Predicting programmers' personality via interaction behaviour with keyboard and mouse

Iftikhar A Khan, Osman Khalid, Waqas Jadoon, Rafi Us. Shan, Abdul Nasir

This exploratory research was conducted to study the relationship between Big Five personality measurement scale and the interaction behaviour of the programmers with keyboard and mouse to examine the possibility of creating a computer based objective personality measurement scale. A field study with two analyses (N = 20), (N = 12) and a lab study (N = 15) were conducted where N is the number of participants who participated in the research. In the field study, interaction data were collected during normal PC use over several days. In the laboratory study, participants worked on a programming task while their interaction with keyboard and mouse was being recorded in background. All of the participants rated their personality online and uploaded the data for analysis. Results found inconsistent behaviour of all of the personality traits except ‘activity level’ in all studies and thus suggested that a programmer’s ‘activity level’ can be predicted from his/her interaction behaviour with keyboard and mouse. This prediction will help in differentiating good programmers from not so good programmers objectively.
Predicting Programmers' Personality via Interaction Behaviour with Keyboard & Mouse

Iftikhar Ahmed Khan, Osman Khalid, Waqas Jadoon, Rafi Us Shan, Abdul Nasir Khan

(iftikharahmed, osman, waqas_Jadoon, shan, anasir)@ciit.net.pk

COMSATS Institute of Information Technology, TOBE Camp Abbottabad, Pakistan

ABSTRACT

This exploratory research was conducted to study the relationship between Big Five personality measurement scale and the interaction behaviour of the programmers with keyboard and mouse to examine the possibility of creating a computer based objective personality measurement scale. A field study with two analyses (N = 20), (N = 12) and a lab study (N = 15) were conducted where N is the number of participants who participated in the research. In the field study, interaction data were collected during normal PC use over several days. In the laboratory study, participants worked on a programming task while their interaction with keyboard and mouse was being recorded in background. All of the participants rated their personality online and uploaded the data for analysis. Results found inconsistent behaviour of all of the personality traits except ‘activity level’ in all studies and thus suggested that a programmer’s ‘activity level’ can be predicted from his/her interaction behaviour with keyboard and mouse. This prediction might help in differentiating good programmers from not so good programmers objectively.

Keywords: Programmers, Personality, Keyboard, Mouse, Interaction, Behaviour, Estimation, Measurement

1. INTRODUCTION

The study of personality is one of the most active areas in the field of psychology (Burger, 2010). One of the most agreed definition of personality is: “consistent behaviour patterns and intrapersonal processes originating within the individual” (Burger, 2010). From definition it can inferred that people have different personalities like some people make friends easily (extroverts) others prefer to be alone (introverts). Similarly, some people are more business oriented while others are service oriented. Therefore, studying personality is significant to identify the differences among people behaviours. This implies that studying various human characteristics can help other...
humans to decide how to deal with a particular personality characters (Burger, 2010). Moreover, these characteristics can help in marketing products (Nunes et al., 2012), finding effects of personality on peoples’ own economy (Phau et al., 2009) and finding the relation between personality and health (McCool, 2008) etc. In order to improve the human-human or human-machine interaction as well as to enhance the productivity, the knowledge about various personality features has gained much significance.

Various personality measurement instruments like (Butcher, 1989), (McCrae and John, 1992), (Cattell, 1989) etc. can be used to collect knowledge about the personality. Personality is a combination of traits and states that can be compared with already formulated personality categories (Cohen et al., 2012). There are various techniques employed by researchers to measure personality, such as MMPI-2 (Butcher, 1989), Big Five IPIP (McCrae and John, 1992), 16 PF (Cattell, 1989), etc. Readers can find a detailed overview about personality measurement in (Cohen et al., 2012) and (Boyle, 2008). All of these measuring instruments require self-reporting. There are various versions of self-reporting scales like the one which require respondents to rate more items and some which require to rate less items. The more items a scale contain, more reliable will be the results. However, several studies showed that shorter versions were highly correlated with full versions (Donnellan et al., 2006). Shorter versions tend to take less time as compared to the versions which have more items. However, even shorter versions require allocation of time on part of the participants. In addition, self-reports pose various validity and reliability problems (Stone et al., 1999).

An alternative solution is automated measurement of personality that automatically measures a user’s personality while doing his/her routine tasks. There is very little research on this approach. Some examples include Heckmann who designed General User Model Ontology (GUMO) to store users’ profile and to represent their psychological information through different kind of technologies like semantic web and userML (User Model Mark-up Language) (Heckmann, 2005). Personality recognizer from linguistic and conversational cues were also introduced (Mairesse et al., 2007). There is also research available that adopted the personality measurement scale on the computers (offline as well as online) but participants still needed to self-report these instruments. An example of such adopted instrument is a website by Goldberg and Saicoer.

1 http://ipip.ori.org/
This research adopted the automated instrument to develop a behaviour based measurement of the personality. This research supports the notion that personality can be expressed from behaviour. For example, Brebner et al. (Brebner, 1983) showed that a change in visual stimulation caused a higher level of motor activity in extroverts. Geen et al. (Geen, 1984) noted that extroverts prefer more intense level of music as compared to introverts in a learning task. Extroverts interaction time with the interfaces were quicker as compared to introverts (Saati et al., 2005). In this paper, we explore the possibility that such interactive behaviours might also be visible through a keyboard and mouse. As compared to other personality measurement techniques, a dispositional approach is adopted to measure personality via usage of keyboard and mouse along with “Big Five” model to describe personality characteristics.

Empirical studies and analysis performed on the gathered data indicated that it is possible to predict some of participants’ personality traits from the interactive behaviour of computer programmers with keyboard and mouse and thus forming the main scientific contribution of this paper. The rest of the paper is organized as follows. In Section 2, we present the state of the art literature review. A brief overview to Big Five Personality Measurement Model is provided in Section 3. Section 4 presents the field study we conducted in our work, whereas in Section 5, we present the conducted laboratory study. Section 6 presents active windows analysis, and a discussion is provided about programmer’s activity level as well as the proposed algorithm in Section 7. Finally, Section 8 presents the limitations of the proposed work along with conclusions and future work.

2. RELATED WORK

The task of personality measurement is complex. Various personality measurement scales exist, such as MMPI-2 (Butcher, 1989), Big Five IPIP (McCrae and John, 1992), 16 PF (Cattell, 1989). Such scales measure personality from a particular dimension. For example, in order to assess biological-based traits, various electronic equipment like EEG or GSR (Heller, 1993) is used. For subjective measurement of personality, subjective self-reporting scales like Big Five Factor models (McCrae and John, 1992) are used. Alternatively, for objective measurement of personality, objective scales like Objective Analytic Battery (OAB) (Cattell, 1955) is used. Additionally, some scales, like multidimensional personality questionnaire is used to improve multicultural effectiveness of the scales to measure personality (Van Der Zee and Van Oudenhoven, 2000). All of these approaches are summarized under approaches like
psychoanalytic approach, the Freudian approach, the humanistic approach, the cognitive approach as well as Behavioural/Social approach and have been discussed in detail by Burger (Burger, 2010).

Personality depicts the habits of living beings specifically humans (Burger, 2010). This means that people learn to behave and respond in a certain way and their responses for most of the times is predictable. As discussed before, psychologists have developed various instruments to measure these behaviours. These behaviours are then matched with a standard set of agreed upon behaviours and are termed as personality (Cohen, 1988). However, significant time and effort is required towards administering and managing these scales (Boyle, 2008). In addition, according to Boyle (Boyle, 2008), there are very limited actual tests of personality and most of the scales according to him are “self-report scales or reports of others’ rating scales”.

To overcome difficulties with self-reporting (as discussed in introduction) and to reduce management and administrative issues, researchers devised various methods. In the present era of computing, various researchers diverted their efforts towards adoption of personality measurement scale for computers. Some studies, such as (Nunes et al., 2012), attempted to standardize the personality information across applications. The aforementioned studies took human psychological aspects into consideration while making decisions by the use of XML (Extensible Mark-up Language). In addition Zhang et al. (Liang et al., 2012) presented a framework to synthesize 2D cartoons and to depict their personality. People now are also able to take paper less personality tests via website\(^2\). Such methods of evaluation might have reduced the managerial and administrative issues, however, these tests are still based on self-reporting. Another issue is of data privacy and handling as data access and control is not in the hand of every researcher and is mostly restricted to the organization who own these scales. The researchers using such scale need to convert textual results into their own format in order to do any analysis on the data. To overcome these problems a more autonomous approach to measure personality is required.

Various automating activities related to personality with the help of computers exist. For example, an interactive story telling system devised by Su et al. (Su et al., 2005) utilized hierarchical fuzzy rule based system to facilitate the personality and emotion control of the dynamic story character.

\(^2\) [http://ipip.ori.org/](http://ipip.ori.org/)
Similarly, Chittaro and Serra [23] presented a system for realistic behavioural programming of virtual characters and based the characters’ programming on personality and probabilistic automata (Chittaro and Serra, 2004). Various researchers such as [24], experimented on imparting personality on virtual computer generated characters (Kshirsagar, 2002). One the same note, Rousseau et al. [25] proposed a model to be used by intelligent automated actors. These actors in turn were able to improvise their behaviour to interact with the users in a multimedia environment (Rousseau, 1996). There is also research on personality and how it is related to online buying behaviour (Bosnjak et al., 2007). For instance, the work presented in [26] imparts various behaviours and personalities to computer characters by using different computational and probabilistic models. However no research in authors’ knowledge addresses autonomous objective prediction of personality of computer users in general and specifically programmers.

Programmers and computer users’ personality is being researched to improve their productivity and to find right person for the right job. For example, Capretz (Capretz, 2003) reviewed empirical studies on personality types of software engineers. Capretz concluded that software engineers are significantly related to various personality types and are more likely to be sensing and thinkers (STs), or Thinking and Judgmental (TJs), or Intuitive and thinkers (NTs) (Capretz, 2003). Software professionals should be assigned roles from which it is evident that they have suitable personality to carry out such tasks (Acuna et al., 2006), (Da Cunha and Greathead, 2007). Another example of software professional personality should be considered while they work comes from (Salleh et al., 2010) which found personality to be significantly related with pair programming students’ academic performance (Salleh et al., 2010).

In our proposed work, we analyse the behaviour of users on most commonly used devices, such as keyboard and mouse. For computer users generally and for software programmers specifically the keyboard and mouse are the most common devices for input and output. The keyboard and mouse behaviour is already used by various researchers to predict moods of computer users. For instance, in (Khan et al., 2013) and (Epp et al., 2011) the authors argued that it is possible to predict some of the personality traits of programmers by recording their behaviour with keyboard and mouse. In this paper, we utilized the Big Five Model of personality measurement. The next section discussed in details the Big Five Model and the reasons why we selected this personality measurement scale.
3. **BIG FIVE PERSONALITY MEASUREMENT MODEL**

Although there is a difference of opinion among researchers in how to measure and what to measure in a person’s personality, but still it is agreed that factor analytic properties of personality are consistent (Burger, 2010). Various investigations on dynamic data like: (Digman, 1990), (Goldberg, 1992) and (McCrae and Costa Jr, 1997) found that five basic dimensions of personality are more or less consistent. These five dimensions of personality are so extensively used by researchers that now it is being referred to as Big Five Model of personality measurement (Burger, 2010), (Boyle, 2008). The most commonly used names of the five dimensions are Neuroticism, Extraversion, Openness, Agreeableness, and Conscientiousness. As the authors Boyle et al. (Boyle, 2008) stated that in the construction of the Big Five scale, researches have not only utilized the theory but also used empirical factor analytic inputs.

Such approach of defining traits is also the most common approach implemented by computing scientists to measure personality via computers (Nunes et al., 2012). According to Nunes et al., trait approach highlights clear psychological differences among individuals and is easier to categorize, organise, and to implement. Because of its wider and universal use, we considered Big Five personality model in our paper for measuring personality of the participants.

Big Five personality measure contains five main traits and each main trait is further divided into six sub-trait thus forming a total of 35 traits including the main trait. This specific 35 trait personality questionnaire is called as NEO PI-R, a revised version by Costa and McCrae (Costa and MacCrae, 1992). The original scale contains 300 items and take on average 30-40 minutes to complete the test. The short version of scale contains 120 items and could take 10-20 minutes to complete. The original scale is more reliable than the shorter one, though, shorter version also meets professional standards of reliability (Costa and MacCrae, 1992). The test can be taken at the following website:\(^3\) IPIP Big Five scale main traits, sub-traits, and few of the characteristics are shown in the Table 1 below.

### Table 1: Big Five personality traits, sub-traits, and some of the related characteristics adopted from Burger (Burger, 2010)

---

\(^3\) [http://www.personal.psu.edu/faculty/j/5/j5j/IPIP/](http://www.personal.psu.edu/faculty/j/5/j5j/IPIP/)
<table>
<thead>
<tr>
<th>Main Trait</th>
<th>Sub-trait</th>
<th>Some Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extraversion</td>
<td>Friendliness</td>
<td>Sociable versus retiring, Fun-loving versus sober</td>
</tr>
<tr>
<td></td>
<td>Gregariousness</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Assertiveness</td>
<td>Affectionate versus reserved, Energetic versus lazy,</td>
</tr>
<tr>
<td></td>
<td>Activity Level</td>
<td>Hurried versus slow, Aggressive versus ambitionless,</td>
</tr>
<tr>
<td></td>
<td>Excitement Seeking</td>
<td>Active versus inactive</td>
</tr>
<tr>
<td></td>
<td>Cheerfulness</td>
<td></td>
</tr>
<tr>
<td>Agreeableness</td>
<td>Trust</td>
<td>Soft-hearted versus ruthless</td>
</tr>
<tr>
<td></td>
<td>Morality</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Altruism</td>
<td>Trusting versus suspicious</td>
</tr>
<tr>
<td></td>
<td>Cooperation</td>
<td>Helpful versus uncooperative</td>
</tr>
<tr>
<td></td>
<td>Modesty</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sympathy</td>
<td></td>
</tr>
<tr>
<td>Conscientiousness</td>
<td>Self-Efficacy</td>
<td>Well organized versus disorganized</td>
</tr>
<tr>
<td></td>
<td>Orderliness</td>
<td>Careful versus careless</td>
</tr>
<tr>
<td></td>
<td>Dutifulness</td>
<td>Self-disciplined versus weak willed</td>
</tr>
<tr>
<td></td>
<td>Achievement Striving</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Self-Discipline</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cautiousness</td>
<td></td>
</tr>
<tr>
<td>Neuroticism</td>
<td>Anxiety</td>
<td>Worried versus calm</td>
</tr>
<tr>
<td></td>
<td>Anger</td>
<td>Insecure versus secure</td>
</tr>
<tr>
<td></td>
<td>Depression</td>
<td>Self-pitying versus self-satisfied</td>
</tr>
<tr>
<td></td>
<td>Self-Consciousness</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Immoderation</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Vulnerability</td>
<td></td>
</tr>
<tr>
<td>Openness to</td>
<td>Imagination</td>
<td>Imaginative versus down-to-earth</td>
</tr>
<tr>
<td>Experience</td>
<td>Artistic Interests</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Emotionality</td>
<td>Preference for variety versus preference for routine</td>
</tr>
<tr>
<td></td>
<td>Adventurousness</td>
<td>Independent versus conforming</td>
</tr>
<tr>
<td></td>
<td>Intellect</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Liberalism</td>
<td></td>
</tr>
</tbody>
</table>

These studies involve collection of the participants’ personality as well as their interactive behaviour with keyboard and mouse. The next sections will present two studies and three analyses performed in this paper. The results are presented in the corresponding study sections.
4. STUDY I – FIELD STUDY

4.1. Material

Most of the material collected and used in this study was also used in (Khan et al., 2013) to measure mood of computer users from keyboard and mouse interaction. The basic idea of the experiment was to record all keyboard and mouse related interaction of a user in log files. Therefore, we developed a software application to record all the user interaction into log files. A mood rating dialog was also designed to appear after every 20 minutes of logging to get participants’ mood response. However, to reduce the concerns of participants over their privacy and security issues various functionalities were developed in the application, such as:

Pause logging: Participants were able to pause application from logging for 5 and 10 minutes as well as for custom specified period (in minutes) by clicking on the relevant menu. This option was helpful for participants who wanted to do private tasks like bank transaction, etc.

Exit logging: Participants were also able to exit the application. Participants were instructed to restart the application when they feel so. They were also told that application will restart itself automatically on next computer boot.

Categorical Storage of keys: Some keys of the keyboard were divided into 4 categories namely: ‘capital alphabets’, ‘lowercase alphabets’, ‘numeric keys’ and ‘special characters’. The keyboard interaction were classified and recorded according to one of the categories instead of recording original keys. In addition, mouse button events were recorded ‘Right Button’ or ‘Left Button’ depending on the right or left button click event. This mechanism further built the trust of participants to continue using this application over several days.

4.2. Experimental setup

Departmental ethics committee namely CS-DARC (Computer Science Department Academic Review Committee) reviewed and approved this experiment. A written consent for all studies was obtained from the participants. The procedure of taking consent this way was approved by CS-DARC. The developed application is run as a background application on the participants’ machines. The categorized keyboard interaction data described earlier were recorded under the column ‘Category of the event’. In addition, the data for three more columns was recorded namely
as: ‘Window name’, ‘Keyboard and Mouse Event’, and ‘Date/Time of the event’. Further
description of these three columns is as follows:

Window name: As modern day operating systems are designed to run different windows
(applications) at a time, the window name field was used to record active window. Names of the
document under work were not recorded. Only general software names under the use were
recorded like ‘Microsoft Power point’, ‘Mozilla Firefox’, etc. The data was recorded based on the
assumption that a person’s behaviour with the usage of keyboard and mouse will be different for
different applications. For example, frequent events of keyboard are expected while using an
editing application, chatting application or Integrated Development Environment (IDE) used by a
programmer during process of coding. Similarly, frequent mouse clicks are expected on a mouse
driven games like age of empires and rise of nations, etc.

Keyboard and Mouse event: When a key is pressed it initiates a ‘key down’ event. When released,
a ‘key up’ event is initiated. Similarly, ‘mouse click’ event is initiated on mouse button click. Such
data was used to record the three events: ‘Key Up’, ‘Key Down’ as well as mouse clicks (‘Left’ or
‘Right’) button.

Date/Time of event: This column was used to record the date and exact time up to milliseconds
on the initiation of an event using the system date and time functionality.

A mood dialogue (Khan et al., 2013) was embedded in the log files as a marker to measure an
interaction. Application also provided the choice to participants to uninstall the application and zip
the log files, or to continue the experiment. Short form of Big Five IPIP-NEO⁴ (Goldberg, 1999)
instrument containing 120 items were used in this research. Participants were instructed to
complete the personality questionnaire online on IPIP website and email the results to the
researcher.

4.3 Participants

A total of 26 computer users participated in the mood study but only 20 consented to take
personality test as well. Therefore this study considered response of only 20 participants in total.
All of the participants were invited via personal contact or emails. The mean age of the
participants’ was 28.5 years and a standard deviation of 2.9 years. Participants’ computer
experience mean was 5 years and standard deviation of 1.6 years. When asked from participants how they rate themselves from the options as programmer, expert computer user, and a novice, 50% of them rated themselves as programmers. Expert computer users were 40% and 10% were novices or beginners. There were only two females who participated in this study.

4.4. Results

4.4.1 Data preparation

Each participant continued logging for at least 5 days. Logging application was designed to create a separate log file for each day with date and time stamp. All these file were merged together to form a single file for a participant thus forming 20 different log files related to 20 participants. On average, about a million lines of interaction of a participant with keyboard and mouse were recorded. Therefore, a separate application was required to extract meaningful data from these files. First step was to find a meaningful point around which data is analysed. This meaningful point was identified as ‘mood rating’ embedded via mood rating dialog (Khan et al., 2013) in to the log files. Average of behavioural data was calculated around every mood rating window\(^4\). An interaction data of 3 minutes before mood rating and 3 minutes after mood rating were taken to form a 6 minute window. The basic measures taken for each window were:

**Self-reported mood:** This was embedded in the log file. In the analysis part, this measure was excluded because of no relevance to the study.

1. **Total number of events (nEvents):** Total events that took place within 6 minute window.
2. **Average time between events (ATBE):** Average of the time difference between occurrence of two consecutive events
3. **Standard deviation time between events (ATBE-STD):** The standard deviation of ATBE
4. **Total windows switched (WS):** A switch by the participants from one window to another window.
5. **Number of backspace and delete key events (NBD):** The events counter for only backspace and delete keys. This is the counter of the possible errors committed by a user and the undo action performed for those errors.

\(^4\) A window was formed around embedded mood rating in the log file and the interaction data before and after it
6. **Numeric and alphabet keys (NAK):** Counter for numeric and alphabet keys

7. **Mouse clicks (MC):** Counter for total number of clicks (including right and left button clicks)

8. **Other keys (OK):** All other keys that include special keys, arrow keys, semi colons, colons etc.

Each participant’s average of the above measures and that of his/her rated personality was put into an excel file for further refinement. There were $N$ rows of entries for each participant’s interaction data, where $N$ is the number of mood rating dialogs answered by participants. There was only 1 row of personality traits. Before averaging a participant’s interaction data over $N$ rows, all the events where interaction was less than 10 events in a window were filtered out. The reason was that the participants might not be actively using keyboard and mouse but instead reading from some website or busy in some other tasks. All events where difference of time between two events was less than 50 milliseconds were also eliminated. The reason being that an experienced typist on average cannot type a meaningful character of a sequence in less than 60 milliseconds (Card et al., 1983). Considering an experienced typist, we reduced filter to 50 milliseconds just in case that someone among the participants could be an experienced typist. In addition, all the events with a difference of time more than 20 seconds were also eliminated on the same grounds of inactivity. After required filtration of each participant data, all the interactions rows were averaged to a single row. In addition to means, standard deviation of each interaction variables was also calculated.

### 4.4.2. Analysis

First step toward analysis of this data was to carry out correlations analysis. As aim of this study is to find out the relation between keyboard and mouse interaction behaviour and personality traits. Therefore, a correlation analysis was conducted between interaction behaviour variables and personality trait variables. The results revealed that various traits and sub-traits are significantly correlated with interaction behaviour of participants with keyboard and mouse. The results are presented from Table 2–Table 7 below.

**Table 2:** Correlations between Extroversion and its sub-traits with that of keyboard and mouse behaviour

<table>
<thead>
<tr>
<th>Interaction Variables</th>
<th>Gregariousness</th>
<th>Activity Level</th>
<th>Excitement Seeking</th>
<th>Cheerfulness</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATBE-Mean</td>
<td>-0.44*</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Table 3: Correlations between Agreeableness and its sub-traits with that of keyboard and mouse behaviour

<table>
<thead>
<tr>
<th>Sub-traits</th>
<th>Agreeableness</th>
<th>Trust</th>
<th>Morality</th>
<th>Cooperation</th>
<th>Sympathy</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATBE-Mean</td>
<td>0.41*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WS-Mean</td>
<td>-0.56**</td>
<td>-0.56**</td>
<td></td>
<td>-0.52*</td>
<td></td>
</tr>
<tr>
<td>WS-Standard Deviation</td>
<td></td>
<td></td>
<td></td>
<td>-0.48*</td>
<td></td>
</tr>
<tr>
<td>NBD-Mean</td>
<td>-0.51*</td>
<td></td>
<td></td>
<td>-0.43*</td>
<td></td>
</tr>
<tr>
<td>NBD-Standard Deviation</td>
<td>-0.40*</td>
<td>-0.40*</td>
<td></td>
<td>-0.40*</td>
<td></td>
</tr>
<tr>
<td>NAK-Mean</td>
<td></td>
<td></td>
<td>-0.45*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NAK-Standard Deviation</td>
<td>0.47*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Note:** ** significant Correlation at 0.01 levels, * significant Correlation at 0.05 levels for N = 20.

### Table 4: Correlations between Conscientiousness and its sub-traits with that of keyboard and mouse behaviour

<table>
<thead>
<tr>
<th>Sub-traits</th>
<th>Conscientiousness</th>
<th>Self-Efficacy</th>
<th>Orderliness</th>
<th>Dutifulness</th>
<th>Cautiousness</th>
</tr>
</thead>
<tbody>
<tr>
<td>nEvents-Mean</td>
<td></td>
<td>-.60**</td>
<td>-.46*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WS-Mean</td>
<td></td>
<td></td>
<td></td>
<td>-.42*</td>
<td></td>
</tr>
<tr>
<td>NAK-Mean</td>
<td></td>
<td>.46*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MC–Mean</td>
<td></td>
<td>.44*</td>
<td>.52*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MC–Standard Deviation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OK-Mean</td>
<td>-.40*</td>
<td>-.43*</td>
<td>-.55**</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Note:** ** significant Correlation at 0.01 levels, * significant Correlation at 0.05 levels for N = 20.
Table 5: Correlations between Neuroticism and its sub-traits with that of keyboard and mouse behaviour

<table>
<thead>
<tr>
<th>Neuroticism</th>
<th>Anxiety</th>
<th>Self-Consciousness</th>
<th>Vulnerability</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATBE-Mean</td>
<td>.43*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WS-Mean</td>
<td>.50*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NAK-M</td>
<td>.51*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MC-M</td>
<td>-.40*</td>
<td>-.53*</td>
<td>-.41*</td>
</tr>
<tr>
<td>MC-Standard Deviation</td>
<td>-.40*</td>
<td>-.40*</td>
<td>-.51*</td>
</tr>
</tbody>
</table>

Note: ** significant Correlation at 0.01 levels, * significant Correlation at 0.05 levels for N = 20.

Table 6: Correlations between Openness to experience and its sub-traits with that of keyboard and mouse behaviour

<table>
<thead>
<tr>
<th>Openness to experience</th>
<th>Artistic Interests</th>
<th>Emotionality</th>
<th>Intellect</th>
<th>Liberalism</th>
</tr>
</thead>
<tbody>
<tr>
<td>nEvents-Mean</td>
<td>.43*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATBE-Standard Deviation</td>
<td>.41*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WS-Mean</td>
<td>-.40*</td>
<td>-.54*</td>
<td>-.60**</td>
<td></td>
</tr>
<tr>
<td>WS-Standard Deviation</td>
<td>-.48*</td>
<td>-.62**</td>
<td>-.58**</td>
<td></td>
</tr>
</tbody>
</table>

Note: ** significant Correlation at 0.01 levels, * significant Correlation at 0.05 levels for N = 20.

Correlations are considered of small, medium, and large size with $r = 0.1$, $r = 0.3$ to 0.49, and $r = 0.5$ or above, respectively (Cohen, 1992). According to the aforementioned criterion, all correlations revealed in Tables 2–7 either have medium or large effects.

5. **STUDY II – LAB STUDY**

Various personality traits were found to be significantly correlated with the keyboard and mouse interaction variables mentioned in section 4.4.1. While in study 1, the interaction of the participants with keyboard and mouse were logged in normal working conditions, in the second study programming tasks of about an hour duration were given to participants as task to minimize
boredom and to reduce activity time. The idea was same to replicate the earlier findings with different experimental setup and conditions. Research shows that personality of people mainly remains constant over years and external conditions like time, environment etc. do not effect personality of people (Shaffer, 2009). Therefore, it was hypothesized that interaction of participants with keyboard may produce similar results as that of experiment 1. The material used, preparation of the data, analysis and results are presented in the following sections

5.1. Material

The customized application used in study 1 with disabled mood dialogs were modified to log keyboard and mouse interaction data. Original data was not recorded and instead categorical data were recorded. Like previous study this study also measured participants’ personality by using short form of IPIP-NEO Big Five personality instrument. In this study programming tasks assigned to participants were to simulate a story (see section 5.2) in ALICE programming language. ALICE is a free object oriented and simulation based language that is often utilized to teach programming to children.

5.2. Experimental setup

Main differences from experiment 1 were in the setup of the experiment. A quiet room without an external interference was reserved for the experiment. All of the participants were either programmers or expert computer users. They were trained to program in ALICE. All participants successfully completed four learning tutorials already present in ALICE documentation for learning purposes. Two of the Aesop stories (‘the hare and the tortoise’ and ‘the cow and the frog’) were assigned to participants to simulate. Music was played in background while participants were performing tasks. The experiment was divided into two sessions with a gap of 10-20 minutes in each session. The logging application was running in background to log keyboard and mouse interaction during the entire duration of the experiment. Same variables as discussed in section 4.1 were also being logged in this study.

5.3. Participants

In this experiment, 15 participants were involved. All participants rated themselves as expert computer users but with no experience of programming in ALICE. The participants mean age was
21 years with standard deviation of 3.1. The mean experience of participants of computer usage was 7.3 years and standard deviation of 2.77. There were no female participants in this study.

5.4. Results

5.4.1 Data preparation

Average of eight behavioural measures such as: (1) number of events, (2) average time between events, (3) standard deviation of the time between events, (4) total windows switched, (5) number of backspace and delete keys, (6) number of alphabetic and numeric keys, (7) number of mouse clicks, and (8) number all other keys were calculated over the duration of the experiment. Before calculating average of behavioural variables over the sessions, the data was filtered with filters like: ‘time difference between the two events’ should be greater than 50 milliseconds and less than 20 seconds. The reason of using these filters is already discussed in section 4.4.1. This data was merged with personality data in an excel file forming a single row with first eight columns of interaction data and last columns containing personality data.

5.4.2 Analysis II

Pearson correlation was conducted between the interaction data and personality data. Various traits and sub-traits of personality were again found to be significantly correlated with that of keyboard and mouse interaction behaviour. Only one result appears to be significantly correlated in both studies – the standard deviation of average time between events and activity level with $r=0.59$ and $p<0.05$. The results from this study and the earlier one indicates that behavioural measures have no relation with personality traits except activity level and standard deviation of average time between events. Activity levels fall within the sub-traits of extraversion. It is related to the level of excitement and energy within individuals (Whalen and Gates, 2007). In both studies activity level was found to be positively correlated with that of standard deviation time between events. To confirm this effect, another analysis was planned by utilizing the data from the experiment 1.

6. ACTIVE WINDOWS ANALYSIS-STUDY I

As the first experiment was field study with data collected from participants working at their offices/homes etc. therefore participants interacted with various windows during the experiment. One of the columns ‘Active windows’ (on which the participants were working on) was being
recorded in the log files. Its purpose was to measure the keyboard and mouse interaction with different applications. A skimming through these data revealed that participants interacted with various different applications. There were various applications which do not need a lot of interaction from participants like audio/video players, as well as Installation/Uninstallation windows etc. The remaining applications were categorized as below:

1. Browsers
2. Word Editing
3. Messaging
4. Programming and Development
5. Games

Some of the windows/software were identifiable by their general name like Microsoft word and Internet Explorer. However, some of the windows were not identifiable as they were tagged like “unidentifiable window” or “False Window”. The reason was that these applications unlike various other applications did not contain sub documents but one general name. For example a document named ‘doc1’ will display a title like: ‘Microsoft Word – doc1’. Extracting the name of the application without identifying the name of document becomes easy. However, some applications like games were logged as ‘Unidentifiable Window’ to protect information related with participants whereas applications without sub-documents with direct names were recorded as ‘False Windows’

Keystroke analysis against ‘Unidentifiable windows’ and ‘false windows’ showed mouse clicks and ‘other keys’ events dominated over all other events. As up, down, left, and right arrows were included in ‘other keys’, therefore these windows were most likely to be games as usually games use a lot of mouse clicks as well as direction keys. Similarly, use of alphabetic and numeric keys was minimal compared to the aforementioned windows thus supporting the claim that the window identifies some game being played. Only two of the entries after analysing the keys were not clearly identifiable and therefore were left out of the analysis.

Different participants showed different frequency of use of the above mentioned five categories. If total number of events by a participant for an application over five days were less than 1000 events, then this application was not included in the analysis. Browsers and Gaming applications
were used by all the 20 participants included in this analysis. However, word editing applications with events greater than 1000 were only used by 16 participants. Chatting/Messaging was used by 19 participants. Programming and development related applications were used by only 12 participants. The filters like: time difference of 50 milliseconds between two events and less than 20 seconds were also applied on this data.

### 6.1 Results

Correlation analysis was conducted between ‘standard deviation of average time of events’ on five categories of applications separately with that of ‘activity level’. Only the one sub-trait of extraversion ‘activity level’ was considered as this variable showed significant correlations in earlier two studies. Results revealed that activity level was only significantly correlated with the standard deviation of the average time between events that occurred on the programming applications ($r = 0.646, p = 0.023$). The outcomes support the notion that generally the participants who tend to be programmers have positive correlation with their activity level.

### DISCUSSIONS-PROGRAMMERS’ ACTIVITY LEVEL

In this study, three types of computer users participated: (a) programmers, (b) experts, and (c) novices. These ranking were participants’ own perceptions. On analysis of the participants’ keyboard and mouse interaction data from experiment 1, it appeared that all the users either were programmers or were users who used computer frequently. All of the participants used word browsers and game applications. Majority participants (a total of 19) used chatting/email applications as well as 16 used word editing applications. Similarly, of all the participants were experienced computer users’ with mean experience of 2 years or more. In the second study, all the participants’ were trained of programming task and eventually carried out that programming task. Therefore, it is safe to presume that finding of this study largely dealt with programmers. Results from all of the analysis showed that it is possible to predict programmers’ activity level from their interaction behaviour with keyboard and mouse.

Activity level is a sub-trait of extraversion. According to (Goldberg, 1999) a person with positive activity level is one who can manage multi-tasking, target goals and achieve them without wasting his time. These are all attributes that belong to good programmers. Researches like (Stolee et al., 2011) and (Staurt, 2010) are some of the examples that discussed on the positive characteristics of
the programmers. Therefore, the results of our research can provide a useful application to predict
good programmers vs. not so good programmers.

Programmers’ activity level might also be dependent on the extent of stress they face during
development. Nan and Harter et al. (Nan and Harter, 2009) demonstrated relationship between
schedule pressure and development activities in software development life cycle (SDLC). They
concluded in their paper that excessive pressure affect development performance negatively. An
indirect inference is that increased pressure may cause a programmer to rapidly press keyboard
keys, indicating high activity level on keyboard and mouse. Therefore, one of the findings of this
research is that the pressure to complete a task within a deadline may cause increase activity level
with the keyboard and mouse.

The variables having significant correlations were analysed to find what frequency of keyboard
interaction is related with which level of activity, where activity level is a participants’ personality
trait rating out of 100. Both analyses of experiment 1 revealed that in conditions where there is no
pressure on the programmers and when their activity level is equal to or above 60 then standard
deviation of difference between events time lies above 1600 milliseconds. In the pressure
development as in experiment 2, standard deviation time was reduced to 530 milliseconds for
programmers having activity level above 60. Considering these statistics, an algorithm to predict
programmers’ activity level is formulated as follows:

**ALGORITHM**

A general algorithm to measure a person’s activity level from his/her interaction behaviour with
keyboard and mouse is formulated. As evident from current research [Khan et. el, 2013] that this
algorithm will most likely to work on programming related application like I.D.Es’.

1: **Begin Program**
2: Let \( E \) a current keyboard or mouse interaction event
3: Let \( E^p \) a previous keyboard and mouse interaction event. If this is the first event then \( E^p = 0 \)
4: Record programmers’ keyboard and mouse interaction over or at least for \( n \) minutes in a file \( F \) where \( n \geq 3 \)
5: Set ‘previous event’ to null
6: For time \( t \) loop from \( t = 0.1 \) second to a finite time but greater than \( n \) minutes
Scan ‘current event’ from file F

If (' previous event' is not null) then

Let TD a variables that will store time difference between two events

TD = absolute (E_p - E_c)

Comment: The line below shows filter applied on the interaction data

If (TD > 50 ms) and (TD < 20s)

Add the difference into an array A_i

End If (TD > 50 ms) and (TD < 20s)

End If (there is previous event)

Put ‘current event’ in ‘previous event’

Scan file for next event and put it into ‘current event’

End for

Let SD = Calculate Standard deviation of all the elements in A_i

If (Development Environment is a Pressure Environment) and

If (SD > 530 milliseconds)

Predication “Programmer has high activity level and might be a good programmer”

End if (SD > 530 milliseconds)

Else (Development Environment is not a Pressure Environment)

If (SD > 1600 milliseconds)

Predication “Programmer has low activity level and might not be a good programmer”

End if (SD > 1600 milliseconds)

End if ((Development Environment == Pressure Environment) and

End Program

7. CONCLUSIONS, LIMITATIONS AND FUTURE RESEARCH

Based on the participants as well as results of study 1, it is evident that most of the personality traits and sub-traits correlate with keyboard and mouse interaction behaviour. However, only activity level was found to be positively correlated in both the experiments and three analyses. Moreover, as these participants had more than two years of experience with computers as well as
almost all participants showed significant usage of applications that need expertise, it can also be
safely concluded that these participants were either programmers or possessed programming skills.
The main conclusion of the research is that programmers’ personality specifically activity level
can be predicted from their interaction behaviour with keyboard and mouse events.

Like all empirical researches, this research also has some limitations. For example, sample
participants were mostly graduate and PhD students that limited the findings specific to only such
people. Another limitation was that these results are mostly representative of male population as
there were only two females in study 1 and no female in study 2. However, this limitation denotes
an overall trend of technology and engineering industry where female population is scarce. In
addition, 40% of the participants in both the studies were same. Therefore, there is a chance of
repetition in the observed correlations. However, as sampling was done in two different
environments, the aforementioned limitation might have negligible effect on overall results of this
research.

Various future possibilities exist from this research. This research laid a foundation for an
application to measure users’ personality (activity level) from their interaction behaviour with
keyboard and mouse. Further experiments to test each of five personality dimensions separately
with the keyboard and mouse interaction might enable researchers to detect personality with more
accuracy.

Acknowledgements

I am thankful to Dr. William-Paul Brinkman of Delft University of Technology, The Netherlands
under whose kind supervision, I have collected the data used in this paper.

8. REFERENCES

1-29.
515 Inventory-2 (MMPI-2): Manual for administration and scoring, Minneapolis: University of
516 Minnesota Press.
518 Studies 58: 207-214.
520 Jersey: Lawrence Erlbaum.
522 Cattell RB. (1955) Handbook for the Objective-Analytic Personality Test Batteries: (including Adult and Child
523 OA Batteries): Institute for Personality and Ability Testing.
529 Psychological Testing: An Introduction to
532 Costa PT and MacCrae RR. (1992) Revised NEO Personality Inventory (NEO PI-R) and NEO Five-Factor
536 Digman JM. (1990) Personality structure: Emergence of the five-factor model. Annual review of
537 psychology 41: 417-440.
539 Big Five factors of personality. Psychological assessment 18: 192.
542 Geen RG. (1984) Preferred stimulation levels in introverts and extroverts: Effects on arousal and 
546 Goldberg LR. (1999) A broad-bandwidth, public domain, personality inventory measuring the lower-level 
547 facets of several five-factor models. Personality psychology in Europe 7: 7-28.
549 Heller W. (1993) Neuropsychological mechanisms of individual differences in emotion, personality, and 
551 Khan IA, Brinkman W-P and Hierons R. (2013) Towards estimating computer users’ mood from interaction 
552 behaviour with keyboard and mouse. Frontiers of Computer Science: 1-12.
554 Smart graphics. ACM, 107-115.
559 McCool LC. (2008) The Health Behaviours of Exercise and Dietary Intake: Links with Personality and 
560 Coping, in a student sample.


Phau I, Sequeira M and Dix S. (2009) To buy or not to buy a...


