are given in column 8 of Table III. We see only for conscientiousness $p \leq 0.01$, This indicates that the testers scored significantly higher in conscientiousness than non-testers. No other significant differences were found.

F. Effect size

An effect size can help us find the magnitude of mean differences. We have applied Hedge's g to find the effect size of the mean differences on the five major personality traits. The calculated effect size for conscientiousness is 0.39, which

TABLE III
POPULATION DISTRIBUTION

Factors	Software testers		Non testers		Combined		Mann Whitney U test*	Effect size	%	Statistical power
	Mean	SD	Mean	SD	Mean	SD				
A	38.51	5.53	37.55	5.21	37.98	5.36	0.191	0.18	58%	0.27
C	38.37	6.03	36.19	5.79	37.17	5.98	0.011	0.39	66%	0.76
E	32.91	7.62	31.51	7.24	32.14	7.43	0.155	0.19	58%	0.27
N	25.07	6.98	25.71	7.3	25.42	7.15	0.549	-0.09	54%	0.10
O	37.54	5.21	37.32	5.43	37.42	5.32	0.892	0.04	50%	0.05

N= Neuroticism, E= Extraversion, O= Openness to experience, A= Agreeableness, C= Conscientiousness, *Between software tester and non testers, % Percentage of non testers who would be below average person in software testing group

is described as a medium-strength effect. A medium effect implies that the mean score on conscientiousness between software testers and non testers are likely to be different. An interpretation of the effect size is presented in Table III. We see 66% of non-testers would be below average conscientiousness of a person in the testing group.

G. Power analysis

Power analysis indicates the probability of avoiding Type II error that is the probability of rejecting the null hypothesis correctly. We applied post-hoc power analysis. The computed power of our hypothesis testing is reported in Table III. We see that for conscientiousness we obtained a high statistical power. For the other factors, such as Openness to experience, the statistical power was relatively low. This means that we can have confidence in our result for significant difference in conscientiousness between testers and non-testers, however less confidence for the other results.

H. Comparison with the general population

The designers of the IPIP personality assessment test discourage establishment of and comparison against any norms. We were unable to find any norm on general population to compare our sample with. The only available results (mean and standard deviation) of application of the 50 item IPIP test on an internet sample of 2448 self-selected participants was reported by Buchanan [10] in an unpublished paper. However, the number of items for each factors was not same as has been used in this study. Therefore we have not directly compared our results with those reported by Buchanan. Even though the test applied by Buchanan used different items, those measure the same five factors. Therefore we calculated the inter correlations in our sample and compared with those reported by Buchanan et al. [11] using the method proposed by Fisher [25]. We found the intercorrelations were higher on our sample, indicating that the personality scores of our sample were more homogenous than those of Buchanan et al [11].

VI. THREATS TO VALIDITY

Threat to internal validity: In web-based surveys a major threat to the internal validity is the random responses of participants. The participants of this survey could register for a draw of two \$100 Amazon gift vouchers that may be a motivating factor for participation. However, the interest and enthusiasm observed in inviting participants contradicts this

view. The fact that only 45.6% (of the total participants) registered for the draw also illustrates that not every participant was motivated by the reward.

Threats to external validity: One possible threat to the external validity of this research is generalization of the findings. This is a common threat to such research conducted with a limited sample. Since it is impossible to conduct the study with an entire population, the findings can always be questioned for any bias caused due to sampling. However, the demographic information provided in Table I, shows our participants were from different countries. Most of those countries have well established software industry of their own. Thus we can consider our sample to be broadly representative.

Threat to construct validity: For this research study we have employed a freely available personality test instrument. The reliability of the test adopted for this research can pose a threat to the construct validity of the research. Although the personality assessment test is unrestricted, the test has become popular and is widely used [30]. Software testing is often conducted in teams and different members perform different roles. The effect of personality in such context could not be accounted for in a broad-ranging survey.

VII. DISCUSSION

The results indicate that there is no significant difference in mean scores on the five factors except conscientiousness between software testers and the non-testers for our sample population. The effect size results indicate that the probability of software testers being high on conscientiousness is higher than the non-testers.

Conscientiousness is related to personal organization of people [18]. Highly conscientious people tend to be more organized, disciplined and hard working. Higher number of software testers with higher conscientiousness conforms to the general assumption, found in our earlier informal survey of practitioners, that software testers generally need to be highly organized, disciplined and hard working.

The qualities exhibited by highly conscientious people seem to be important to succeed in any profession. However, these qualities might be particularly important for software testers. Capretz and Ahmed [15] analysed the job responsibilities of software testers as mentioned in job advertisements and suggested that sensing and judging type of people categorized with MBTI personality assessment test will be more successful as software testers. Sensing and judging types of MBTI are associated with conscientiousness of five major personality traits [26]. Our study's finding that software testers are more highly conscientious than others in the IT profession supports Capretz and Ahmed's perception.

Greathead and Cunha [20] found that intuitive and thinking type students assessed with MBTI were better at performing code review task. According to Furnham [26], Conscientiousness is associated with thinking and sensing type of MBTI. Sensing type is opposite to intuitive type in the dichotomous scale of MBTI. The observation of Greathead and Cunha, thus partly supports the influence of conscientiousness on effectiveness of software testing tasks.

Our study, in isolation, cannot explain why testers score higher on conscientiousness than other software developers. We cannot tell from these results whether it is a result of people who score more highly on conscientiousness self-selecting testing as their preferred career path, whether greater job performance in the area leads to them staying and advancing in testing longer rather than seeking alternative IT employment, or that those with higher conscientiousness enjoy testing more and thus prefer to stay working in the area. However, in our previous research study conducted with ICT students [35], we found a weak positive correlation between conscientiousness and bug location rate, and a weak negative correlation between conscientiousness and weighted fault density. Both bug location rate and weighted fault density were used as measures of effectiveness in software testing. This also indicates that the influence of conscientiousness on being successful software testers might be significant.

From the review of literature and our previous observation there is indication that Conscientiousness influences software testing effectiveness. The observation that software testers are more conscientious than others involved in software development makes the indication stronger.

VIII. CONCLUSION

The purpose of this research was to find whether there was any difference in personality traits, as assessed using the 50 item IPIP test, between software testers and others involved in software development. We conducted a web-based survey to collect personality profiles of nearly 200 IT practitioners. From the analysis of their personality profiles we found a significant difference on the conscientiousness trait between the software testers and other software developers. Software testers in our sample scored significantly higher than other software developers. We plan to replicate the study with a larger sample. Further studies are also needed to determine any actual link between personality and testing performance.

It would be desirable to repeat the study with a larger sample of IT professionals as our post-hoc power analysis showed results for the other factors were less reliable. We used a small Five Factor inventory to assess personality of our participants, in order to make the study less time-consuming and more attractive. However, all of the big Five Factor traits have sub-traits that may influence choice of tester vs. non-tester role. A replication using a much more detailed Five Factor inventory (e.g. the 300 question inventory) may allow us to identify sub-traits under one or more areas that differ significantly. Our study shows that practicing testers seem to be more conscientious than other practicing IT professionals. How this impacts the quality of their actual work, whether recruiting more or less conscientious testers matters or not, and whether less conscientious people employed as software testers can still perform acceptably well cannot be determined. Studies that explicitly link personality and job performance are required to explore these issues further.

IX. ACKNOWLEDGEMENTS

The authors thank Swinburne University of Technology and the Foundation for Research, Science and Technology for their funding supporting aspects of this research.

REFERENCES

- [1] MBTI. http://www.personalitypathways.com/MBTI_intro.html.
- [2] F. Ahmed, L. Capretz, and P. Campbell. Evaluating the demand for soft skills in software development. *IT Professional*, 14(1):44–49, Jan.-Feb.
- [3] H. Almodaimeegh and J. Harrald. Predicting debugging success: an investigation of the relationship between learning styles, personality traits, and computer program debugging. In G. Siemens and C. Fulford, editors, *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2009*, pages 2702–2710, Honolulu, HI, USA, June 2009. AACE.
- [4] P. G. Armour. The unconscious art of software testing. *Communications of the ACM*, 48(1):15–18, 2005.
- [5] A. Beer and R. Ramler. The role of experience in software testing practice. In *SEAA '08: Proceedings of the 2008 34th Euromicro Conference Software Engineering and Advanced Applications*, pages 258–265, Washington, DC, USA, 2008. IEEE Computer Society.
- [6] A. Bertolino. Software testing research: achievements, challenges, dreams. In *FOSE '07: 2007 Future of Software Engineering*, pages 85–103, Washington, DC, USA, 2007. IEEE Computer Society.
- [7] R. Black. Being a “good” tester: Attitudes, skills, and growth. <http://www.rbc-us.com/documents/BeingaGoodTester.pdf>.
- [8] S. Boslaugh and P. A. Watters. *Statistics in a Nutshell*. O’Reilly, Sebastopol, CA, USA, 2008.
- [9] G. J. Boyle. Myers-Briggs Type Indicator (MBTI): Some psychometric limitations. *Australian Journal of Psychology*, 30, Mar. 1995.
- [10] T. Buchanan. Online implementation of an IPIP five factor personality inventory. <http://www.docstoc.com/docs/36825859/Big-Five—Online-Implementation-of-an-IPIP-Five-Factor>, March 2013.
- [11] T. Buchanan, J. A. Johnson, and L. R. Goldberg. Implementing a five-factor personality inventory for use on the internet. *European Journal of Psychological Assessment*, 21(2):115–127, 2005.
- [12] C. M. Bush and L. L. Schkade. In search of the perfect programmer. *Datamation*, 31(6):128–132, Mar. 1985.
- [13] L. F. Capretz. Personality types in software engineering. *Int. J. Hum.-Comput. Stud.*, 58(2):207–214, 2003.
- [14] L. F. Capretz and F. Ahmed. Making sense of software development and personality types. *IT Professional*, 12:6–13, 2010.
- [15] L. F. Capretz and F. Ahmed. Why do we need personality diversity in software engineering? *Journal of Information Systems Education*, 35:1–11, March 2010.
- [16] C. G. Cegielski and D. J. Hall. What makes a good programmer? *Commun. ACM*, 49:73–75, October 2006.
- [17] J.-F. Collard and I. Burnstein. *Practical Software Testing*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2002.
- [18] P. T. Costa and R. R. McCrae. *The NEO personality inventory manual*. Psychological Assessment Ressources, Odessa, FL, 1985.
- [19] S. S. J. O. Cruz, F. da Silva, C. Monteiro, P. Santos, and I. Rossilei. Personality in software engineering: Preliminary findings from a systematic literature review. In *Evaluation Assessment in Software Engineering (EASE 2011)*, 15th Annual Conference on, pages 1–10, April.
- [20] A. D. D. Cunha and D. Greathead. Does personality matter? An analysis of code-review ability. *Communications of the ACM*, 50(5):109–112, May 2007.
- [21] D. P. Darcy and M. J. Ma. Exploring individual characteristics and programming performance: implications for programmer selection. *Hawaii International Conference on System Sciences*, 9:314a, 2005.
- [22] A. Deak. What characterizes a good software tester? a survey in four norwegian companies. 8763:161–172, 2014.
- [23] M. Denscombe. *The Good Research Guide for Small-Scale Social Research Projects*. Open University Press, Milton Keynes, UK, 2003.
- [24] R. Feldt, L. Angelis, R. Torkar, and M. Samuelsson. Links between the personalities, views and attitudes of software engineers. *Inf. Softw. Technol.*, 52:611–624, June 2010.
- [25] R. A. Fisher. On the probable error of a coefficient of correlation deduced from a small sample. *Metron*, 1:3–32, 1921.
- [26] A. Furnham. The big five versus the big four: the relationship between the Myers-Briggs Type Indicator (MBTI) and NEO-PI five factor model of personality. *Personality and Individual Differences*, 21(2):303 – 307, 1996.
- [27] D. George and P. Mallery. *SPSS for Windows Step by Step: A Simple Guide and Reference*. Allyn and Bacon, 1999.
- [28] J. Gliem and R. Gliem. Calculating, interpreting, and reporting Cronbach’s alpha reliability coefficient for likert-type scales, Oct. 2003.
- [29] L. R. Goldberg. *A broad-bandwidth, public-domain, personality inventory measuring the lower-level facets of several five-factor models*, volume 7, pages 7–28. Tilburg University Press, 1999.
- [30] L. R. Goldberg, J. A. Johnson, H. W. Eber, R. Hogan, M. C. Ashton, C. R. Cloninger, and H. G. Gough. The international personality item pool and the future of public-domain personality measures. *Journal of Research in Personality*, 40(1):84 – 96, 2006.
- [31] J. Itkonen, M. V. Mantyla, and C. Lassenius. How do testers do it? An exploratory study on manual testing practices. In *Proceedings of the 2009 3rd International Symposium on Empirical Software Engineering and Measurement*, ESEM ’09, pages 494–497, Washington, DC, USA, 2009. IEEE Computer Society.
- [32] J. C. Jane Chandler and I. Benest. Extrovert or introvert? the real personalities of computing students. In *4th Annual LTSN-ICS Conference*, NUI Galway, 2003.
- [33] J. A. Johnson. The IPIP NEO personality assessment tools. <http://www.personal.psu.edu/j5j/IPIP/>.
- [34] T. Kanij, R. Merkel, and J. Grundy. A preliminary study on factors affecting software testing team performance. In *ESEM*, pages 359–362, 2011.
- [35] T. Kanij, R. Merkel, and J. Grundy. An empirical study of the effects of personality on software testing. In *Accepted to be published at 26th Conference on Software Engineering Education and Training*, 2013.
- [36] M. L. Lyons. The DP psyche. *Datamation*, 31(16):103–105, 1985.
- [37] R. T. Martin Pol and E. V. Veenendaal. *Software Testing: A guide to the TMap Approach*. Addison-Wesley (E), 2001.
- [38] R. R. McCrae and O. P. John. An introduction to the five-factor model and its applications. *Journal of Personality*, 60(2):175–215, 1992.
- [39] R. Merkel and T. Kanij. Does the individual matter in software testing? <http://www.swinburne.edu.au/ict/research/sat/technicalReports/TC2010-001.pdf>.
- [40] V. C. Petersen. MBTI - distorted reflections of personality? Working Papers 2006-5, University of Aarhus, Aarhus School of Business, Department of Management, Sept. 2006.
- [41] B. Pettichord. Testers and developers think differently: Understanding and utilizing the diverse traits of key players on your team. *Testing & Quality*, 2, January-February 2000.
- [42] M. Rehman, A. Mahmood, R. Salleh, and A. Amin. Mapping job requirements of software engineers to big five personality traits. In *Computer Information Science (ICCIS), 2012 International Conference on*, volume 2, pages 1115–1122, June.
- [43] J. Rooksby, M. Rouncefield, and I. Sommerville. Testing in the wild: The social and organisational dimensions of real world practice. *Comput. Supported Coop. Work*, 18:559–580, December 2009.
- [44] R. Sach, M. Petre, and H. Sharp. The use of MBTI in software engineering. In *22nd Annual Psychology of Programming Interest Group 2010*, September 2010.
- [45] 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. *Empirical Software Engineering*, pages 1–39, 2012.
- [46] L. F. Sanz. Personal skills for computing professionals. *IEEE Computer*, (10):110–112.
- [47] H. Shah and M. J. Harrold. Studying human and social aspects of testing in a service-based software company: a case study. In *Proceedings of the 2010 ICSE Workshop on Cooperative and Human Aspects of Software Engineering*, CHASE ’10, pages 102–108, New York, NY, USA, 2010. ACM.
- [48] L. Shoaih, A. Nadeem, and A. Akbar. An empirical evaluation of the influence of human personality on exploratory software testing. In *Multitopic Conference, 2009. INMIC 2009. IEEE 13th International*, pages 1 –6, Dec. 2009.
- [49] F. Shull, J. Singer, and D. I. Sjøberg. *Guide to Advanced Empirical Software Engineering*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2007.
- [50] A. S. Sodiya, H. Longe, S. Onashoga, O. Awodele, and L. Omotosho. An improved assessment of personality traits in software engineering. *Interdisciplinary Journal of Information, Knowledge and Management*, January 2007.