Summer 10-6-2011

A COMPARISON OF COMPUTING AND NON-COMPUTING STUDENTS' PERSONALITIES BASED ON THE FIVE-FACTOR MODEL

Patricia Alexander

Vreda Pieterse

Hugo Lotriet

A COMPARISON OF COMPUTING AND NON-COMPUTING STUDENTS' PERSONALITIES BASED ON THE FIVE-FACTOR MODEL

Alexander, Patricia M, University of Pretoria, Lynnwood Road, Pretoria, South Africa, patricia.alexander@up.ac.za

Pieterse, Vreda, University of Pretoria, Lynnwood Road, Pretoria, South Africa, vreda.pieterse@up.ac.za

Lotriet, Hugo H, University of South Africa, Corner of Walker and Mears Street, Pretoria, South Africa, lotrihh@unisa.ac.za

Abstract

The research project uses words provided by first year South African university students describing their own personalities. These words are used to compare the personalities of computing and non-computing students on the basis of the Five-Factor Model. The researchers used qualitative methods to code and group the words in terms of the Five-Factor Model, then used statistical methods to analyse and compare the data of computing and non-computing students. The findings of the study suggest that there are significant differences in personalities between students with computing and non-computing majors, and that gender also plays an important role. Most students (both computing majors and non-computing majors) choose to describe themselves as agreeable, conscientious and extraverted. However, significant differences emerged in terms of less commonly chosen personality factors, with computing majors students highlighting their openness to new ideas and their confidence significantly more often than other students do. Computing majors also see themselves as less conscientious than non-computing majors. These seem to be new findings.

Keywords: computing students, non-computing students, Five-Factor Model, gender, South Africa.

1 INTRODUCTION

The importance of personality type in terms of vocational behaviour has been recognized in extensive research, e.g., Ham, Junankar, and Wells, (2009) and Zhao, Seibert and Lumpkin (2009) and has also been related specifically to computing-related careers (Alexander et al., 2010; Rosenbloom et al., 2008). From the perspective of a particular vocational discipline, knowledge of the personality types of students that choose to study that discipline has obvious benefits in terms of understanding cognitive and affective aspects of the student population, marketing, doing career guidance (Larson et al., 2010) and crafting the pedagogy of courses within the discipline, e.g., Lau and Yuen (2009).

In the case of students in computing disciplines, the understanding of issues around career choice has become even more urgent, given the world-wide drop in student numbers in spite of a demand for highly skilled and qualified workers to meet the technology requirements of all organizations.

The shortage of such students in computing-related degrees (computing here is used in a sense that includes Information Systems, Computer Science, Information Science and Computer Engineering) has been as keenly felt in South Africa as elsewhere in the world. As South Africa is one of the emerging economies, the implications of shortages of skilled computing students are significant in terms of providing for the country's needs in terms of not only technology and its management, but also support for socio-economic development that is a high priority on the societal agenda. Limited ICT skills thus form part of the 'concern that permeates Information Systems in Developing Countries literature', (Avgerou, 2008) .

As a result, various research projects have been initiated in South Africa to consider issues related to skills demands, training and vocational behaviour in general of students in computing (Alexander et al., 2010; Alexander & Pieterse, 2010; Blignaut & Naude, 2008; Chetty et al., 2007; Coetzee et al., 2010; Merkofer & Murphy, 2009; Mitrovic, 2010). This paper reports on one aspect of a larger project that involves collaboration between four South African universities on understanding the vocational behaviour of computing students. The project is in its second year, and involves surveys of large numbers of first year students on their career choices and expectations.

This particular paper reports on information that the students provided that could be linked to personality. Specifically the paper focuses on establishing whether there are significant personality differences between students taking computing majors (CM) and non-computing students (NCM) in terms of the Five Factor personality model (Costa & Mc Crae, 1985; Digman & Takemoto-Chock, 1981; Ehrhart & Makransky, 2007). Given the prominence of the Five-Factor Model (FFM) and the world-wide phenomenon of dropping student numbers in computing, we believe that the results presented in this paper are not only of interest to South African audiences, but could guide vocational research in computing-related disciplines elsewhere in the world as well.

The paper has the following sections: (1) A literature survey briefly summarizes the main concepts of the five-factor theory and introduces literature regarding personality and computing; (2) The methodology used for this paper is described; (3) We present findings based on the analysis of our survey data; (4) Some conclusions are presented.

2 LITERATURE REVIEW

2.1 Personality

The Five Factor Model (FFM) of personality is one of the most prominent contemporary personality theories. It was popularised by the work of Paul Costa and Robert McCrae (Robertson & Kinder,

1993). It describes and explains individual differences in how people think, feel, and behave in different situations in terms of five factors. The factors are Neuroticism (N), Extraversion (E), Agreeableness (A), Conscientiousness (C), and Openness to Experience (O).

Description	Facets	Level Descriptors	
Non-Neuroticism		**N+**	**N-**
levels of proneness to psychological distress	Anxious	Relaxed, calm, secure	Worrying, uneasy, nervous
	Angry and Hostile	Composed	Quick to feel anger
	Depressed	Slow to be discouraged	Easily discouraged
	Self-Conscious	Confident	Easily embarrassed
	Impulsive	Resist urge	Easily tempted
	Vulnerable	Handle stress easily	Sensitive, Difficulty coping
Extraversion		**E+**	**E-**
quantity and intensity of energy directed outwards into the social world	Warm	Affectionate, Friendly	Reserved, Formal
	Assertive	Speaks up	Stay in background, shy
	Gregarious	Prefers company, outgoing	Solitary, withdrawn
	Active	Vigorous pace	Leisurely pace
	Excitement seeking	Craves excitement, energetic	Low need for thrills
	Positive	Cheerful, Optimistic	Less exuberant
Openness to Experience		**O+**	**O-**
the active seeking and appreciation of experiences for their own sake	Fantasy	Imaginative, Inventive	Focus on here and now
	Aesthetics	Appreciate art and beauty	Appreciate practical value
	Feelings	Values all emotions	Discount feelings
	Actions	Prefers variety, try new things	Prefers familiarity, cautious
	Ideas	Broad intellectual curiosity	Narrow intellectual focus
	Values	Open to re-examining values	Dogmatic, conservative
Agreeableness		**A+**	**A-**
the kinds of interactions an individual prefers from compassion to tough mindedness	Trusting	See others as honest.	Cynical, Sceptical
	Straightforward	Frank	Guarded, Stretches truth
	Altruistic	Friendly, Willing to help others	Reluctant to get involved
	Compliant	Yields under conflict; defers	Aggressive, Competitive
	Modest	Self-effacing, Humble	Feels superior to others
	Tender-Minded	Empathetic, compassionate	Hard-headed, Rational
Conscientiousness		**C+**	**C-**
degree of organization, persistence, control and motivation in goal directed behaviour	Competent	Efficient, capable and effective	Feels unprepared
	Orderly	Well-organised; neat; tidy	Unorganised; unmethodical
	Dutiful	Reliable	Easy going.
	Achieving	Driven to be successful	Low need for achievement
	Self-Disciplined	Focused on completing tasks	Procrastinates, careless
	Deliberate	Thinks carefully before acting	Spontaneous; Hasty

Table 1: Summary of our understanding of the FFM and the code families created for analysing data

Although the FFM is widely accepted, different scholars give different names to the five factors and interpret the meaning of the individual factors differently. Table 1 summarises our understanding of the FFM and identifies the code families we created for analysing our data. The five factors and their facets given in this table are those defined by Costa and McCrae (1992). In Table 1 we include the textual description for each of the factors provided by Costa and Widiger (2002, cited by Vick, 2005). For further clarification we also include words describing the positive and negative pole of each of the facets mostly as defined by Howard and Howard (1995). These are called level descriptors.

A wide variety of personality theories exist but we consider the FFM appropriate for our research because it is often used as a benchmark and it also enables meta-research. The initial development of the FFM can be credited to Tupes and Christal (1961, cited by Taylor, 2004) and it was brought into personality research by Norman (1963). However, it only became popular after Paul Costa and Robert McCrae proposed their NEO Personality Inventory. This inventory was subsequently revised to measure personality in terms of the currently accepted FFM (Costa & Mc Crae, 1992). Costa and McCrae generated an enormous amount of research, co-authoring with almost all the prominent personality theorists, to determine correlations between the FFM and a variety of established theories that preceded the FFM.

When we analysed our data in terms of the FFM we came across a significant number of student quotations that related to spirituality and values. We turned to the literature to inform a decision as to how we should deal with these quotations and found that spiritual variables are considered to be distinct from the FFM factors. For example, Piedmont (1999) developed a construct that he dubbed Spiritual Transcendence. He established that this new construct was independent of personality in terms of the FFM and proposed that it be considered a sixth major factor of personality. Similarly MacDonald (2000) identified five dimensions of spirituality and demonstrated that four of these seem to be conceptually unique relative to the FFM and argued that the FFM is incomplete, lacking a domain that addresses spirituality. A six-factor model called HEXACO, was proposed by Ashton and Lee (2005). It involved a reorganization of the FFM factors, with the addition of some new variance. The factors are Honesty-Humility (H), Emotionality (E), eXtraversion (X), Agreeableness (A), Conscientiousness (C), and Openness to Experience (O). Lee, Ogunfowora and Ashton (2005) emphasise that the HEXACO model is not simply a "Big Five plus Honesty-Humility" model, as the HEXACO Agreeableness and Emotionality differ in their rotational orientation and content from the two corresponding factors in the FFM structure. Research by Ashton and Lee (2007) claims that the HEXACO model predicts several personality phenomena that are not explained within the FFM.

Code Family	Codes	Number of Quotations
A-	Bully at times, Controlling, Critical, Demanding, Difficult, Egotistical, Greedy, Impatient, Manipulative, Not Temperate, Proud, Sarcastic, Sceptic	97
A+	Accommodating, Advisable, Agreeable, Caring, Charming, Compassionate, Cute, Easy, Flexible, Follower, Forgiving, Friendly, Helpful, Humble, Indulgent Person, Kind, Loving, Loyal, Non-Racist, Obedient, Participating, Patient, Polite, Respectful, Selfless, Sharing, Soft Hearted, Sweet, Straightforward, Sympathetic, Team Player, Trusting, Warm Hearted, Willing	2388
V+	Authentic, Belief, Believer, Christian, Fair, Family, Honest, Moral, Religious, Sincere, Spiritual, Truthful	348

Table 2: Extract showing examples of codes grouped into code families

Despite the views mentioned above, we maintain that spirituality is strictly speaking *not* a personality factor. Nonetheless, we decided to create a code family, namely V+, for investigating the differences in attitude between our student populations with regard to values and spirituality. Table 2 includes the codes that we grouped into the V+ code family. Since there were no quotations or codes that sensibly fell under V-, this was not added. An explanation of the coding process follows later in the Methodology section.

2.2 Related Studies

Studies reported in the literature regarding distinctive personality characteristics of people pursuing careers in computing or scientific fields have been conducted based on a variety of different personality theories. The Myers Briggs Type Indicator (MBTI) has been used in many studies over the years (e.g., Lyons (1985), Turley and Bieman (1995), Capretz (2003), Galpin et al (2007) and Choi et al (2008)). A number of studies was conducted based on Holland's RIASEC model (e.g., Porter and Umbach (2006), Rosenbloom et al. (2008) and Lounsbury et al (2008)). Other models have also been used, such as, a study by Wynekoop and Walz (2000) involved Gough and Heilbrun's Adjective Checklist (ACL),one by Ferratt and Short (1988) which investigated differences at the hand of motivators within need areas, as well as one by Blignault and Naude (2008) that considered Boyd's DISC model.

Studies that investigate the personality of professionals in computing careers and involving the FFM mainly focus on how personality attributes correlate with success in specific computing activities, e.g. with programming performance (Darcy & Ma, 2005), with tasks specific to stages in the software development process (Sodiya et al., 2007), and with teamwork in Software Engineering (Acuña et al., 2009).

The contribution made by this paper is to compare personality differences of people who have opted to do tertiary computing qualifications with those of people who have opted to do tertiary non-computing qualifications. As far as we know this has not been done in terms of the FFM before.

3 RESEARCH METHODOLOGY

This research used qualitative data collected from one question included in questionnaires sent out early in the South African 2010 academic year. The respondents were 1985 first year students registered for introductory computing courses at four diverse South African universities, which together are believed to cover the full spectrum of types of university in the country. University A is an extremely large university that offers courses by means of distance teaching only. This university has students from all segments of the population, including some from outside South Africa. A high percentage of their students work full time and their average age is high in comparison with the average ages of the other universities. This university offers a wide range of programmes from degrees to shorter diploma and certificate courses. University B is a University of Technology in a major city which offers degree and diploma courses. The students at this university are heterogeneous in terms of demographic description. University C has had students almost exclusively from disadvantaged backgrounds throughout its 100 years of existence and is in a smaller centre in a low-income area. University D in contrast is considered a prestigious research university, in a major city and which has historically had advantaged students. Only in the last twenty years has this university had a more inclusive registration policy.

In the sample there were 633 male CM students 283 female CM students, 546 male NCM students and 523 female NCM students. The lower number of female CM students than male is evidence of the global problem that females are under-represented in computing classes at university level. The question that forms the basis for the analysis in this paper asked the respondents to "Give any five

words to describe your personality". Most students gave single words but occasionally they used a phrase.

In terms of the typology of Leech and Onwuegbuzie (2009) the study employed a 'partially mixed sequential dominant' design. The method is dominantly quantitative and employs statistical methods to analyse the data during the later phases., .However, the data that were gathered are qualitative. The respondents were allowed the freedom of choosing any words to describe their personalities. The transition from qualitative data to quantitative analysis was achieved through a pragmatic combination of qualitative and quantitative methods, with an initial qualitative phase involving the categorisation of these words according to accepted qualitative principles with the categories associated with the FFM. The systematic approach followed to achieve this is briefly described below.

The strength of the approach followed (Johnson and Onwuegbuzie, 2004) was that the data gathered closely represented the respondents' own personal opinion of themselves without imposing a theory or a framework. The potential weakness of the results being influenced by the researchers' own biases was addressed through a systematic research protocol that involved (1) coding of data by more than one person (as described in following sections); and (2) categorisation of student quotes by more than one person. The high level of inter-coder reliability that was obtained in terms of our code families suggests that the final set of categories used for statistical analysis is largely insensitive to personal bias. For the more dominant phase of the research (i.e. the statistical data analysis) de facto statistical procedures were followed.

During the qualitative phase we used Atlas.ti vs. 6 as a tool to assist us in analyzing this data. The primary document consisted of a separate line of text for each student, consisting of an identifier, demographic data and the five personality descriptors (single words or short phrases). Two independent coders, one of whom is an author of this paper, and one a senior student, then used Atlas.ti to assign codes to quotations. Each quotation was assigned only one code per coder. The second coder was provided with 214 codes created by coder 1, but was free to add to them and ended with a set of 356 codes. After eliminating duplicate and unused codes, coder 1 had 209 codes and coder 2 had 257 codes of which 166 were common to both coders.

The output from this phase of the process was two large sets of codes and the links to quotations in the two separate primary text documents, one of which had been spelling checked using a word processor before coding was done and the other not. Two working documents (summaries) were then created giving a code, number of quotations linked to that code and synonyms, corresponding to actual data or quotations (an extract is shown in Table 2).

Three researchers, one of whom was a coder in the first phase, were asked to use these summary documents and classify the codes that resulted from the first phase into the ten main categories (code families) related to the FFM discussed in section 2.1. Each code was to be assigned to a single code family. This was done independently by each researcher and repeated for the two summaries. Two code families were used for each of the five factors of the FFM, representing the high and low levels of each factor. This resulted in ten code families, namely, E+, E-, O+, O-, A+, A-, C+, C-, N+, N-. Two further code families were added – one was for ethical or value-based words (e.g. 'truthful' as discussed in Section 2.1) and one (N/A – not applicable) was for codes did not relate to any one of the other code families (as might be the case where a student described an attribute that was not in any way related to personality, e.g. 'married' or 'tall'). The researchers each applied their own interpretation as to what is meant by the high and low levels of the different code families and, as a result, they did not always agree on the code families to which the codes should be assigned.

Once the three researchers had completed this phase of the exercise, they compared their results and reached consensus regarding assigning each code into a single code family.

The percentage of original student quotations falling into each code family was then compared using the two original coders' coding to determine inter-coder reliability. It was not considered important whether the two coders agreed at the lower level of assigning quotations to codes, as these codes

would not be used in further analysis. Checking for significant differences was only done at the level of code families. According to the calculation on the number of quotations per code family, the correlation of the classifications between the two sets of codes was more than 0.998. This correlation suggests that the code families are insensitive to any disagreements in classification of words into the different families by the researchers.

Once the reliability of the code family classification was confirmed, Atlas.ti was used to instantiate the code families and links to the quotations. This concluded the qualitative and interpretative process, where data analysis was primarily data classification, and the quantitative part of the analysis could commence using queries to determine the distribution of code families according to various demographic variables.

4 FINDINGS

4.1 Coding

Table 2 shows the codes that were grouped into the A+, A- and V+ code families. It serves as an example of how codes were classified into our code families. Since students had absolutely free choice in creating descriptors it makes sense that the number of initial descriptors is unpredictable and the number of codes reflects this to some extent. Note that the number of quotations in each category refers to individual quotations and does not match the percentages in the tables that follow. The percentages in the tables that follow reflect the percentage of respondents who included one or more descriptors for each code family.

4.2 Computing Major compared with Not Computing Major

Code Family	CM	NCM	p
A+	72.4%	71.3%	.608
C+	62.8%	63.5%	.759
E+	63.6%	63.4%	.930
N+	37.2%	32.9%	.049*
O+	33.0%	26.6%	.002*
V+	15.9%	17.0%	.508

Table 3: High level code families - Computing Major compared with Not Computing Major (indicates $p < 0.05$)*

Code Family	CM	NCM	p
A-	4.9%	4.7%	.837
C-	2.9%	1.0%	.002*
E-	20.1%	16.9%	.070
N-	6.5%	7.3%	.498
O-	7.0%	7.8%	.549
V-	0%	0%	

Table 4: Low level code families - Computing Major compared with Not Computing Major (indicated $p < 0.01$)*

Table 3 shows the percentages of students from the two groups, Computing Major (CM) and Not Computing Major (NCM), that described themselves using at least one descriptor from the given code family. Only the high level (positive) code families are considered in this table.

Table 3 shows that both sets of students used descriptors from the A+, C+ and E+ code families nearly twice as often as descriptors from N+ or O+ and four or more times as often as the V+ code family. Hence, most students described themselves as Agreeable, Conscientious and Extraverted. This observation is in line with that of Darcy and Ma (2005) who also observed higher than usual levels of Conscientiousness, Agreeableness and Extraversion in their programming participants.

Table 3 shows that for two of the less commonly chosen code families, namely, N+ and O+, there are significant differences at the $p < 0.05$ level between the CM and the other students. In both cases CM students more often selected descriptors belonging to these code families than the NCM students did, whereas the groups were remarkably similar with respect to the other code families. Hence CM students elect to highlight their openness to new ideas and their confidence significantly more often than NCM students do. Bearing in mind that the sensing-intuition dimension of the MBTI correlates strongly with low and high levels of openness to new ideas (Furnham et al., 2003), this is in contrast with Choi et al (2008) who reported far more sensing and fewer intuitive types among their computing participants.

As could be predicted, considerably larger numbers of students described themselves in terms of descriptors that fell into code families representing high levels (+) of the five factors than those whose descriptors fell into the corresponding code families representing low levels (-). Therefore we regard Table 3 as more reliable and revealing than Table 4.

As in the case for Table 3, Table 4 shows the percentages of students from the two groups that described themselves using at least one descriptor from the given code family. Personality descriptors relating to E- (Introverted) were more commonly chosen than code families representing any of the other low levels of factors. In the light of the stereotypical computer user, and the comment by Lounsbury et al (2008) that two-thirds of computer professionals are estimated as being introverted, it is not surprising that some students are willing to describe themselves as such. However, if we compare the E- percentages (Table 4) with E+ percentages (Table 3) we see that whereas 20.1% of the CM students used at least one descriptor from E-, 63.6% used a descriptor from E+[1]. The corresponding figures for NCM students are: E- (16.9%) and E+ (63.4%). Therefore, we can see that all students are more likely to select words from the high level of the E factor than its low level when given complete freedom to use any words to describe themselves. Furthermore, although the stereotypical CM introverted students featured in our data, extraverted CM are far more prominent and not significantly different from NCM. This observation also corroborates the finding of Lounsbury et al (2008) who found no significant differences in terms of extraversion between their sample of computer professionals and other occupations as a whole.

Table 4 shows that there is a significant difference between CM and NCM students regarding the C- code family only. Although an admittedly small number of students chose C- descriptors, CM students chose them nearly three times more often than NCM students. CM students see themselves as less conscientious than NCM students. This seems to corroborate the findings from Table 3 and seems to indicate a greater preference in the CM group for experimenting as a learning style rather than the more disciplined or structured approach which is more highly favoured by the NCM group. This observation seems to be unique in the sense that it has not been prominently reported in the literature before. However, it is important to note that it stems from a fairly small minority.

4.3 Comparing Gender within Major

A strange phenomenon is evident in Table 5: for code families A+, C+, N+, O+ and V+, only one of the student groups (either CM or NCM) showed significant differences between males and females. Only E+ had no significant difference in terms of gender for either group. In most cases (with C+ the exception) there was a very marked difference between the p values for the CM and NCM groups.

Taking those code families where the NCM males were significantly different from the NCM females, A+ (Agreeableness) was more often selected by the female students than the males but was the most

[1] Since students each gave five different descriptors, one individual could choose descriptors one or more of which fall into the high level (E+) and others into the low level (E-) category.

popular code family for all four groups. N+ (Not Neurotic) and O+ (Open to new experiences) were more often selected by the NCM males than NCM females.

	CM			NCM		
Gender	M	F	p	M	F	p
Code Family						
A+	70.9%	75.3%	.168	66.2%	76.8%	.000*
C+	59.5%	69.6%	.004*	61.0%	66.5%	.071
E+	63.2%	64.2%	.774	60.8%	66.3%	.074
N+	37.6%	36.7%	.794	38.0%	27.5%	.000*
O+	34.4%	29.8%	.164	30.0%	23.2%	.014*
V+	17.9%	11.8%	.019*	16.5%	17.4%	.700

Table 5: High level code families– Comparing Gender within Major (indicates p< 0.05)*

	CM			NCM		
Gender	M	F	P	M	F	p
Code Family						
A-	5.1%	4.2%	.522	4.9%	4.5%	.750
C-	4.0%	.7%	.006*	1.7%	.2%	.015*
E-	21.7%	17.0%	.099	17.3%	16.4%	.710
N-	7.1%	5.2%	.274	8.0%	6.6%	.387
O-	6.5%	8.3%	.313	8.7%	6.6%	.194

Table 6: Low level code families– Comparing Gender within Major (indicates p< 0.05)*

The comparisons of males and females for these three code families match commonly held gender stereotypes. Interestingly, CM students appear less prone to these gender stereotypes than the NCM students are (there is no significant difference between CM males and CM females for these code families).

The two cases where CM males differed significantly from CM females were C+ and V+. These also showed opposite 'directions', that is, CM males **more frequently** chose V+ descriptors than their female classmates but the CM males **less frequently** chose C+ than their female classmates. V+ was chosen by the minority of all students (between 11.8% and 18%) with female CM students choosing to describe themselves this way considerably less often than any other group (by gender and major) and the CM males choosing this most often. The C+ group was a popular set of descriptors (60 – 70% of students used it) and here CM females took top spot (chose it most often) and CM males had the lowest percentage, whereas C+ was the second most popular code family for the three other subgroups. CM males chose E+ more often than C+.

How should this be interpreted? The use of a separate code family, V+, shows one difference quite clearly whereas it would probably be lost if it was merged with the FFM factors.. CM females portray themselves as the most conscientious of all subgroups while CM males portray themselves as the most virtuous of all subgroups. This might show that to some extent CM men are more ideas oriented and the women rather more down to earth. This might be seen to support the findings in Table 4.

In the NCM group, however, there are no corresponding significant differences. Particularly in the case of V+, NCM students show no gender differences.

Table 6 shows only one significant difference, namely, for C-, on the low level code families. But, in direct contrast with the results for the high level code families, the gender groups of both CM and NCM groups manifested significantly different results for this code family. These C- categories have the lowest percentages of students across all the code families representing both high and low levels of the personality factors, but the female students were much less likely to select a descriptor in this category than the males were, regardless of major. This corroborates the findings for Table 4. Females in both the CM and NCM cohorts of students believe that conscientiousness is very important.

5 CONCLUSION

5.1 Study limitations

This study is based on data derived from an open question asking the students' to describe their own personalities in five words. Answers are, therefore, short and self-reported, and can be biased in response to context. As the questionnaire as a whole focussed on career choice in terms of choice of majors, the responses may be those that respondents thought, even subconsciously, were most suitable for the chosen major and hence may reflect the student's own stereotypes for that discipline. These responses cannot be assumed to be 'true' reflections of the essential and unchanging personality of each respondent (if indeed such a set of personality traits exist). In fact no data collected from human respondents is likely to be 'untainted' by the subject's awareness of the research process, the researcher's expectations or larger social influences. In addition, this data was collected from students who had recently started their studies. The data cannot predict whether they would be successful or enjoy the courses and careers which they are starting.

5.2 Observations

Our findings show that CM and NCM students are similar in terms of descriptors from A+, C+ and E+, i.e. most students choose to describe themselves as agreeable, conscientious and extraverted. However, significant differences emerged in terms of less commonly chosen code families, namely, N+ and O+, with CM students electing to highlight their openness to new ideas (O+) and their confidence (N+) significantly more often than NCM students do.

Within the relatively small subset of the data obtained from students who used descriptors of low levels of the factors, a significant difference in the frequency of use of C- descriptors was observed. Apparently CM students see themselves as less conscientious than NCM students do. This seems to be a novel finding. Unfortunately, despite a corresponding observable difference when considering the high level (C+) descriptors, this difference is not statistically significant and further research is needed to confirm its existence.

When comparing gender differences within the CM and NCM groups an interesting anomaly was observed. Although significant differences regarding commonly held gender stereotypes in terms of the A+, N+ and O+ categories, were visible within the NCM group, there is no significant difference between CM males and CM females for these categories. This may be an indication that the females choosing computer-related careers are less likely to be the stereotypical female or are more sensitive than NCM female students about portraying themselves as fitting these stereotypes.

We included an additional category that may be interpreted as a personality factor beyond the FFM in our analysis. We labelled this V+ and used it for ethical or value-based words. This is not an entirely new approach since other researchers, as mentioned in Section 2.1, have proposed similar amendments to the FFM. There was no need for a V- code family since there were no students who described themselves in terms of low levels of ethics or values. The decision to include V+ as a code family proved to be useful because it enabled us to show an interesting difference between female CM and male CM students. We observed that female CM students portray themselves as being more

conscientious than all other subgroups while CM men portray themselves as being more virtuous than others. This might show that, to some extent, CM men are more ideas and ideals oriented and the CM women rather more down to earth.

5.3 **Recommendations**

The combined findings derived from our observations regarding gender differences within the CM group could potentially be used to address the low number of female students enrolling in computer courses. On the one hand we could look at the reported personality traits of those females who have chosen to major in a computing discipline and emphasize a systematic and conscientious approach rather than the ability to be innovative or discover solutions in order to attract more women who are similar to these. However, we acknowledge that it might make sense to look at the personal requirements of those who have not selected a computer major rather than at those who have already been recruited in order to increase enrolments.

We believe that the use of the FFM provides insights into differences between self-perception of personality between CM and NCM students that are useful in terms of teaching CM courses at universities as these give insight into learning styles. However, the pedagogical design of programmes cannot be based solely on how people see themselves or chose to portray themselves and also needs to find out what they actually respond to (which may be a better reflection as to their real natures). This goes beyond the intention of this research. Nor should the results from this research be used to define a narrow, 'typical' CM student.

While we prefer using the FFM as opposed to the HEXACO model, we recommend that analysis based on the FFM should include the use of a value-based code family. We also propose that future research should be carried out at different stages in a student cohort's time at university to see whether the personality differences become more or less pronounced among more successful and committed groups of students (those who remain in the programme). Although our study was limited to respondents from South Africa universities, the world-wide decline in computing students and the universal interest in and the cross cultural validity of the FFM imply that the study is of interest to a wider audience.

References

Acuña, S., Gómez, M. & Juristo, N. (2009). How do personality, team processes and task characteristics relate to job satisfaction and software quality? Information and Software Technology, 51(3), 627 - 639.

Alexander, P., Holmner, M., Lotriet, H., Matthee, M., Pieterse, V., Naidoo, S., et al. (2010). Factors affecting career choice: Comparison between students from computer and other disciplines. Journal of Science Education and Technology, OnlineFirst 16 Oct, 2010.

Alexander, P. & Pieterse, V. (2010). Indications of personality trait difference between ICT and other students. Paper presented at the 18th European Conference of Information Systems (ECIS' 2010), 7-9 June 2010, Pretoria, South Africa.

Ashton, M. & Lee, K. (2005). Honesty-humility, the big five, and the Five-Factor Model. Journal of Personality, 73(5), 1321-1354.

Ashton, M. & Lee, K. (2007). Empirical, theoretical, and practical advantages of the HEXACO model of personality structure. Personality and Social Psychology Review, 11(2), 150-166.

Avgerou, C. (2008). Information systems in developing countries: a critical research review. Journal of Information Technology, 23, 133 - 146.

Blignaut, P. & Naude, A. (2008). The influence of temperament style on a student's choice of and performance in a computer programming course. Computers in Human Behavior, 24, 1010-1020.

Capretz, L. (2003). Personality types in Software Engineering. Int J Human-Computer Studies, 58, 207-214.

Chetty, M., Buckhalter, C., Best, M., Grinter, R. E. & Guzdial, M. (2007). Description of Computer Science higher education in sub-Saharan Africa: Initial explorations. [Electronic Version]. Retrieved 2010/12/01 from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.125.8376&rep=rep1&type=pdf.

Choi, K., Deek, F. & Im, I. (2008). Exploring the underlying aspects of pair programming: The impact of personality. Information and Software Technology 50(11), 1114–1126.

Coetzee, L., Seetharam, D., Butgereit, L. & Van Gelder, A. (2010). The SAFIPA-MERAKA Institute code-sprints program; A mechanism to enhance the development capacity of emerging developers - observations and lessons learnt. Paper presented at the IST-Africa 2010.

Costa, P. T. & Mc Crae, R. R. (1985). The NEO Personality Inventory Manual. Odessa FL: : Psychological Assessment Resources.

Costa, P. T. & Mc Crae, R. R. (1992). Revised NEO Personality Inventory (NEO PI-R) and NEO Five-Factor Inventory (NEO-FFI): Professional Manual. Odessa: FL: Psychological Assessment Resources.

Darcy, D. & Ma, M. (2005). Exploring individual characteristics and programming performance: Implications for programmer selection. Paper presented at the 38th HICSS- 2005.

Digman, J. M. & Takemoto-Chock, N. K. (1981). Factors in the natural language of personality: Re-analysis and comparison of six major studies. Multivariate Behavioral Research 16 149-170.

Ehrhart, K. & Makransky, G. (2007). Testing vocational interests and personality as predictors of person-vocation and person-job fit. Journal of Career Assessment 15(2), 206–226.

Ferratt, T. W. & Short, L. E. (1988). Are information systems people different? An investigation of how they are and should be managed. Management Information Systems Quarterly, 12(3), 427-443.

Furnham, A., Moutafi, J. & Crump, J. (2003). The relationship between the Revised NEO-Personality Inventory and the Myers-Biggs Type Indicator. Social Behavior and Personality 31(6), 577-584.

Galpin, V., Sanders, I. & Chen, P. (2007, June 23–27 2007). Learning styles and personality types of Computer Science students at a South African university. Paper presented at the ITiCSE '07: The 12th annual SIGCSE conference on Innovation and Technology in Computer Science Education Dundee, Scotland, United Kingdom.

Ham, R., Junankar, P. N. & Wells, R. (2009). Occupational Choice: Personality Matters [Electronic Version]. IZA Discussion paper series, IZA DP No. 4105. Retrieved 30 November 2010 from ftp://repec.iza.org/RePEc/Discussionpaper/dp4105.pdf.

Howard, P. J. & Howard, J. M. (1995). The Big Five quickstart: An introduction to the Five-Factor Model of Personality for human resource professionals. Charlotte, NC: Centre for Applied Cognitive Studies.

Johnson, R.B. & Onwuegbuzie, A.J. (2004). Mixed-methods research: A research paradigm whose time has come. Educational researcher, 33(7), 14 - 26.

Larson, L. M., Wu, T. F., Bailey, D. C., Gasser, C. E., Bonitz, V. S. & Borgen, F. H. (2010). The role of personality in the selection of a major: With and without vocational self-efficacy and interests. Journal of Vocational Behaviour, 76(2).

Lau, W. W. F. & Yuen, A. H. K. (2009). Exploring the effects of gender and learning styles on computer programming: Implications for programming pedagogy. British Journal of Educational Technology, 40(4), 696 - 712.

Lee, K., Ogunfowora, B. & Ashton, M. C. (2005). Personality traits beyond the Big Five: Are they within the HEXACO space? Journal of Personality, 73(5), 1437-1463.

Leech, N.L. & Onwuegbuzie, A.J. (2009). A typology of mixed method research designs. Qual Quant, 43, 265 - 275.

Lounsbury, J. W., Studham, R. S., Steel, R. P., Gibson, L. W. & Drost, A. W. (2008). Personality traits and career satisfaction of Information Technology professionals [Electronic Version]. Retrieved 2009/08/15 from http://info.ecareerfit.com/eCareerFit/IT%20Personality%20Traits%20%20Career%20Satisfaction.pdf

Lyons, M. L. (1985, August 15). The DP Psyche. *Datamation,* 103-110.

MacDonald, D. A. (2000). Spirituality: Description, measurement, and relation to the Five Factor Model of Personality. Journal of Personality, 68(1), 153-197.

Merkofer, P. & Murphy, A. (2009). The e-skills landscape in South Africa: The issues of demand and supply and the use of international benchmarks to inform the South African e-skills development context. Zeitschrift für Politikberatung, 4, 685 - 695.

Mitrovic, Z. (2010). Positioning e-skills within an organisation: An information systems management viewpoint. South African Journal of Information Management, 12(1).

Norman, W. T. (1963). Toward an adequate taxonomy of personality attributes. Journal of Abnormal and Social Psychology, 66 574-583.

Piedmont, R. L. (1999). Does Spirituality Represent the Sixth Factor of Personality? Spiritual Transcendence and the Five-Factor Model. Journal of Personality, 67(6), 985-1013.

Porter, S. R. & Umbach, P. D. (2006). College major choice: an analysis of person-environment fit. Research in Higher Education, 47(4), 429-449.

Robertson, I. T. & Kinder, A. (1993). Personality and job competences: The criterion-related validity of some personality variables. Journal of Occupational and Organizational Psychology, 66(3), 225-244.

Rosenbloom, J. L., Ash, R. A., Dupont, B. & Coder, L. (2008). Why are there so few women in information technology? Assessing the role of personality in career choices. Journal of Economic Psychology 29(4), 543 - 554.

Sarloglou, V. (2010). Religiousness as a Cultural Adaptation of Basic Traits: A Five-Factor Model Perspective. Personality and Social Psychology Review, 14(1), 108-125.

Sarloglou, V. & Muñoz-García. (2008). Individual differences in religion and spirituality: An issue of personality traits and/or values. Journal for the Scientific Study of Religion, 47(1), 83-101.

Sodiya, A. S., Longe, H. O. D., Onashoga, S. A., Awodele, O. & Omotosho, L. O. (2007). An improved assessment of personality traits in software engineering. Interdisciplinary Journal of Information Knowledge and Management 2, 163 - 178.

Taylor, N. (2004). The construction of a South African five-factor personality questionnaire. Unpublished Masters Dissertation, Rand Afrikaans University, Johannesburg.

Turley, R. T. & Bieman, J. M. (1995). Competencies of exceptional and nonexceptional Software Engineers. The Journal of Systems and Software 28(1), 19-38.

Vick, E. H. (2005). Implementing lexical and creative intentionality in synthetic personality. Unpublished PhD Dissertation, University of Central Florida, Orlando, Florida.

Wynekoop, J. L. & Walz, D. B. (2000). Investigating traits of top performing software developers. Information Technology & People 13(3), 13 13:186-195.

Zhao, H., Seibert, S. E. & Lumpkin, G. T. (2009). The relationship of personality to entrepreneurial intentions and performance: A meta-analytic view. Journal of Management, 36(2), 381 - 404.