Creativity Tests Versus Cognitive Computing: How Automated Personality Mining Tools Can Enhance Team Composition

Rangina Ahmad
University of Braunschweig
rangina.ahmad@tu-bs.de

Dominik Siemon
University of Braunschweig
d.siemon@tu-bs.de

Susanne Robra-Bissantz
University of Braunschweig
s.robra-bissantz@tu-bs.de

Abstract

Optimal composition of teams is an issue most enterprises face. Research conducted on this topic has identified personality as one of the key factors influencing team performance. The Big Five model, a framework for assessing personality, has standardized five personality traits, of which openness is reported to have a positive relationship with creativity. Creativity is regarded as one of the most relevant qualities for innovation. However, creativity as an ability manifested by performance on creativity tests is associated with difficulties. We therefore present cognitive systems as an alternative way, to not only find creative potential but also as a strategy to enhance team composition. Within our pilot study, we attempted to find a linkage between variables of creativity tests and the Big Five personality traits. Although our findings showed no salient correlations between these variables, we believe that automated personality mining tools would outperform creativity tests in the long run.

1. Introduction

The quality of establishing and maintaining efficient affiliations with peers is partially shaped by an individual's disposition [54]. This being the case, significant research has been conducted in the field of personality psychology, coming to the conclusion that personality traits are a relevant factor for team performance and outcomes [8, 47, 61]. Accordingly, increasing attention has been paid to one of the enduring questions that particularly organizations and institutions face; that is, how a proper constellation of personalities may lead to an ideal team composition, whose potential could be used in the most effective and efficient manner.

In trait theory it is assumed that people's behavior and feelings can be explained to some extent in terms of underlying personality traits, which are regarded as relatively stable features of an individual's personality [10]. Multifactorial models such as the Big Five model have formalized traits in order to measure personality. Derived through factorial studies, five fundamental traits or dimensions have been defined for a comprehensive assessment of individuals: Openness (refers to the extent to which a person is open to experiencing a variety of activities, and prefers novelty over convention), Conscientiousness (a person's tendency to act in an organized or thoughtful way), Extraversion (refers to the extent to which people enjoy company, and seek excitement and stimulation), Agreeableness (a person's tendency to be compassionate and cooperative toward others), and Neuroticism (or also referred to as Emotional Range: the extent to which a person's emotions are sensitive to the individual's environment) [10, 28].

Examining the link between the Big Five dimensions and numerous outcomes, a number of studies have reported a positive relationship between the trait openness and creativity [20, 35, 44]. This means an individual who scores high on the personality dimension openness is in general described as imaginative, original, intellectually curious, has artistic interests and is able to think about new ideas [10, 44]. Previous studies have therefore suggested that openness can be interpreted as a proxy of creativity [20, 35]. Creativity in turn fosters idea generation and innovative thinking. In fact, innovation is considered to be a constant process rooted in the continuous need of creative ideas of individuals and groups [67]. Especially in an organizational context, innovation is necessary for a company to survive and to have a competitive advantage. Organizational innovation has been therefore defined as the "successful implementation of creative ideas" [1, p. 126]. This implies that people who tend to approach and solve problems in a creative manner are specifically of great value for enterprises.

An individual's creativity level is often assessed through performance measures derived from tests of creative thinking, which is associated with the creative process itself (e.g. thought mechanisms). However, creativity as an ability manifested by performance on tests is attended with difficulties [6]. Some researchers argue that due to the lack of a

URI: http://hdl.handle.net/10125/49903
ISBN: 978-0-9981331-1-9

general creativity skill, tests of creativity ability are missing validity [3]. Further, creativity tests are time-consuming (having a completion time of up to 45 minutes), test results can be furthermore error-prone (biased by test anxiety) and also be biased by an evaluator's cognitive abilities (due to the limited information processing and memory capabilities of human beings) [9, 27, 48]. Throughout several decades of research on creativity, various investigators emphasized the "creative personality" and suggested to focus on the creative person rather than solely on the creative process, as personality traits predictably relate to creative achievement [2, 18, 22, 44].

General expressions of individuals' personalities can be found in many aspects of everyday interactions with the physical and social environment [43]. For example, researchers have shown that people's personality traits can be identified by examining their words used in daily life [55]. More exactly, the frequency with which certain categories of words are used as well as the variations in word usage in writings can predict aspects of personality [21, 69]. Leveraging the findings in Psycholinguistic studies, an individual's personality traits can automatically be computed from one's linguistic footprints left for instance on social media [27]. Smart machines – based on cognitive computing, which is considered to be a technological evolution – are able to automatically infer personality traits from an individual's text by applying linguistic analytics and personality models [21, 29, 62, 69].

Motivated by previous studies that have found a positive relationship between the personality trait openness and creativity, the purpose of this paper is to examine whether there exist further correlations between the components of a creativity test and the Big Five personality traits and its sub-categories. The present pilot study was therefore conducted to investigate if a person's creativity test scores are related to their Big Five personality traits. We assume that automated personality mining systems would outperform creativity tests in many ways such as in accuracy, performance and validity. Moreover, personality mining tools would not only facilitate finding "creative persons" but could also help build a profile of an individual's personality, which can be used to effectively compose teams in organizations and thus reduce group inefficiency and ineffectivity. With our conducted pilot study we want to propose a systematical approach and "determine initial data for the primary outcome measure" [37, p. 308] for future research.

2. Theoretical Background

2.1 Personality

When it comes to the term personality, it has proven to be difficult defining it accurately. Psychologists have argued for a long time specifying what exactly is understood by this word, but no definition that has been offered yet, is universally accepted. Hall and Lindzey [1957] state that "no substantive definition of personality can be applied with any generality" [11, p. 9]. Due to the complexity and diversity of human personality, it is sheer impossible to unite this construct in a single, coherent theoretical framework [60]. Throughout the history of psychology, diverse approaches have competed with each other, leading to various groupings within the field of personality psychology. The dispositional approach considers *trait* as the key concept of the field of personality, whereby traits are also related to the processes of personality measurement. Accordingly, personality is what makes a human being's behavior, thoughts and feelings (relatively) consistent, but at the same time is the construct that differentiates individuals from another [1]. The generally accepted taxonomy of the multifactorial Big Five model provides a guide to the comprehensive assessment of individuals and was found to be stable across cultures, as well as instruments and observers [43, 45]. The Big Five traits are summarized in table 1.

Table 1. The Big Five Personality Traits

Dimension	High Scores	Low Scores
Openness	imaginative, creative, original, prefers variety, curious, liberal	down-to-earth, uncreative, conventional, uncurious, conservative
Conscientiousness	hard-working, well-organized, punctual, ambitious, persevering	negligent, lazy, disorganized, late, aimless, quitting
Extraversion	affectionate, joiner, talkative, fun loving, active, passionate	reserved, loner, quiet, sober, passive, unfeeling
Agreeableness	soft-hearted, trusting, generous, acquiescent, lenient, good-natured	ruthless, suspicious, stingy, antagonistic, critical, irritable
Neuroticism	anxious, temperamental, self-pitying, self-conscious, emotional	calm, even-tempered, self-satisfied, comfortable, unemotional

2.2 Big Five Workforce Outcomes and Team Composition

Research has shown that personality traits are predictive of many life outcomes, including workplace-related outcomes. In virtue of the hardworking, persevering nature of conscientious individuals, their behavior is associated with goal completion and problem solving, and can thus be seen as beneficial and predictive of job performance [26]. Extraversion has been positively correlated with leadership abilities [40] and job satisfaction [34]. Extraverts also tend to search for social relationships with co-workers, leading to beneficial interpersonal interactions [53]. Further, people who are emotionally stable (e.g. calm and steady and thus score low in neuroticism) might create a relaxed atmosphere that promotes cooperation and therefore might engage in less disruptive behavior [57]. Highly agreeable individuals tend to easily adapt, seek to maintain social harmony and reduce within-group competition, which leads to cooperative behavior [10, 52]. Lastly, a number of studies have reported a positive relationship between the trait openness to experience and divergent thinking, which refers to creative problem solving [20, 35, 44]. Accordingly, openness may be related to team performance to the extent that individuals high on this trait are more adaptable and can make the changes required to continue in a dynamic team environment due to their high levels of creative behavior [7, 39].

Personality has been furthermore found to be a valid predictor of team performance, which is defined as the extent to which a team achieves its goals or mission [14]. This finding has lead to an increased research on groups in the few past decades [8, 47, 61]. As organizations structure work through the use of work teams, the need for effective strategies to staff these teams has become more relevant [5, 36]. Work or group teams are defined as units of two or more individuals who interact interdependently to achieve a common objective for their organizations [4, 63]. Investigators thus have focused on research on how to increase team performance via team composition, which is the mix of individual characteristics to put into a work team such as personality (e.g. traits), demographics (e.g. sex or race), and abilities (e.g. intelligence or expertise) [24]. Conversely, Bradley and Hebert [1997] allude that the cause of ineffective teams may be the product of inappropriate team composition, and further emphasize the impact of personality on team performance [8].

A question that often arises in this matter is whether homogeneous or else heterogeneous (also called diverse) work groups are more effective when it comes to team formation [46]. As both approaches may have positive as well as negative effects on group performance, these processes pose major challenges to research in organizational behavior [65]. As opposed to homogeneous work groups, heterogeneous teams bring a wider variety of solutions to a given problem and enhance team creativity, research suggests [42]. Accordingly, when convergent thinking is required (the ability to come up with a single but correct solution to an actual problem or given potential) homogeneous teams should outperform heterogeneous groups, but diverse groups should outperform homogeneous teams when divergent thinking is required (which involves the generation of multiple answers to an often loosely defined problem) [46, 50].

Divergent thinking is theorized to result in high levels of creative behavior [20, 35]. This implies that people who tend to approach problems in a creative manner are specifically of great value for organizations. Yang and Choi [68, p. 298] further found that "creativity is a predictor to explain a significant improvement of team performance." Consequently, increased interest and attention to the determinants of creative potential has been paid by enterprises, as creative employees generate novel ideas that can be the starting point for innovation [2].

2.3 Creativity and Creativity Tests

Creativity is a widely used term, however, researchers have been unanimous in their precise definition of the word. In his *Analysis of Creativity* [1961], Rhodes observed that theories on creativity have focused pre-eminently on four aspects: Person, Process, Press and Products [58]. The Four P's represent the essential cornerstones of creativity and are intertwined and influence each other. However, only the person- and process-oriented definitions of creativity are most appropriate for present purposes, as the measurement of creativity is central to this paper. The person-oriented term focuses on the nature or disposition of the creative person, while the process-oriented term defines creativity according to the process itself [58]. Creativity theories of this strand for instance include stages or models of the creative process [49]. Cognitive studies of creativity therefore focus on thinking abilities such as convergent and divergent thinking. Creativity tests relate to the creative process, aiming to measure creative thinking [13]. Further, creative outcomes must be "novel-original and useful-adaptive" to a specific task [19, p. 290]. The thinking abilities that are involved in order to be considered as creative,

consist of the generation of novelty (via divergent thinking) on the one hand and the evaluation of novelty (via convergent thinking) on the other hand [12]. Facaoaru [1985] and Lonergan et al. [2004] purport that the idea of a two-step creativity testing procedure, which assesses the area of "overlap" between the two thinking styles, is widely accepted as both convergent and divergent thinking lead to production of ideas [12, 16, 41].

2.4 Limitations of Creativity Tests and Advantages of Automated Personality Mining Tools

The search for creative potential by means of creativity tests has been carried on in many domains. However, creativity as an ability manifested by performance on tests is also associated with formidable difficulties [6]. Some investigators for instance argue that tests of general creativity ability lack validity, because unlike intelligence there is no general creativity skill to be measured [3].

Another limitation creativity tests generally face is that with a completion time of up to 45 minutes, such tests are not only extremely time-consuming, but can also be biased by test anxiety: In regards to convergent thinking for instance, a study has found that effects of neuroticism on IQ test performance were fully accounted for by test anxiety [48]. In another study from 2008 by Chamorro-Premuzic and Reichenbacher, the researchers assessed whether personality traits affect participants' performance differently under stressful and calm conditions. The researchers found out that neuroticism correlated significantly with divergent thinking (negatively) only under threat of evaluation [9]. Thus, a candidate's test results can be error-prone or influence their creative thinking, depending on the environment a person conducts a test (e.g. in an Assessment Centre).

Further negative aspects of questionnaires are cognitive biases on the part of the evaluator: Regarding team composition, evaluators (e.g. recruiters, project managers etc.) must find the most suitable personalities in order to build an effective team. Human being's decision making is however influenced by limited information processing and memory capabilities [27]. Considering the enormous number of potential candidates (including job applicants from Social Media), a situation where a candidate's test results have to be assessed and where the applicant's suitability for a position in a team has to be evaluated, might be overwhelming and not effectively enough handled by evaluators [17].

In order to avoid these negative effects associated with creativity tests, the question has been raised whether smart machines – more precisely automated personality mining services – are the "better" tools when it comes to finding creative potential and suitable personalities for a work group. Assertions and evidence of well-respected investigators who did research on personality traits associated with creative achievement support the suggestion to focus on the creative person rather than solely on the creative process: At a very early research stage, Guilford [1950, 1967] drew attention to the importance in creativity of factors such as personality, and suggested a trait approach for the study of creativity [12, 22, 23]. Amabile [1988] analyzed qualities of individuals that influence creativity and came to the conclusion that among other factors, personality traits were the one that promoted creativity the most. In his famous paper *A Meta-Analysis of Personality in Scientific and Artistic Creativity* [1998], Feist concluded that a "creative personality" does exist and that personality traits predictably relate to creative achievement. McCrae [1987] finally suggested that although divergent thinking may indicate aptitude for creativity, it is the trait openness that serves as the catalyst leading to creative expression [35, 44].

Recall, when it comes to tasks that require creative problem solving, heterogeneous groups propose better quality solutions than homogeneous groups. A heterogeneous work team in turn consists of various factors, including diverse personalities (i.e. traits). This suggests that a team aiming for creative output should comprise a wide range of personalities with the ability to have both thinking styles. A two-step creativity testing procedure may indicate aptitude for creative potential, but it does not give further details about an individual's personality in order to be able to compose a team – and this is where the deployment of automated personality mining tools could come in handy.

A well-accepted theory of psychology and other fields is that human language reflects the personality, emotional state and thinking style of a person, as the frequency with which certain categories of words are used can provide clues to these characteristics [21, 69]. To turn to the issue at hand, personality mining algorithms analyze words written by persons and automatically infer portraits of individuals that reflect their personality characteristics [33]. Further, they not only can infer the "creativity trait" openness from text, but also deliver an overall personality profile, which then could be used for a more effective team composition. Further advantages of automated personality mining tools would be savings in time and money. Smart machines also expand human

capabilities. Such cognitive computer systems, which are designed to participate in organizations by carrying out complex cognitive tasks could be an enormous benefit when it comes to avoiding ineffective and inefficient team building [51].

2.5 Creativity Tests and IBM Watson

Traditionally, many creativity tests are conducted in the form of paper and pencil tests. On grounds of the development of digital technologies, however, many standardized tests can now be administered on computers or online. Hence, using a computerized test for the present study resulted in much faster data collection. The idea management platform Creaboration, developed by researchers and students at the University of Braunschweig, was used as a data collection platform for the study. It comprises a creativity test based on a two-step testing procedure, that participants used in the context of the study. Two out of five tasks – Guilford's Alternative Uses Test [1967] and Wallach-Kogan's Test [1965], hereafter referred to as Guilford Test and Wallach Test, respectively – focus on measuring divergent thinking, while the remaining three tests are based on Dow and Mayer's Insight Problems Task [2004] that is aimed at measuring convergent thinking [15, 23, 66]. The Guilford Test asks participants to think of as many uses as possible for a simple object, in this case to think of all possible uses for a brick. The Wallach Test in contrast requires listing all items containing the component "wheel". The Insight Problems Task is a test categorizing convergent thinking into three groups of tests aiming to measure mathematical, verbal and spatial insights. Both, the divergent thinking tests as well as the convergent thinking tests are carried out consecutively. Creaboration also serves as a tool to upload a written idea. Users are being asked to write down their opinions and ideas on a generic topic such as how to make planet earth healthier. Their answers can be then uploaded on the platform. For the conducted study, the participants' uploaded ideas were used as input texts to be analyzed by IBM Watson's personality mining tool.

With the creation of Watson, IBM has built a smart machine that answers human language questions by scouring a vast amount of sources of related information in a short amount of time, and as a result is capable to deliver accurate answers as well as new insights from extremely large volumes of information. The cognitive system applies advanced deep Content Analysis, Information Retrieval, Artificial Intelligence, Machine Learning and Natural Language Processing (NLP) technologies to the field of open-domain question answering [19]. As opposed to traditional IT systems (such as search engines), cognitive systems operate at a different level, as they enable individuals a more natural human-machine interaction. If Watson has enough data and contextual knowledge related to the question, the user gets a directed result – either an answer to the question or a follow-up question to support clarify the individual's intent [27]. Further, one of the essential principles of a cognitive system is the process of human natural language. NLP techniques interpret the relationship between massive amounts of natural language elements and infer the meaning of a word, phrase, sentence, or document by recognizing the grammatical rules such as their syntax, context and usage patterns [19, 27].

IBM offers Application Programming Interfaces (APIs) in the domain of language, such as the Personality Insights (PI) service – an automated personality mining tool used for the purposes of the present study. PI extracts personality traits based on how a person writes by using linguistic analytics and personality theory. The PI service provides an API that enables deriving insights from enterprise data, digital communications (e.g. e-mail, text messages, blog posts, tweets etc.) or other texts given access to it [29]. From a potentially noisy corpus of text, the service can automatically infer portraits of individuals that reflect their personality traits [29]. The current version of the PI service is based on an open-vocabulary approach – a technique called "Global Vectors for Word Representation", or GloVe, which is an open-source word-embedding algorithm for obtaining a vector representation for the words in the input text [56]. Analyzed text of an individual should reveal the author's personal experiences, responses and thoughts, and should thus be rather "reflective" than formal [30]. A subjective writing style thus reflects the author's personality traits best, as it requires a certain amount of thought into the words that they choose, while scientific articles are only marginally suited for inferring personality [31, 55, 62, 64, 69].

3. Method

The intention of this paper is to investigate whether there exists a linkage between the components of the creativity test – that is the Guilford Test, Wallach Test, verbal test, mathematical test and spatial test – and the Big Five variables, i.e. openness, conscientiousness, extraversion, agreeableness, neuroticism and its 30 sub-categories or facets.

The pilot study is based on a Mixed Methods Research Design called Exploratory Sequential

Design, which is a two-phase approach: The design begins with collecting qualitative data, and then builds to a second quantitative phase [11]. The present study first collected (primarily) qualitative data (i.e. the participant's ideas and results of the creativity test) via the platform Creaboration. The result data was then analyzed quantitatively in the second phase of the study.

3.1 Subjects

A sample of 13 participants aged 14-34 years (8 females and 5 males; age, M = 27,8; SD = 5,4) took part in this study. Apart from one pupil who attended school, 54% of the participants were students majoring in different fields of study such as pharmacy, medicine, business and computer science. The remaining 39% were graduates working as a doctor, consultant, biomedical scientist, PhD student and a school teacher. Further, merely people whose first language was English or who have been living in an Anglophone country for more than three years were considered as subjects for the experiment. This has two major reasons: IBM Watson's algorithm is presently capable of analyzing a written text only in a limited amount of languages – that is Spanish, Japanese, Arabic and English. Due to personal unfamiliarity with the first three languages, we decided to only focus on English for this study. Secondly, only native respectively advanced English speakers were qualified for the experiment, since the tasks required fluent English. This issue was very much pivotal for the evaluation: A person who only speaks basic or moderate English, and has to look up every other word in the dictionary in order to work on the tasks would have biased the results, as Watson's evaluation is strongly premised on linguistic analytics. Based in Australia, USA or the UK, all 13 participants were acquired via e-mail and through direct approach, the majority of them being friends, relatives or acquaintances.

3.2 Study Structure

The study consisted of two tasks: First of all, the participants had to write down a comprehensive idea on how to save planet earth. The account for letting people ponder over this issue was simply because it is a pervasive, omnipresent topic the majority of humans have an opinion on. In addition, it poses a topical subject where none of the participants need to have specific expertise, in order to be able to generate ideas. A more specified question was intentionally not given, so that the subjects could be as creative and innovative as possible in their answers. However, since Watson needs at least 1600 words to provide a decent analysis, we suggested in the scope that certain aspects such as acceptability, implementability, workability and/or the feasibility of ways to protect the environment could be kept in mind while drafting the ideas. The intention hereby was to deliver more input for the writers without narrowing down the topic, and to ensure that a minimum of the 1600-word mark was reached. By virtue of this time-consuming task, it was clearly communicated that participants should not feel compelled to submit their ideas all at once.

A further essential condition mentioned in the scope was that every contribution must be of original content – no copy and paste was allowed, the submitted texts had to represent each and every participant's own writing style. So for instance, the text could express the writer's likes and dislikes, include their attitudes, sentiments, observations or their opinion on the respective topic – eventually the text should be considered as a reflection of the author's selection of words. Using information out of respective references or seeking for inspiration in the Internet was permitted.

The second task demanded considerably less time (approx. 15-20 min.) than task one. The subjects went through the Guilford, Wallach, mathematical, verbal and spatial test by answering the questions consecutively and without any help. Both, uploading the idea as well as doing the creativity test were conducted through Creaboration. The subjects were given 5 months to complete both tasks. After having completed the process of qualitative data collection, the gathered data on Creaboration has then been analyzed quantitatively by using tools such as IBM Watson's PI service, Microsoft Excel and the statistical program R. Every participant's data was perused individually.

3.3 Procedure

In order to find correlations between a participant's Big Five variables and the results of their creativity test, we let Watson's PI service analyze the original texts of the subjects with the result that the program delivered us the required percentiles of the Big Five variables (via JSON). Creaboration, however, does not transfer and evaluate input into numeric values by default. In order to create comparability with the percentiles of the Big Five variables, an appropriate valuation system was required for the results of the creativity test, which reads as follows: For a correct answer during the mathematics or verbal test (both of them consisting of five questions each), the participant

received the item 0,2. If another question was equally answered correctly, the item was incremented up to 0,4 and so forth. This means, if all five answers were accurate, the subject received a score of 1. Accordingly, incorrect answers were rated with the item 0 (e.g. a participant answered four out of five questions of the mathematical test right, and all five questions of the verbal test correctly; this resulted in a score of 0,8 and 1, respectively). Due to the fact that the spatial test comprised three instead of five questions, the values in this case were divided into three items: 0,34 (for one correct answer), 0,67 (for two correct answers) and 1 (if all questions have been answered accurately). Wrong answers were once again rated with 0 (e.g. a participant answered two questions in the spatial test correctly, which resulted in a score of 0,67). Unlike the three previously mentioned convergent thinking tests, the Guilford and Wallach Test did not aim to ask questions where only one correct answer was possible, but rather aimed to test the participant on their divergent thinking abilities in the form of word-associations. As the Guilford Test asked the subjects to enter all possible uses for a brick, and the Wallach Test demanded to list all items containing the component "wheel", the valuation of this data required a different approach: The results were only measured by fluency, meaning that the amount of words entered per participant were divided by the total number of entered words by all participants, i.e. for the Guilford Test the subjects listed a number of 55 words altogether. After deducting redundant words (e.g. "building houses" was mentioned 11 times), the total number of entered words by all participants summed up to 27. So if a subject entered 7 different uses for a brick, their score resulted in 0,26 (=7/27). The same approach holds true for the Wallach Test, where the total (not redundant) number of words were 52. Listing 9 words individually for example would have resulted in a score of 0,17 (=9/52). These computed values of the creativity test scores as well as of the Big Five variables in the second phase of the study served as the basis for the Pearson's product-moment correlation with a 95% CI.

3.3 Results and Discussion

Apart from a moderate positive relationship with the Big Five trait extraversion, both of the two divergent thinking tests Guilford ($r = 0.58$, $p = 0.039$) and Wallach ($r = 0.42$, $p = 0.1577$) showed no significant correlations with the remaining variables of the Big Five. The mathematical test showed a moderate positive correlation ($r = 0.56$, $p = 0.046$) with the openness-facet intellect. Further, the results of the verbal test depict similar moderate positive correlations with the extraversion-facet excitement-seeking ($r = 0.59$, $p = 0.035$), as well as with the openness-facet intellect ($r = 0.55$, $p = 0.052$). Of all tests, the spatial test has the most considerable number of correlations with the Big Five variables. Although correlations are moderate, the relationship between the spatial test and the Big Five variables are positive and negative: With $r = 0.65$, $p= 0.016$, the convergent thinking test has the strongest positive correlation with the Big Five trait extraversion. It has a further positive relationship ($r = 0.5$, $p = 0.082$) with the openness-facet adventurousness. Big Five variables that are negatively correlated with the spatial test are neuroticism ($r = -0.65$, $p = 0.016$), the agreeableness-facet altruism ($r = -0.48$, $p = 0.100$), conscientiousness ($r = -0.51$, $p = 0.076$) and its facet dutifulness ($r = -0.59$, $p = 0.034$).

Examining the pattern of results, it seems fair to conclude that the study has shown no particularly salient correlations between the Big Five and creativity test variables, due to the lack of any strong correlation coefficients. However, what is in fact salient is that contrary to our expectation, the two divergent thinking tests show either very weak or no correlation at all with the trait openness and its facets. The positive relationship between the openness-facet intellect and the mathematical test in contrast support the assumption that a convergent thinking test seems to be indeed a predictor for intelligence. Further, there seem to be a moderate positive relationship between the spatial test and extraversion as well as a positive correlation between the verbal test and the extraversion-facet excitement-seeking. However, this is not backed up by any findings in research so far. In fact, Roberts [2002] investigated the relationship between extraversion and ability and came to the conclusion that a correlation between extraversion and verbal and spatial ability tests is likely to be inconsistent and fragile [59]. The same general inconsistency applies for the negative relationship between the spatial test and the traits neuroticism and conscientiousness.

3.4 Conclusion and Outlook

The present study aimed to investigate the assumption, whether automated personality mining tools could replace creativity tests. The conducted study, however, has shown no particularly salient relationships between the Big Five and creativity test variables. Nevertheless, a few limitations of the study should be taken into account when interpreting the results. First of all, data were obtained from 13 participants – this low sample size suggests that

statistical power is relatively weak. Low statistical power in turn may have contributed to the lack of statistical significance.

Although research emphasizes a strong correlation between openness and in particular divergent thinking tests, the present study was unable to find support for this. A plausible reason for this may be the fact that neither the Guilford Test nor the Wallach Test was conducted thoroughly by the subjects. It can be seen from this that the majority of participants solely listed a small amount of words (1 – 4) on average when asked for alternate uses to find creative uses for the respective items. The fluency score however, aimed to measure creative potential is determined by counting the number of ideas that were produced. Thus, the reason for the sparse word listing might not necessarily be a lack of creativity on the part of the participants but on any other grounds such as environmental conditions that could have influenced their fluency ability (e.g. being in a hurry, etc.). A further limitation concerns the word count of the texts submitted by the subjects (1600 words on average). The number of words that an idea contains plays an important role. Despite the fact that a 1600-word text delivers a decent analysis, IBM Watson's PI requires at least 3500 and preferably 6000 words to give a reliable estimate of personality [32]. This suggests that although the analysis of the participants' personality profile in this study was not weak, it could have been stronger if the submitted texts contained 6000 words at the minimum – for a more high-quality assessment of an individual's personality.

Sustained increase in the use of work groups require strategies particularly designed to effectively select group members [38]. Automated strategies manage the complexity of today's data more accurately and effectively than manual methods such as creativity tests. Although our presented approach – assuming creativity tests can be entirely replaced by automated personality mining tools – could not be backed up by our pilot study, we strongly believe that future studies with a bigger sample size would give more useful insights as to how variables of creativity tests and Big Five personality traits are linked with each other. Our key contribution in this paper, however, was to propose the usage of cognitive systems and specifically automated personality mining systems in order to identify the Big Five personality traits and to be then able to enhance team composition with the help of this information.

Despite our suggestion of the deployment of smart machines for team composition, we are aware that, given the current state of team composition literature, it is still yet unclear which specific configurations of the Big Five personality traits can be used at the design stage of teams in order to increase group work performance [7]. Predicting "perfect" groups may not be possible at the present time, a reduction of group work ineffectivity however is. A future goal of cognitive computing is thus to foster human-machine interaction, by building cognitive capabilities into many different applications and systems [28]. For example, personality mining systems could assist human resources departments within a recruitment process as effective pre-screening tools by searching specifically for creative potential and by delivering personality profiles of suitable candidates. Current cognitive systems such as IBM Watson's PI service are still in the early days of technological evolution. But as we can see, the potential of cognitive computer systems is not nearly fully exhausted and yet to be researched, especially in regard to team composition.

5. References

[1] Allport, G.W. Pattern and growth in personality. (1961).

[2] Amabile, T.M. A model of creativity and innovation in organizations. *Research in organizational behavior 10*, 1 (1988), 123–167.

[3] Baer, J. Commentary: Divergent thinking tests have problems, but this is not the solution. (2008).

[4] Baker, D.P. and Salas, E. Principles for measuring teamwork: A summary and look toward the future. *Team performance assessment and measurement: Theory, methods, and applications*, (1997), 331–355.

[5] Barrick, M.R., Stewart, G.L., Neubert, M.J., and Mount, M.K. Relating member ability and personality to work-team processes and team effectiveness. *Journal of applied psychology 83*, 3 (1998), 377.

[6] Barron, F. and Harrington, D.M. Creativity, intelligence, and personality. *Annual review of psychology 32*, 1 (1981), 439–476.

[7] Bell, S.T. Deep-level composition variables as predictors of team performance: a meta-analysis. *Journal of applied psychology 92*, 3 (2007), 595.

[8] Bradley, J.H. and Hebert, F.J. The effect of personality type on team performance. *Journal of Management Development 16*, 5 (1997), 337–353.

[9] Chamorro-Premuzic, T. and Reichenbacher, L. Effects of personality and threat of evaluation on divergent and convergent thinking. *Journal of Research in Personality 42*, 4 (2008), 1095–1101.

[10] Costa, P.T. and McCrae, R.R. The revised neo personality inventory (neo-pi-r). *The SAGE handbook of personality theory and assessment 2*, (2008), 179–198.

[11] Creswell, J.W., Plano Clark, V.L., Gutmann, M.L., and Hanson, W.E. Advanced mixed methods research designs. *Handbook of mixed methods in social and behavioral research*, (2003), 209–240.
[12] Cropley, A. In praise of convergent thinking. *Creativity research journal 18*, 3 (2006), 391–404.
[13] Cropley, A.J. Recognizing Creative Potential: An Evaluation of the Usefulness of Creativity Tests. *High ability studies 7*, 2 (1996), 203–219.
[14] Devine, D.J. and Philips, J.L. Do smarter teams do better a meta-analysis of cognitive ability and team performance. *Small Group Research 32*, 5 (2001), 507–532.
[15] Dow, G.T. and Mayer, R.E. Teaching students to solve insight problems: Evidence for domain specificity in creativity training. *Creativity Research Journal 16*, 4 (2004), 389–398.
[16] Facaoaru, C. *Kreativität in Wissenschaft und Technik: Operationalisierung von Problemlösefähigkeiten u. kognitiven Stilen*. Huber, 1985.
[17] Faliagka, E., Tsakalidis, A., and Tzimas, G. An integrated e-recruitment system for automated personality mining and applicant ranking. *Internet research 22*, 5 (2012), 551–568.
[18] Feist, G.J. A meta-analysis of personality in scientific and artistic creativity. *Personality and social psychology review 2*, 4 (1998), 290–309.
[19] Ferrucci, D.A. Introduction to "this is watson." *IBM Journal of Research and Development 56*, 3.4 (2012), 1–1.
[20] George, J.M. and Zhou, J. When openness to experience and conscientiousness are related to creative behavior: an interactional approach. *Journal of applied psychology 86*, 3 (2001), 513.
[21] Golbeck, J., Robles, C., Edmondson, M., and Turner, K. Predicting personality from twitter. *Privacy, Security, Risk and Trust (PASSAT) and 2011 IEEE Third Inernational Conference on Social Computing (SocialCom), 2011 IEEE Third International Conference on*, IEEE (2011), 149–156.
[22] Guilford, J.P. Creativity. *American Psychologist 5*, 9 (1950), 444–454.
[23] Guilford, J.P. *The nature of human intelligence*. McGraw-Hill, New York, NY, US, 1967.
[24] Halfhill, T., Sundstrom, E., Lahner, J., Calderone, W., and Nielsen, T.M. Group personality composition and group effectiveness: An integrative review of empirical research. *Small group research 36*, 1 (2005), 83–105.
[25] Hall, C.S. and Lindzey, G. Theories of personality. (1957).
[26] Hogan, J. and Holland, B. *Using theory to evaluate personality and job-performance relations: a socioanalytic perspective.* American Psychological Association, 2003.
[27] Hurwitz, J., Kaufman, M., and Bowles, A. *Cognitive computing and big data analytics*. John Wiley & Sons, 2015.
[28] IBM Watson. Personality Insights. Personality models. IBM Watson Developer Cloud. 2017. https://www.ibm.com/watson/developercloud/doc/personality-insights/models.html#outputBigFive.
[29] IBM Watson. Personality Insights v3 API Explorer. 2017. https://watson-api-explorer.mybluemix.net/apis/personality-insights-v3.
[30] IBM Watson. Personality Insights. IBM Watson Bluemix. IBM Watson Developer Cloud. 2017. https://personality-insights-livedemo.mybluemix.net/.
[31] IBM Watson. Personality Insights. Usage guidance. IBM Watson Developer Cloud. 2017. https://www.ibm.com/watson/developercloud/doc/personality-insights/guidance.html.
[32] IBM Watson. Personality Insights. Overview for users. BM Watson Developer Cloud. 2017. https://www.ibm.com/watson/developercloud/doc/personality-insights/user-overview.html.
[33] IBM Watson. Personality Insights. The science behind the service. IBM Watson Developer Cloud. https://www.ibm.com/watson/developercloud/doc/personality-insights/science.html#science.
[34] Judge, T.A., Heller, D., and Mount, M.K. *Five-factor model of personality and job satisfaction: a meta-analysis.* American Psychological Association, 2002.
[35] King, L.A., Walker, L.M., and Broyles, S.J. Creativity and the five-factor model. *Journal of research in personality 30*, 2 (1996), 189–203.
[36] Klimoski, R. and Zukin, L. Selection and staffing for team effectiveness. *Supporting work team effectiveness*, (1999), 63–94.
[37] Lancaster, G.A., Dodd, S., and Williamson, P.R. Design and analysis of pilot studies: recommendations for good practice. *Journal of Evaluation in Clinical Practice 10*, 2 (2004), 307–312.
[38] Lawler, E.E., Mohrman, S.A., and Ledford, G.E. Strategies for high performance organizations: Employee involvement, TQM, and reengineering programs in Fortune 1000 corporations. *San Francisco: Josey-Bass*, (1998).
[39] LePine, J.A. Team adaptation and postchange performance: effects of team composition in terms of members' cognitive ability and personality. *Journal of Applied Psychology 88*, 1 (2003), 27.
[40] Lim, B.-C. and Ployhart, R.E. Transformational leadership: relations to the five-factor model and team performance in typical and maximum contexts. *Journal of Applied Psychology 89*, 4 (2004), 610.
[41] Lonergan, D.C., Scott, G.M., and Mumford,

M.D. Evaluative aspects of creative thought: Effects of appraisal and revision standards. *Creativity Research Journal 16*, 2–3 (2004), 231–246.
[42] Martins, L.L., Gilson, L.L., and Maynard, M.T. Virtual teams: What do we know and where do we go from here? *Journal of management 30*, 6 (2004), 805–835.
[43] Matz, S., Chan, Y.W.F., and Kosinski, M. Emotions and Personality in Personalized Services. In M. Tkalčič, B. De Carolis, M. de Gemmis, A. Odić and A. Košir, eds., *Human–Computer Interaction Series*. Springer Switzerland, 2016, 35–54.
[44] McCrae, R.R. Creativity, divergent thinking, and openness to experience. *Journal of Personality and Social Psychology 52*, 6 (1987), 1258–1265.
[45] McCrae, R.R. and John, O.P. An introduction to the five-factor model and its applications. *Journal of personality 60*, 2 (1992), 175–215.
[46] Moreland, R.L., Levine, J., and Wingert, M. Creating the ideal group: Composition effects at work. *Understanding group behavior 2*, (2013), 11–35.
[47] Moreland, R.L. and Levine, J.M. Problem identification by groups. *Group process and productivity*, 9 (1992), 17–47.
[48] Moutafi, J., Furnham, A., and Tsaousis, I. Is the relationship between intelligence and trait Neuroticism mediated by test anxiety? *Personality and Individual Differences 40*, 3 (2006), 587–597.
[49] Mumford, J.A., Turner, B.O., Ashby, F.G., and Poldrack, R.A. Deconvolving BOLD activation in event-related designs for multivoxel pattern classification analyses. *NeuroImage 59*, 3 (2012), 2636–2643.
[50] Nemeth, C.J. Minority dissent as a stimulant to group performance. *Group process and productivity 95*, (1992), 111.
[51] Nobre, F.S., Tobias, A.M., and Walker, D.S. The impact of cognitive machines on complex decisions and organizational change. *AI & society 24*, 4 (2009), 365–381.
[52] Organ, D.W. and Lingl, A. Personality, satisfaction, and organizational citizenship behavior. *The journal of social psychology 135*, 3 (1995), 339–350.
[53] Organ, D.W. and Ryan, K. A meta-analytic review of attitudinal and dispositional predictors of organizational citizenship behavior. *Personnel psychology 48*, 4 (1995), 775–802.
[54] Ozer, D.J. and Benet-Martinez, V. Personality and the prediction of consequential outcomes. *Annu. Rev. Psychol. 57*, (2006), 401–421.
[55] Pennebaker, J.W., Boyd, R.L., Jordan, K., and Blackburn, K. *The development and psychometric properties of LIWC2015*. 2015.
[56] Pennington, J., Socher, R., and Manning, C.D. Glove: Global Vectors for Word Representation. *EMNLP*, (2014), 1532–1543.
[57] Reilly, R.R., Lynn, G.S., and Aronson, Z.H. The role of personality in new product development team performance. *Journal of Engineering and Technology Management 19*, 1 (2002), 39–58.
[58] Rhodes, M. An Analysis of Creativity. *The Phi Delta Kappan 42*, 7 (1961), 305–310.
[59] Roberts, M.J. The relationship between extraversion and ability. *Personality and Individual Differences 32*, 3 (2002), 517–522.
[60] Ryckman, R.M. *Theories of personality*. Cengage Learning, 2012.
[61] Shen, S.-T., Prior, S.D., White, A.S., and Karamanoglu, M. Using personality type differences to form engineering design teams. *Engineering Education 2*, 2 (2007), 54–66.
[62] Staiano, J., Lepri, B., Aharony, N., Pianesi, F., Sebe, N., and Pentland, A. Friends don't lie: inferring personality traits from social network structure. *Proceedings of the 2012 ACM conference on ubiquitous computing*, ACM (2012), 321–330.
[63] Sundstrom, E., De Meuse, K.P., and Futrell, D. Work teams: Applications and effectiveness. *American psychologist 45*, 2 (1990), 120.
[64] Tausczik, Y.R. and Pennebaker, J.W. The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of language and social psychology 29*, 1 (2010), 24–54.
[65] Van Knippenberg, D. and Schippers, M.C. Work group diversity. *Annu. Rev. Psychol. 58*, (2007), 515–541.
[66] Wallach, M.A. and Kogan, N. Modes of thinking in young children: A study of the creativity-intelligence distinction. (1965).
[67] Weerawardena, J. and Mavondo, F.T. Capabilities, innovation and competitive advantage. *Industrial Marketing Management 40*, 8 (2011), 1220–1223.
[68] Yang, S. and Choi, S.O. Employee empowerment and team performance: Autonomy, responsibility, information, and creativity. *Team Performance Management: An International Journal 15*, 5/6 (2009), 289–301.
[69] Yarkoni, T. Personality in 100,000 words: A large-scale analysis of personality and word use among bloggers. *Journal of research in personality 44*, 3 (2010), 363–373.