Abstract

With recent advances of Artificial Intelligence (AI) and its application in medical education, understanding how AI can best support learners in achieving educational outcomes becomes essential. In the context of tutoring, when AI augments academic faculty, mimicking faculty personality aspects may positively impact learners’ achievement. The aim of this study was to identify and examine what faculty traits to implement in AI agents. A validated instrument, based on the Big Five Inventory (BFI), was used to support participants (n=72) recalling personality factors of familiar, exceptional educators they’ve encountered. Results from median dimension score analysis displayed a preference for four broad trait dimensions: Conscientiousness (88%), Agreeableness (86%), Extraversion (75%) and Openness (74%). A strong positive preference for Agreeableness and Conscientiousness and negative preference for Neuroticism related to the Digmanns’ superfactor Alpha which is related to social development. This study supports the need for an affect element in AI driven pedagogical agents. Furthermore, it clarifies the specific preferences of Singapore based medical students. These findings will contribute to the further development of human-like AI research.

Keywords: Higher Education, Innovation, Artificial Intelligence, Medical Education, Flipped Classroom, Virtual Patient, Virtual Tutor, Personality.

1 INTRODUCTION

Medical education is regarded as a challenge given the ever-increasing volume of knowledge in medical education and the struggle educators face to deliver relevant knowledge to undergraduates in a relevant manner and context. The evolution of medical education has seen multiple iterations with respect to both how content is delivered and how best to prepare the physicians of tomorrow for the world they will inherit.

1.1 Recent Advances In Medical Education

The age-old method of classroom lecture, which is ideal for a passive learning, is still commonly used in undergraduate medical studies [1]. However, despite adherence to the tried and tested strategy of classroom lecture models, this instructor-centered model of the past is slowly giving way to a more learner-centered model. In this evolving medical education model, the students are tasked to assume more responsibility for enhancing their own learning [2]. Learners must now review materials prior to the face-to-face classroom interactions. In turn, this increases the responsibility of medical educators to provide richer media experiences in the form of e-textbooks, narrated and interactive presentation slides and 24/7 access to e-learning platforms. The value of these face-to-face interactions is then augmented as educators can now engage the learner in more active teaching and learning [3]. With recent advances in technology, medical education is now exploring how to implement Artificial Intelligence (AI) within e-learning platforms. E-learning platforms are defined in a plethora of ways, however e-learning, as defined in a study conducted by Ruiz et al. [2], refers to a multitude of Internet technology-based learning platforms that also encompass terms like web-based learning, online learning, distributed learning, computer-assisted instruction and/or Internet-based learning. At present, AI is commonly used to augment the delivery of learning resources to users by improving the recommendation system through the utilization of different arrays of data such as “dwell time”, “type of media” and “number of resources accessed”.

1.2 E-Learning platforms and the need to personalize Virtual Tutors

Combining the advent of new mobile technologies (i.e 3G, 4G, smartphones and tablets) and vast distribution of mobile applications (apps) through app stores, there has been a substantive increase in the implementation and use of mobile apps in daily lives, including in medical education) [4].
medical apps initially served as a mobile resource depository, they have since evolved to providing services that can aid in the development of clinical reasoning skills through the use of virtual entities (i.e., virtual patients and virtual tutors). While the role of virtual patients primarily involves conveying case symptoms of a patient to a student doctor, it is starkly different from the more challenging role of a virtual tutor found in medical education apps. Here, the purpose of the virtual tutor is to augment the role of a real-life tutor. In the case of medical education apps, it is not uncommon that virtual tutors take the form of chatbots which provide responses appropriate to the learner’s input. However, this embodiment of the tutor usually expresses as a strict and rigid chatbot, based on hand-coded rules and statistical models built over heuristic parameters [5]. These chatbot responses are robotic, repetitive and lack the flair and contextual vocabulary found in real-life tutors [6]. Research has shown that in teaching clinical reasoning, the presence of tutors plays a more important role beyond simply providing knowledge. Both the quality of feedback and the nature of the delivery has a heavy impact on how much a student will learn during clinical reasoning sessions [7]. Furthermore, success in the implementation of virtual patients is significantly attributed to the presence of real-life tutors. This significant attribution is based on how they adapt to learners’ aptitude, expectation and personality ([6], [8]). Therefore, a common strategy in current virtual tutoring apps is to craft scripts that are personalized to each student, and which simulate the real-life scenarios traditionally experienced during clinical reasoning tutorials. However, due to the sheer amount of effort and time required to generate responses for each given learner input, these scripts often sound repetitive and predictable, especially after extended use of the app. Thus, one of the main challenges is to provide a feedback system that is not repetitive and reflects, to some degree, the personality of a real-life tutor.

1.3 Aspects of personalization in Artificial Intelligence

Research suggests that personality traits of teaching figures have significant impact on teaching effectiveness [9]. Therefore, if AI-driven learning apps are to help students learn effectively, it follows that whenever and wherever possible ways of means of crafting elements of personality to the AI agent should be explored. In so doing, an AI that begins to successfully emulate positive personality traits may result in increased engagement. This occurs because the AI is now able to reflect and reinforce those elements that are characteristically associated with positive teacher-student interactions as experienced with tutors who adapt to the learners in the face-to-face interactions.

The ideal pedagogical agent would have a personality of its own. It would interact with the learner in ways that are compatible with the learner’s aptitude, expectation and personality by providing personalized and non-repetitive feedback.

In this journey to achieve and model an ideal, effective pedagogical agent, this study utilizes the Big Five Inventory (BFI) to examine and identify what key personality traits Asian medical students perceive to be inherent in good tutors. Based on these identified key personality traits, the responses of an AI-driven pedagogical agent can be altered and provide the means for increased engagement throughout the student interactions with the application.

2 METHODOLOGY

2.1 Participants

Year 1 students from a medical academic environment were approached to report on their perceived key personality traits of exceptional educators. In total, data from 90 individuals were collected. Gender and age of the participants were not collected as data collected were anonymous. Participants were all of minimum diploma or Singapore-Cambridge General Certificate of Education Advanced Level (GCE-A level) to ensure that participants had adequate experience of education before entering the study.

2.2 Materials

The Big Five Inventory (BFI) is a model on personality that claims that important personality traits in people can be embodied into their language [10]. Statistical study of personality items using factor analysis identified 5 major group for personality: Openness, Conscientiousness, Extraversion, Agreeableness and Neuroticism. The inventory consists of 44 items which participants had to rate on a Likert scale of 1 to 5 where 1= “Strongly disagree” and 5= “Strongly agree” [11].
2.3 Procedure
Participants completed the BFI prior to the start of their lesson. The instrument was administered by a research assistant at an auditorium, the location of the students’ lesson were conducted. The students accessed the instrument through a QR code that the research assistant presented on screen and completed the instrument online [12]. Students were allocated no more than 15 minutes to complete the inventory. However, the actual BFI does not have a time limit and this time limit was imposed due to administrative arrangements with the faculty.

2.4 Scoring
Based on the collated data, each response was scored across all dimension. The contributing questions to each dimension are as follows, where the “R” in “6R” would mean a reverse scored value of question 6 [11]:

- Extraversion: 1, 6R, 11, 16, 21R, 26, 31R, 36
- Agreeableness: 2R, 7, 12R, 17, 22, 27R, 32, 37R, 42
- Conscientiousness: 3, 8R, 13, 18R, 23R, 28, 33, 38, 43R
- Neuroticism: 4, 9R, 14, 19, 24R, 29, 34R, 39
- Openness: 5, 10, 115, 20, 25, 30, 35R, 40, 41R, 44

The Median Dimension Scores were then calculated by measuring dimension scores against their total possible scores and multiplied by 100 to give a percentage. This value in percentage shows how close the results are to the full score. Refer to Example 1 and 2 for detail on how dimension.

Example 1: For Extraversion, it consists of questions 1, 6, 11,16, 21, 26, 31, 36. Since there are 8 questions and for each question the maximum score a student can give is 5, the maximum score for that dimension is 5*8= 40. Student A scores 25 on extraversion. Since extra-version, has maximum score of 40. The dimension score for Student A on extraversion would be 62.5 [25/40*100=62.5].

2.5 Data Screening
Data was screened for any univariate outliers and incomplete responses. Out of the 90 data responses collected, 18 had to be omitted due to incomplete responses. The final sample size was 72 which provided over 13 cases per variable.

3 RESULTS
3.1 Identifying the key personality traits
Initial descriptive analysis of Highest Rated Dimensions and Median Dimensions Scores of the BFI revealed a strong preference for certain dimensions, Conscientiousness and Agreeableness, over others (see Fig. 1 and Fig. 2).
3.2 Examination of relationship of within and outside of dimensions in BFI

Data from 44-item BFI were first subjected to correlational analysis using SPSS. Results from the correlation matrix suggests that 8 out of 10 correlations were statistically significant however only 6 out of 8 statistically significant correlations were equal or greater to $r(70)= +0.30$, $p<0.05$, two tailed. Results also show that 3 out of 4 correlations with Neuroticism had significant negative correlation and the correlation between Neuroticism and Conscientiousness dimensions statistically strong(negative) at $r(70)= -0.720$, $p<0.00$.

Table 1. Total Variance Explained

<table>
<thead>
<tr>
<th>Component</th>
<th>Total</th>
<th>% of Var</th>
<th>Cumulative %</th>
<th>Total</th>
<th>% of Var</th>
<th>Cumulative %</th>
<th>Total</th>
<th>% of Var</th>
<th>Cumulative %</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.489</td>
<td>49.776</td>
<td>49.776</td>
<td>2.489</td>
<td>49.776</td>
<td>49.776</td>
<td>2.047</td>
<td>40.940</td>
<td>40.940</td>
</tr>
<tr>
<td>2</td>
<td>1.008</td>
<td>20.169</td>
<td>69.945</td>
<td>1.008</td>
<td>20.169</td>
<td>69.945</td>
<td>1.450</td>
<td>29.005</td>
<td>69.945</td>
</tr>
<tr>
<td>3</td>
<td>.708</td>
<td>14.158</td>
<td>84.103</td>
<td>.708</td>
<td>14.158</td>
<td>84.103</td>
<td>.533</td>
<td>10.654</td>
<td>94.757</td>
</tr>
<tr>
<td>4</td>
<td>.533</td>
<td>10.654</td>
<td>94.757</td>
<td>.533</td>
<td>10.654</td>
<td>94.757</td>
<td>.262</td>
<td>5.243</td>
<td>100.000</td>
</tr>
</tbody>
</table>

Extraction Method: Principal Component Analysis

Table 2. Correlational Matrix of the 5 Dimensions of BFI

<table>
<thead>
<tr>
<th></th>
<th>E</th>
<th>A</th>
<th>C</th>
<th>N</th>
<th>O</th>
</tr>
</thead>
<tbody>
<tr>
<td>E</td>
<td>1</td>
<td>.117</td>
<td>.359**</td>
<td>-.215</td>
<td>.367**</td>
</tr>
<tr>
<td>A</td>
<td>1</td>
<td>.442**</td>
<td>-.467**</td>
<td>.288*</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>1</td>
<td>-.720**</td>
<td>.347**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>1</td>
<td></td>
<td>-.270*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>O</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

E: Extraversion, A: Agreeableness, C: Conscientious-ness, N: Neuroticism, O: Openness

* Significant at the 0.05 level (2-tailed).

** Significant at the 0.01 level (2-tailed).
The data was then subjected to Principal Component Analysis. Prior to this, the suitability of the data was further inspected. The Kaiser Meyer-Olkin value was 0.69 which exceeds the recommended minimum of 0.6 ([13], [14]). Additionally, Bartlett’s Test of Sphericity reached statistical significance, this supports the factorability of the correlation matrix.

Exploratory Factor analysis (Table 1) revealed the presence of 2 Components with eigenvalues exceeding 1, explaining 49.8% and 20.2% of the variance respectively. Other Components had eigenvalues of lower than one (0.70, 0.53 and 0.262) and they account for 14.2%, 10.7% and 5.2% of the variance respectively. Based on the eigen-values and coupled with further analysis using Cartell’s scree test and scree plot, only 2 components were extracted and the rest removed.

The two component solution explained 70% of the variance, with Component 1 contributing 49.8% and Component 2 contributing 20.2%. Further, Varimax rotation was performed (Table 3) and inspection on the rotated component matrix showed strong loadings on Component 1 for “Neuroticism”, “Agreeableness” and “Conscientiousness” with loading values of -0.87, 0.77 and 0.80 respectively. “Extraversion” and “Openness” showed substantial loading on Component 2 with values of 0.88 and 0.73 respectively.

<table>
<thead>
<tr>
<th>Component</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neuroticism</td>
<td>-.868</td>
<td></td>
</tr>
<tr>
<td>Conscientiousness</td>
<td>.797</td>
<td>.351</td>
</tr>
<tr>
<td>Agreeableness</td>
<td>.771</td>
<td></td>
</tr>
<tr>
<td>Extraversion</td>
<td></td>
<td>.875</td>
</tr>
<tr>
<td>Openness</td>
<td></td>
<td>.732</td>
</tr>
</tbody>
</table>

4 DISCUSSION AND CONCLUSION

The results from our Factor Analysis, indicate that there are 2 components that affect the Big Five dimensions. This is consonant with the literature and supports the evidence that there exists a higher order factor of personality ([15], [16]).

Digman [17] refers to these Components as a higher order of factors or superfactors and labels them as Alpha and Beta. While Alpha highly relates to the Big Five dimensions of Neuroticism, Agreeableness and Conscientiousness, Beta relates to Extraversion and Openness. Our results strongly resemble the model proposed by Digman [17]. Furthermore, the Alpha attributes were proposed to have a close association with social development such as stress and negative emotion coping, conforming to social norms and being warm and friendly. While Beta attributes are associated with personal growth such as exploration, flexibility, adapting to novel situations, questioning social norms, seeking out stimulating experiences and having a tendency to experience positive emotions [18].

From this pilot study we can derive that students perceive the Alpha or social elements of a tutor to be more prevalent in the tutor than the Beta or personal growth aspects. This is also further confirmed by existing literature that supports the idea of the affect element, such as teacher’s motivation, beliefs and self-efficacy of a tutor, which are strongly linked to teaching effectiveness [9].

This pilot study, of Asian medical students, provides a number of useful insights:
- It explains which of the BFI dimensions students think are key in a good educator.
- It provides additional evidence for the existence of the superfactors proposed by Digman [17]
- It establishes grounds for further research on Asian Medical Students tutor preferences.

Hence, apart from simply providing knowledge, future AI driven pedagogical agents should be developed as adaptable conversation models, based on personality models described above.
Coupling linguistic principals and computer science approaches to AI-Learner dialogue construction, these conversations can be made potentially more engaging and interactive by providing the affect element that was missing in development of previous pedagogical agents.

In summary, this application of personality on to AI-driven pedagogical agents will assist in bridging the gap that exists in current digital and e-learning platforms; a gap that fosters feelings of alienation and isolation amongst learners and which eventually leads to increased attrition rates [6].

5 FUTURE RESEARCH

Future research in this area would need to focus on: (i) Larger sample size and in different context, (ii) Explore other theoretical underpinnings such as other personality models such as the Myers-Briggs Type Inventory (MBTI) and the Minnesota Multiphasic Personality Inventory (MMPI), (iii) More studies to validate the real-life applications of personality in AI, (iv) Which personality type is most effective in improving engagement in e-learning platforms and (v) Effectiveness in implementing personality in AI pedagogical agents to academic achievements.

REFERENCES

[12] Qualtrics. Provo, Utah, USA: Qualtrics; 2018