

Interactive Learning Environments

ISSN: 1049-4820 (Print) 1744-5191 (Online) Journal homepage: https://www.tandfonline.com/loi/nile20

# Automatic modeling learner's personality using learning analytics approach in an intelligent Moodle learning platform

Ahmed Tlili, Mouna Denden, Fathi Essalmi, Mohamed Jemni, Maiga Chang, Kinshuk & Nian-Shing Chen

**To cite this article:** Ahmed Tlili, Mouna Denden, Fathi Essalmi, Mohamed Jemni, Maiga Chang, Kinshuk & Nian-Shing Chen (2019): Automatic modeling learner's personality using learning analytics approach in an intelligent Moodle learning platform, Interactive Learning Environments, DOI: 10.1080/10494820.2019.1636084

**To link to this article:** https://doi.org/10.1080/10494820.2019.1636084

Published online: 03 Jul 2019.

Submit your article to this journal

Article views: 12

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=nile20

INTERACTIVE LEARNING ENVIRONMENTS
https://doi.org/10.1080/10494820.2019.1636084

Routledge
Taylor & Francis Group

Check for updates

# Automatic modeling learner's personality using learning analytics approach in an intelligent Moodle learning platform

Ahmed Tlili [a], Mouna Denden[b], Fathi Essalmi [b], Mohamed Jemni[b], Maiga Chang[c], Kinshuk[d] and Nian-Shing Chen[e]

[a]Smart Learning Institute of Beijing Normal University, Beijing, People's Republic of China; [b]Research laboratory of Technologies of Information and Communication & Electrical Engineering (LaTICE), Tunis higher school of engineering (ENSIT), University of Tunis, Tunis, Tunisia; [c]School of Computing and Information Systems, Athabasca University, Athabasca, Canada; [d]University of North Texas, Denton, TX, USA; [e]Department of Applied Foreign Languages, National Yunlin University of Science and Technology, Douliou, Taiwan

**ABSTRACT**

The ability of automatically modeling learners' personalities is an important step in building adaptive learning environments. Several studies showed that knowing the personality of each learner can make the learning interaction with the provided learning contents and activities within learning systems more effective. However, the traditional method of modeling personality is using self-reports, such as questionnaire, which is subjective and with several limitations. Therefore, this study presents a new unobtrusive method to model the learners' personalities in an intelligent Moodle (iMoodle) using Learning Analytic (LA) approach with Bayesian network. To evaluate the accuracy of the proposed approach, an experiment was conducted with one hundred thirty-nine learners in a public university. Results showed that recall, precision, F-measure and accuracy values are in acceptance range for three personality dimensions including extraversion, openness, and neuroticism. Moreover, the results showed that the LA approach has a fair agreement with the Big Five Inventory (BFI) in modeling these three personality dimensions. Finally, this study provides several recommendations which can help researchers and practitioners develop effective smart learning environments for both learning and modeling. For example, it is needed to help identify more features of the hardest personality traits, such as agreeableness, using gamification courses.

**ARTICLE HISTORY**
Received 13 November 2018
Accepted 24 March 2019

**KEYWORDS**
Personality; learner modeling; learning analytics; smart learning; assessment; Moodle

## Introduction

Information and Communication Technologies (ICTs) have opened new learning methods for learners, such as online learning where thousands of learners are learning using Learning Management Systems (LMSs). These LMSs provide learners many different types of activities, such as doing assignments, answering quizzes, and engagement in discussions using chats and forums. However, the distributed nature of online learning has raised new challenges. One of the major challenges is low retention with dropout rates of 90% or more (Ebben & Murphy, 2014; Veletsianos & Shepherdson, 2016). This can be due to the lack of personalizing the given learning contents (e.g. course design, motivational aspects, etc.) according to the learners' individual differences and needs (Eriksson, Adawi, & Stöhr, 2017). Of course, unlike classrooms, it becomes much harder for teachers in online learning environments to supervise, control and adjust the learning process for thousands of learners

**CONTACT** Ahmed Tlili ahmed.tlili23@yahoo.com
© 2019 Informa UK Limited, trading as Taylor & Francis Group

(Vozniuk, Govaerts, & Gillet, 2013). The National Academy of Engineering (2014) stated that providing personalized learning is one of the fourteen most important challenges of the twenty-first Century. Many personalization parameters are reported in the literature, which are used by various personalized learning systems. One of these parameters is "personality" which is widely identified as an important indicator of individual differences (Irani, Telg, Scherler, & Harrington, 2003). Kim, Lee, and Ryu (2013) have argued that personality can affect preferences of learning materials as well as the way of processing information and making decisions. Tlili, Essalmi, Jemni, Kinshuk and Chen (2016) highlighted the importance of considering the learner's personality in computer-based learning.

While the traditional and most used method of modeling the learner's personality is questionnaire (Tlili et al., 2016), it is possible to use the learners' learning actions and data in LMS. The LMS activity log file represents learners' online learning behavior. This file can then be used to implicitly model the learners' personalities. Once learners' personalities are modeled, an adaptive system can provide personalized learning contents for them based on their personalities. The analysis of learning activity log data is often referred to as Learning Analytics (LA), which is defined as "the measurement, collection, analysis and reporting of data about learners and their context, for purposes of understanding and optimizing learning and the environments in which it occurs" (Siemens & Long, 2011). This study presents an ongoing project to make Modular Object-Oriented Dynamic Learning Environment (Moodle) intelligent (iMoodle) by implicitly modeling learners' personalities based on their learning data and using LA approach with Bayesian network. To model learners' personalities, this study adopts the Five Factor Model (FFM) which is a widely known psychological model in the literature (McCrae & John, 1992). A pilot experiment is then conducted at a public university to evaluate the accuracy of the developed approach for learners' personalities modeling using iMoodle. The personalization process is beyond the scope of this study.

The remainder of this paper is as follows: Section two presents a literature review regarding personality and LA. Section three presents the architecture of our iMoodle based on the LA approach. Section four describes the conducted experiment, while Section five reports the obtained results and discuss them. Finally, Section six concludes the paper with a summary of the findings, limitations and future directions based on this research.

## Related work

In order to achieve the objective of modeling learners' personalities based on their LMS data, this section starts with reviewing literature in relation to personality and LA.

### *Personality*

While there is no agreed upon definition of personality in the literature, two of the classic definitions belong to Allport (1961) and Child (1968). Allport (1961) considered personality as a unique psychological system located inside individuals. Child (1968) on the other hand considered personality as an internal factor that gives consistency over time for the individual's behavior. According to Zafar and Meenakshi (2012), personality is an integrated part of individuals. It comes with them to a particular situation and leaves with them when they go. Clarkson and Clarkson (1996) considered personality in their book as "the way one thinks-that is, how you gather information, organize it, and make decisions with it." Personality accounts for the "natural differences" among learners and teachers driving how information is perceived and acted on (Wankat & Oreovicz, 2004). Bayne (2004) claimed that the differences of learners' personalities result in different ways of learners' involvement in the learning progress regardless of their personal interests or the degree of cognitive development. Kolb (1984) expanded the experiential learning theory by incorporating aspects of personality type theory. Tlili et al. (2016) highlighted the importance of taking into consideration the learners' personalities in computer-based learning environments.

Various personality models are reported in the literature to understand individuals' behaviors and characteristics. One of these models adopted in this study is the Five Factor Model (FFM). FFM is the most used psychological model (Franić, Borsboom, Dolan, & Boomsma, 2014). It attributes a variety of personality characteristics to five dimensions as follows:

Extraversion refers to individual's degree of activeness, assertiveness, interpersonal skills, warmth, energetic, sociability, enthusiastic, outgoing, talkative and positive emotions. People high in extraversion are characterized as more optimist, energetic, tend to show high level of commitment to social groups and activities (Watson & Clark, 1997), risk takers (Walsh, 2012) and prefer hot colors (Choungourian, 1967). Furthermore, they are considered as more interested in details (Laney, 2002).

Agreeableness refers to the way in which a person interacts with his/her environment in terms of compliance, trust, altruism, kindliness, modesty and generosity. People high in agreeableness tend to be more willing to help others, cooperative, sympathetic and confident (McCrae & John, 1992). This dimension also relies on the use of online conversations (Okdie, Guadagno, Bernieri, Geers, & Mclarney-Vesotski, 2011). Specifically, people low in agreeableness are more likely to use online conversation because it allows them to hide their disagreeable nature and communicate more effectively compared to the face to face communication.

Conscientiousness refers to individual's degree of self-discipline, orderliness, organization and achievement striving. People high in conscientiousness are characterized as more organized, punctual, hardworking, ambitious and responsible (Patrick, 2011). Therefore, they may have high task performance and job satisfaction levels (Barrick & Mount, 1991) and better academic results (Busato, Prins, Elshout, & Hamaker, 2000; Heaven, Mak, Barry, & Ciarrochi, 2002). On the other hand, they are less risk takers (James & Mazerolle, 2002; Raja & Johns, 2004).

Neuroticism refers to individual's degree of emotional stability, anxiety, hostility, depression, impulsivity, self-consciousness and emotional vulnerability. People high in neuroticism tend to be more worrying, less satisfied with their work and evoke more negative life events (Emmons, Diener, & Larsen, 1985). Neuroticism is also positively correlated with attitudes toward inaction when facing challenging tasks (Ireland, Hepler, Li, & Albarracín, 2015).

Openness to experience refers to individual's degree of intellectual curiosity, imagination, interest in new experiences, originality (McCrae & John, 1992; Watson & Clark, 1997). People high in openness tend to be more logic, creative and seek out new experiences. Openness is also positively correlated with learning motivation (Barrick, Mount, & Judge, 2001) and academic success (De Fruyt & Mervielde, 1996; Farsides & Woodfield, 2003; Schuerger & Kuna, 1987). This means that people high in openness are more likely to be motivated to learn and have better academic results.

To provide personalized learning process based on personality, the learner's personality should be modeled first. Thus, the next subsequent section presents how learners' personalities are modeled in computer-based learning environments.

### *Personality modeling*

Based on a literature review, Tlili et al. (2016) found that the most used method of modeling the learner's personality is the self-report, namely questionnaire. Despite that this method is accurate in modeling the learners' personalities, it can be not motivating and the learners may not reveal their true information, especially when they think that they will not benefit from answering (Chen & Lin, 2017). Therefore, several researchers have reported that using behavioral patterns may be more effective in modeling personality (Chen & Lin, 2017; Scherer & Giles, 1979; Vinciarelli & Mohammadi, 2014). For instance, Essalmi, Tlili, Ayed, and Jemni (2017), and Bunian, Canossa, Colvin and Seif El-Nasr (2017) have used gaming behaviors to model the learner's personality. Also, Gao et al. (2013) have used social media behaviors to model personality.

In Moodle platform, personality can be modeled using a personality test add-on to the questionnaire plug-in. This plug-in allows creating a questionnaire to be answered on the Moodle.

Additionally, Moodle offers several learning analytics tools which mainly focus on assessing learners' performance and evaluating different skills and competencies. For instance, GISMO (Dietz-Uhler & Hurn, 2013) is a visualization tool for Moodle which is used by teachers to analyze the learning process of all learners. It is incorporated within Moodle as an additional block. It generates graphical representations to evaluate learners' behaviors, based on their log data. Besides, several researchers, such as Conijn, Snijders, Kleingeld, and Matzat (2017), have used learning analytics in Moodle to predict learner performance. However, to the best of our knowledge, no research or Moodle tool is reported in the literature which aims to implicitly model learners' personalities based on their online LMS data. Therefore, this study presents an intelligent Moodle (iMoodle) which implicitly models the learners' personalities based on their LMS data and using an LA approach based on Bayesian Network (BN). BN is one of the most used methods to deal with the uncertainty of the learner model (Chrysafiadi & Virvou, 2013). It is a direct acyclic graph where nodes represent the variables and arcs represent the probabilistic correlation between variables (Pearl, 1988). BN is considered as a powerful tool for knowledge representation (Cheng et al., 2002). One of its advantages is its ability to combine different sources of knowledge and their suitability for small and incomplete data sets (Khodakarami & Abdi, 2014). The architecture of iMoodle is presented in the next section.

## Architecture of iMoodle

As shown in Figure 1, iMoodle differs from the classic Moodle in giving immediate learning assistance for teachers and help them control the learning process through providing immediate dashboards. Additionally, iMoodle models at-risk learners, learners who may fail to pass their final exams, and gives them additional personalized learning contents as notifications (Tlili, Essalmi, Jemni, Chang, & Kinshuk, 2018). Furthermore, iMoodle aims to model learners' personalities to provide later on personalized learning contents and gamified elements. Specifically, iMoodle includes an LA system

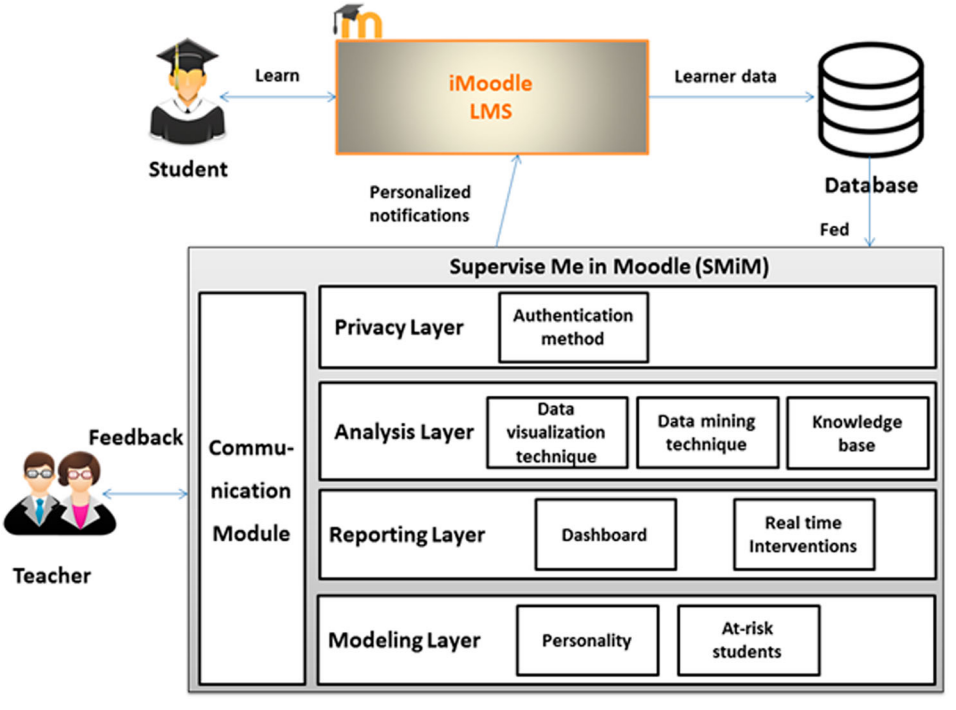



**Figure 1.** iMoodle architecture.

named Supervise Me in Moodle (SMiM) to achieve the functionalities mentioned above. This study mainly focuses on modeling the learners' personalities based on their LMS data, as described below.

In order to implicitly model learners' personalities using LA approach based on the BN approach, we must first build a graph that contains different features associated to each personality dimension and the relationship between them. Second, we must indicate the probability of the relationships strength previously modeled on each node in the graph. Therefore, the first step is identifying the learners' key features that are worth modeling and their states.

As this is an exploratory study, these features were identified based on the characteristics of each personality dimension reported in the FFM (presented in the personality section). For instance, a learner with high extraversion is more likely to have many friends, more interested in details, like to speak with others and be active. This would make them actively involved in chats and forums activities by creating new posts, replying others' posts, and read supplementary course materials (give more course details). Therefore, we infer these behaviors, namely participation in forums (PF), participation in chat (PC) and access to supplementary course materials (ASCM), as the key features of extraversion in our LA approach. Additionally, learners high in agreeableness are cooperative and tend to help others. Therefore, we analyze chat and forum behaviors, namely participation in forums (PF) and participation in chat (PC) as the key features for identifying agreeableness dimension. Learners high in conscientiousness are more likely to be more organized and punctual, less risk-taking during learning and have better academic results. This would make them frequently enter the platform (iMoodle) and complete the learning assignments in time. Additionally, learners high in conscientiousness would take all their time to study learning assignments correctly in order to secure good academic results. Therefore, we infer delay in assignment delivering, score in homework assignment including quizzes, the accomplishment degree of assignments (whether a learner finished all the learning assignments or not), time used in solving the quizzes and number of the entrance to the system as the key features for identifying conscientiousness. Table 1 presents the definitions and the extracted key features for each personality dimension and the states (the range of values of the extracted features).

The key features for each personality dimension, presented in Table 1, are then encoded in the network structure, as shown in Figure 2. This network models the relationship between the learners' identified features and each personality dimension. For each node feature, different possibilities of states are mentioned. For instance, PC has two possible states, namely participation and no participation.

In the second step of building a BN, the probability values of each node in the different conditional probability tables should be computed. This was done via a training dataset based on experimental results, as shown in Figure 3. Specifically, fifty learners answered the Big Five Inventory (BFI) to identify their personalities (BFI is further detailed in the experimental section). These learners then learned using iMoodle platform where their learning behavior data was stored along with their personality results (already identified using BFI) to determine the conditional parameters of the BN. Table 2 shows the Conditional Probability Table (CPT) for the "Extraversion" node. For example, Table 2 indicates that if a learner accesses and reads many supplementary materials and participates in chat, this particular learner is identified to be high extraversion with a probability of 64%.

The collected data are then fed to the developed LA system based on BN. In this context, the following Bayesian rule was used.

$$P(Cj|d) = (P(d|Cj)P(Cj))/(P(d))$$

where $P(Cj|d)$ is the posterior probability of instance $d$ being in class $Cj$; $P(d|Cj)$ is the likelihood, which is the probability of generating instance $d$ given a class $Cj$; $P(Cj)$ is the prior probability of occurrence of class $Cj$; $P(d)$ is the prior probability of occurrence of instance $d$.

Figure 3 summarizes the above description of LA approach with Bayesian network to model the learner's personality.

**Table 1.** The key features for each personality dimension.

| Dimension | Feature | Definition | States |
|---|---|---|---|
| Extraversion | Participation in forums (PF) | Learner participation in forums. | • Participation<br>• No participation |
| | Participation in chat (PC) | Learner participation in chat rooms. | • Participation<br>• No participation |
| | Access to supplementary course materials (ASCM) | Learner access to supplementary course materials which is optional. | • Many: more than 75% of materials<br>• Few: between 25% and 75%<br>• None |
| Agreeableness | Participation in forums (PF) | Learner participation in forums. | • Participation<br>• No participation |
| | Participation in chat (PC) | Learner participation in chat rooms. | • Participation<br>• No participation |
| Conscientiousness | Delay in assignment delivering (DAD) | The delay in delivering the given assignments. | • No: final date is not exceeded<br>• Yes: final date is exceeded |
| | Score (S) | The learner score in course assignments. | • High: greater than 15 (in a 1–20 scale)<br>• Medium: between 10 and 15<br>• Low: Less than 10 |
| | Accomplishment of assignments (AOA) | The accomplishment of all the assignments in a course. | • Yes: all the assignments are completed<br>• No: not all the assignments are completed |
| | Time solving the quizzes (TSQ) | The total amount of time the learner spent on the course quiz pages. | • High: more than 75% of the time assigned for the quiz.<br>• Medium: between 50% and 75% of the time assigned for the quiz.<br>• Low: less than 50% of the time assigned for the quiz. |
| | Number of entrance to the system (NES) | The number of entrance to the system in a week. | • High: 6 or more times<br>• Medium: between 3 and 5 times<br>• Low: 1 or 2 times |
| Neuroticism | Participation in chat (PC) | Learner participation in chat rooms. | • Participation<br>• No participation |
| | Number of entrance to the system (NES) | The number of entrance to the system in a week. | • High: 6 or more times<br>• Medium: between 3 and 5 times<br>• Low: 1 or 2 times |
| | Delay in assignment delivering (DAD) | The delay in delivering the assignment. | • No: final date is not exceeded<br>• Yes: final date is exceeded |
| Openness | Time solving the quizzes (TSQ) | The total amount of time the learner spent on the course quiz pages. | • High: more than 75% of the time assigned for the course.<br>• Average: between 50% and 75% if the time assigned for the course.<br>• Low: less than 50% of the time assigned for the course. |
| | Score (S) | The learner score in course assignments. | • High: greater than 15 (in a 1–20 scale)<br>• Medium: between 10 and 15<br>• Low: less than 10 |
| | Access to supplementary course materials (ASCM) | Learner access to supplementary course materials which is optional. | • Many: more than 75% of materials<br>• Few: between 25% and 75%<br>• None |

Additionally, to overcome this "zero frequency problem" in a BN, the LA approach uses the statistical technique namely the "Laplace smoothing", which adds one to each count (Manning, Raghavan, & Schütze, 2008). Specifically, the LA system SMiM applies data visualization, specifically pie chart, to show teachers the personality distribution of their class, hence provide the needed interventions accordingly. It also applies several strategies to avoid LA design issues from the data preparation perspective, highlighted in (Tlili, Essalmi, Jemni, & Chen, 2018). For instance, to protect the learner's privacy, SMiM uses authentication methods to allow only authorized persons to have access to the collected data and results. Additionally, since the collected data and the obtained analytics results,

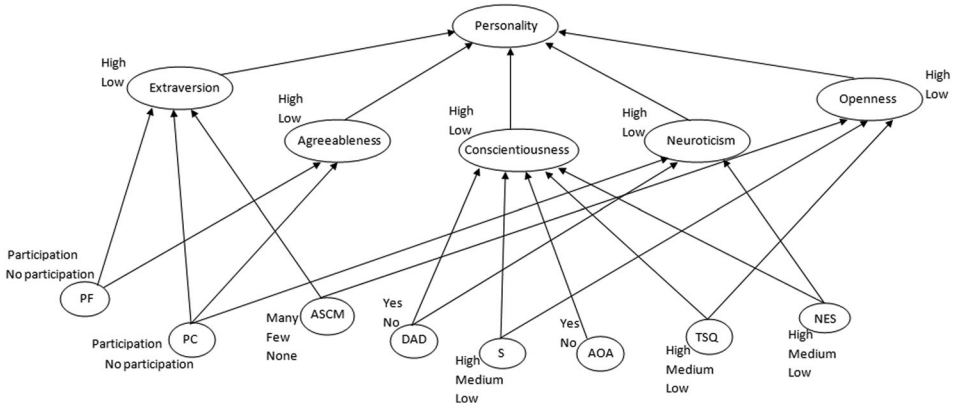



**Figure 2.** Bayesian network for modeling learners' personalities.

recommendations and interventions must have a pre-defined time for how long they are going to be stored and used, the collected data and generated reports within SMiM are stored for a pre-defined period (one academic year) before they are automatically deleted. Moreover, to make the applied LA process more transparent for learners, SMiM gives them the possibility to see their collected LMS data. The next section presents the followed experimental method to validate this study.

## Method

A pilot experiment was conducted at a public Tunisian University in order to model learners' personalities based on their collected LMS data. This section presents the participants of this experiment. In addition, it describes the followed procedure, the used instruments and the data analysis.

### *Participants*

Participants of this study were one hundred and thirty-nine undergraduate learners (55 males and 84 females) majoring in computer science and aged between eighteen and twenty-three. Sixty-four

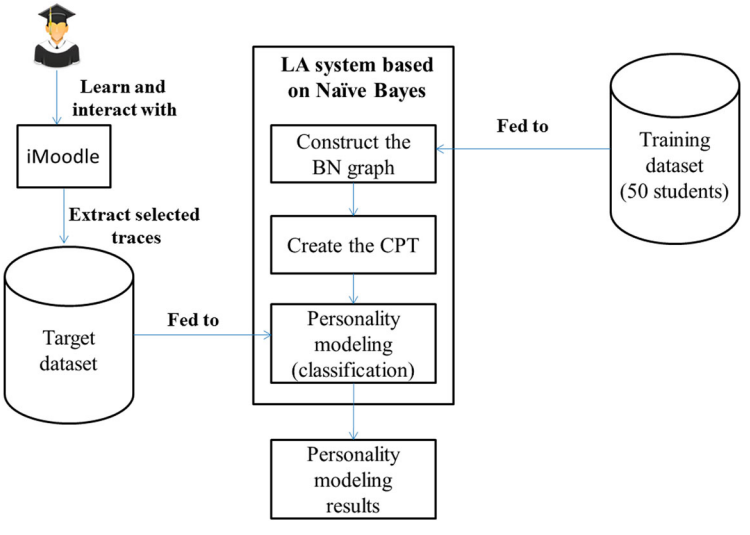



**Figure 3.** Modeling learner's personality using LA approach with Bayesian network.

**Table 2.** CPT for the extraversion node.

| | | | Extraversion | |
|---|---|---|---|---|
| ASCM | PF | PC | High | Low |
| Many | Participation | Participation | 0.2 | 0.4 |
| | Participation | No participation | 0 | 0.04 |
| | No participation | Participation | 0.64 | 0.4 |
| | No participation | No participation | 0.04 | 0.08 |
| Few | Participation | Participation | 0.08 | 0 |
| | Participation | No participation | 0 | 0 |
| | No participation | Participation | 0 | 0.04 |
| | No participation | No participation | 0.04 | 0 |
| None | Participation | Participation | 0 | 0 |
| | Participation | No participation | 0 | 0 |
| | No participation | Participation | 0 | 0 |
| | No participation | No participation | 0 | 0 |

learners were enrolled in the Basic Software course (BS) which was delivered to them in their first year (out of three), fifty-six learners were enrolled in the Object Oriented Design Methodology course (OODM) which was delivered to them in their second year (out of three) and nineteen learners were enrolled in the Information Monitoring Methodology course (IMM) which was delivered to them in their last year. The three taught courses were prepared by the same teacher where: (1) Object Oriented Design Methodology (OODM) aims to help learners learn the Unified Modeling Language (UML) diagrams, such as use case and class diagrams; (2) Basic Software (BS) aims to help learners learn the assembly language and computer architecture; and (3) Information Monitoring Methodology (IMM) aims to help learners learn monitoring techniques to collect the required information which helps in decision making.

### *Experimental design and data analysis*

As shown in Figure 4, at the beginning of each course, learners' personalities were modeled using the Big Five Inventory (BFI) which is a validated and widely used questionnaire in the literature (John, Donahue, & Kentle, 1991). This took between ten and fifteen minutes. It is a 5 points Likert-scale questionnaire from 1 (strongly disagree) to 5 (strongly agree). The learners then took the provided courses (BS, OODM, IMM) using iMoodle platform for the rest of the semester (three months). After that, the learners' personalities were modeled based on their learning data (on iMoodle) using the developed LA approach with Bayesian network. Finally, the personality modeling results obtained from both the BFI and the developed approach were compared.

Since there is no guidance in the BFI scoring for determining whether an individual has a high or low personality trait (e.g. high extraversion or low extraversion), the standard *z*-score was computed. Learners with $z > 0$ were considered as learners with high personality measures (e.g. high extrovert, high openness, etc.), while learners with $z < 0$ were considered as learners with low personality measures (e.g. low extrovert, low openness, etc.). In our case, no learners were found with $z = 0$. Table 3 presents the mean, standard deviation and Cronbach's alpha value for each personality trait. As shown in Table 3, the questionnaire's measurements are reliable since all Cronbach's alpha values are above 0.7 (Yu, 2001).

## Results and discussion

In personality modeling, the main goal is personality level identification (high or low). Therefore, Chi-square test is used as an assessment criterion to compare between the assessed personality levels from the results of LA approach and that assessed from the BFI results. The obtained Chi-square test results are presented in Table 4. As shown in Table 4, there is no significant difference

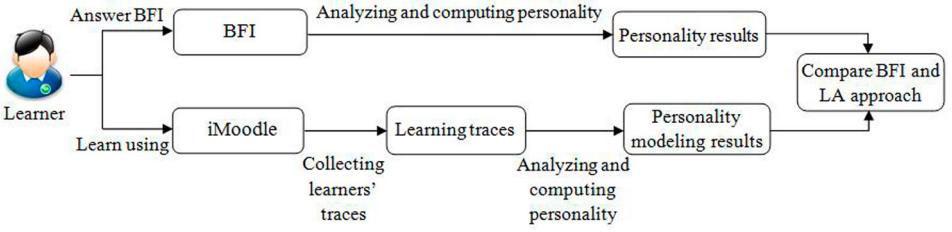



**Figure 4.** Experimental procedure.

**Table 3.** Personality traits mean and standard deviation ($N = 139$).

| | Male ($n = 55$) | | Female ($n = 84$) | | Overall | |
|---|---|---|---|---|---|---|
| Personality | Mean | SD | Mean | SD | Mean | SD |
| Extraversion, $\alpha = .88$ | 3.23 | .53 | 3.27 | .55 | 3.26 | .55 |
| Agreeableness, $\alpha = .82$ | 3.71 | .53 | 4.06 | .42 | 3.94 | .49 |
| Conscientiousness, $\alpha = .84$ | 3.62 | .69 | 3.63 | .58 | 3.62 | .62 |
| Neuroticism, $\alpha = .78$ | 2.50 | .71 | 2.78 | .69 | 2.68 | .71 |
| Openness, $\alpha = .86$ | 3.73 | .40 | 3.61 | .41 | 3.65 | .41 |

between the developed LA approach and BFI in modeling four personality dimensions, namely conscientiousness, extraversion, openness and neuroticism (In agreeableness, $p$ value is less than .05).

Additionally, the personality classification results of LA approach are compared to the obtained results from BFI (validated instrument). In a classification process, four variables should be computed to assess and compare the results of a classifier (our system) with trusted external judgments, namely: (1) number of True Positives (TP) is the number of instances correctly labeled as belonging to a given class; (2) number of False Positive (FP) is the number of instances incorrectly labeled as belonging to a given class; (3) number of False Negative (FN) is the number of instances which were not labeled as belonging to the class, but should have been; and, (4) number of True Negative (TN) is the number of instances that were not labeled as belonging to the class and should not have been. Finally, the precision, recall, F-measure and accuracy are defined and calculated, as mentioned in Olson and Delen (2008), based on the above variables, as follows:

- Recall (also known as True Positive Rate) is the percentage of the instances that are correctly classified within a class, over the total number of instances belonging to that class.

$$\text{Recall} = \frac{TP}{TP + FN}$$

**Table 4.** Chi-square test results.

| Personality traits | Level | Frequency in BFI | Frequency in our system | Chi-square value ($\chi^2$) | $P$ value |
|---|---|---|---|---|---|
| Conscientiousness | High | 64 | 71 | .05 | .26 |
| | Low | 75 | 68 | | |
| Extraversion | High | 75 | 58 | .3 | .39 |
| | Low | 64 | 81 | | |
| Openness | High | 73 | 92 | .06 | .55 |
| | Low | 66 | 47 | | |
| Agreeableness | High | 86 | 11 | 89.06 | .01 |
| | Low | 53 | 128 | | |
| Neuroticism | High | 68 | 71 | .18 | .11 |
| | Low | 71 | 68 | | |

- Precision is the proportion of the instance, which correctly belongs to a class, over the entire number of instances that were classified in that class

$$\text{Precision} = \frac{TP}{TP + FP}$$

- F-Measure is computed as a combined measure of precision and recall.

$$\text{F - measure} = \frac{2}{(1/\text{precision}) + (1/\text{recall})}$$

- Accuracy is the percentage of correctly classified instances

$$\text{Accuracy} = \frac{TP}{TP + FP + TN + FN}$$

Using 10-fold cross-validation, the obtained results of precision, recall, F-measure of the proposed LA approach are presented in Table 5. As shown in Table 5, the developed LA approach accuracy is above 0.5 in modeling three personality dimensions, namely extraversion, openness and neuroticism. Specifically, recall should be used to illustrate the overall performance of the approach. It is seen from Table 5 that the recall of modeling high agreeableness personality is very low (.09).

Furthermore, to determine the agreement degree or the inter-rater reliability between the developed LA approach and BFI, the Cohen's Kappa (K) variable (Cohen, 1960) is calculated. Landis and Koch (1977) stated that if Kappa < 0 indicates no agreement, from 0.0 to 0.2 indicates slight agreement, from 0.21 to 0.40 indicates fair agreement, from 0.41 to 0.60 indicates moderate agreement, from 0.61 to 0.80 indicates good agreement, and from 0.81 to 1.0 indicates perfect agreement. Table 6 presents the results of the Kappa variable. As shown in Table 6, the LA approach has a fair agreement with BFI in modeling personality, specifically in three personality dimensions, namely extraversion, openness and neuroticism. However, the LA approach has a slight agreement with BFI in modeling the personality dimensions conscientiousness and agreeableness. Particularly, it is seen in Table 5 that these two dimensions have the lowest accuracy rate and recall.

To better understand and interpret the obtained results, further discussions are described as follows:

Most studies in computer-based learning research were using questionnaires to model the learner's personality, very few studies tried to automatically model the learners' personalities based on their learning traces (Tlili et al., 2016). Particularly, the previous studies did not evaluate the accuracy of the constructed learner model by comparing their results with the results from other already validated instruments, like BFI in this case (Chen & Lin, 2017). Additionally, to the best of our knowledge, no study has reported the application of LA to automatically model learner's personality in Moodle platform. Therefore, this study has closed these two gaps by automatically modeling learners' personalities while they are learning in the Moodle platform, and the accuracy of the obtained

**Table 5.** Personality modeling results using the LA approach.

| Personality traits | Level | Precision | Recall | F-measure | Accuracy |
|---|---|---|---|---|---|
| Conscientiousness | High | .43 | .50 | .46 | .48 |
| | Low | .53 | .46 | .49 | |
| Extraversion | High | .61 | .44 | .51 | .53 |
| | Low | .5 | .65 | .56 | |
| Openness | High | .59 | .48 | .52 | .53 |
| | Low | .5 | .6 | .54 | |
| Agreeableness | High | .66 | .09 | .15 | .41 |
| | Low | .39 | .92 | .54 | |
| Neuroticism | High | .5 | .52 | .5 | .51 |
| | Low | .52 | .49 | .5 | |

**Table 6.** Cohen's kappa agreement results.

| Personality | Cohen's kappa value | Agreement |
|---|---|---|
| Conscientiousness | .16 | slight |
| Extraversion | .36 | fair |
| Openness | .35 | fair |
| Agreeableness | .1 | slight |
| Neuroticism | .34 | fair |

learners' personalities is also evaluated. The findings showed that extraversion and openness traits have the highest identification accuracy with an accuracy of 53%, followed by neuroticism with an accuracy of 51%. Agreeableness and conscientiousness, however, are the hardest trait to model, with an accuracy of 41% and 48% respectively. Interestingly, a previous study on automatically modeling personality using linguistic cues (not using a learning management system like this study) had also found that extraversion and openness were the easiest traits to model (Mairesse, Walker, Mehl, & Moore, 2007). However, some of the findings are also different from the previous studies. For instance, in the study of Chen, Davis, Hauff and Houben (2016) where they reported a difficulty in modeling the agreeableness personality dimension due to all the selected features were not correlated to this dimension; this study reported an opposite finding and managed to model the agreeableness dimension from the selected features with a low accuracy rate of 41%.

To further understand the obtained modeling results using the developed LA approach, the learners' data was analyzed. It is seen that 93 % of the learners did not use the chat and forums. This can be because the given three courses did not promote the use of chat and forum. Consequently, the obtained personality modeling results may be biased, for instance, in agreeableness personality modeling results which is based on the forum and chat traces, as shown in Table 1. Additionally, as shown in Table 7, the majority of the obtained wrong personality modeling results was from females. In this context, Ramírez-Correa, Arenas-Gaitán, and Rondán-Cataluña (2015) highlighted the effect of gender on learners' learning behaviors in online courses.

Since this study is exploratory and little information is previously known, the obtained experimental data from this pilot experiment was validated using three methods namely, Chi-square, 10-fold cross-validation and Cohen's Kappa. The results showed that the three personality dimensions, namely extraversion, openness, and neuroticism have an accuracy rate above 50% compared to the BFI. Additionally, these three dimensions have a "fair" agreement with BFI based on calculating Kappa for inter-reliability or agreement (above 0.30). Furthermore, no significant difference was found, using Chi-square, between the developed LA approach and BFI in modeling these dimensions (extraversion, openness and neuroticism). Therefore, we can conclude that recall, precision, F-measure and accuracy, are considered in an acceptable range for these three personality dimensions. We can also conclude that only these three dimensions are considered with reliable results using the designed LA approach. Besides, the analysis of these results has revealed some promising recommendations that should be considered by researchers, educators and practitioners to enhance the automatic modeling process of learners' personalities in learning management systems, such as Moodle. When a course is designed to be used for this purpose, there are some recommendations for achieving a better result as described below.

**Table 7.** Distribution of wrong personality results between male and female gender.

| | Number of wrong personality modeling results | |
|---|---|---|
| Personality | Female (%) | Male (%) |
| Extraversion | 69.7 | 30.3 |
| Agreeableness | 71.4 | 28.6 |
| Conscientiousness | 70 | 30 |
| Neuroticism | 62.1 | 37.9 |
| Openness | 69.1 | 30.9 |

- Since the participation rate in the forum and chat was very low in this study, it is recommended that collaborative learning activities using forum and chat facilities should be designed to promote participation. This can help collecting rich and representative data from the chat and forum activities, which would then help in modeling the learner's personality.
- Several characteristics of some personality dimensions (e.g. achiever for conscientiousness dimension) could not be linked with any feature in our course. Therefore, course designers should consider several properties that may help in generating more learning traces to facilitate personality modeling. For instance, gamifying a course, by adding some game design elements, can help identify more features of the hardest personality traits. Specifically, using the badge game element and the number of collected badges while learning may reveal whether a learner is an achiever or not. This trace can help in modeling the conscientiousness dimension which may not be found in a non-gamified course.
- Since the majority of misidentified results were found in female learners, gender differences should be considered carefully while building a knowledge base for a personality modeling system.
- Since there were only two features used for automatically identifying the agreeableness dimension, the results may be biased. It is recommended that each personality dimension to be automatically modeled using learning analytics approach should be based on several features to enhance the accuracy results and decrease the bias in a particular feature.

## Conclusion, limits and future directions

The traditional method of modeling learners' personalities is self-report using questionnaires. However, this method may include some incertitude because people may be influenced by several factors, such as having wrong perceptions about themselves (Tlili et al., 2016). On the other hand, people's behaviors may reflect their habits and believe. Therefore, this study presents a new unobtrusive method of modeling the learners' personalities in an intelligent Moodle (iMoodle) using learning analytics approach. The obtained results showed that iMoodle based on the LA approach with Bayesian network can model learners' personalities with an acceptable precision and a fair agreement compared to BFI for only three personality dimensions, namely, extraversion, openness and neuroticism.

This study can advance research in the smart learning field by developing environments which can be used for both learning and automatically modeling learners' personalities. Particularly, this can be useful for smart learning systems to adapt, for instance, the proposed teaching strategies according to each learner' personality. Additionally, this study can advance research in the educational technology field by offering new tools which can be used to model the learner's personality instead of the traditional method, namely questionnaire. Furthermore, this study highlights several recommendations that may help researchers and practitioners develop effective smart learning environments which can be used for learning and modeling at the same time.

This study on the other hand has several limitations that should be acknowledged and further researched. For instance, this study used only one algorithm (Naive Bayes algorithm) to model learners' personalities. Additionally, this study did not consider gender while automatically modeling personality. However, despite these limitations, this study provides a solid ground for researchers and practitioners to develop smart learning environments for learning and automatically modeling personality using learning analytics.

Future research directions may focus on: (1) enhancing the accuracy and reliability results of our LA approach by implementing new algorithms and considering the above-mentioned recommendations; and, (2) providing adaptive learning within iMoodle platform based on each modeled learner's personality. For instance, a learner with high neuroticism tends to be more worry and less satisfied with his/her work. This can be overcome by automatically providing game elements that

promote confidence while learning, such as points and badges. Consequently, this learner may feel more confident so as to have better learning outcomes.

## Disclosure statement

No potential conflict of interest was reported by the authors.

## ORCID

*Ahmed Tlili* http://orcid.org/0000-0003-1449-7751
*Fathi Essalmi* http://orcid.org/0000-0002-1153-787X

## References

Allport, G. W. (1961). *Pattern and growth in personality*. New York, NY: Holt, Rinehart & Winston.
Barrick, M. R., & Mount, M. K. (1991). The big five personality dimensions and job performance: A meta-analysis. *Personnel Psychology*, *44*, 1–26.
Barrick, M. R., Mount, M. K., & Judge, T. A. (2001). Personality and performance at the beginning of the new millennium: What do we know and where do we go next? *International Journal of Selection and Assessment*, *9*, 9–30.
Bayne, R. (2004). *Psychological types at work: An MBTI perspective: Psychology@work series*. London: International Thomson Business.
Bunian, S., Canossa, A., Colvin, R., & El-Nasr, M. S. (2017). *Modeling individual differences in game behavior using HMM*. Proceedings of the 13th AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment (AIIDE-17). Palo Alto, CA: AAAI Press.
Busato, V. V., Prins, F. J., Elshout, J. J., & Hamaker, C. (2000). Intellectual ability, learning style, achievement motivation and academic success of psychology students in higher education. *Personality and Individual Differences*, *29*, 1057–1068.
Chen, G., Davis, D., Hauff, C., & Houben, G. J. (2016). *On the impact of personality in massive open online learning*. Proceedings of the 2016 conference on user modeling adaptation and personalization (pp. 121–130). New York, NY: ACM.
Chen, Z., & Lin, T. (2017). Automatic personality identification using writing behaviours: An exploratory study. *Behaviour & Information Technology*, *36*(8), 839–845.
Cheng, J., Greiner, R., Kelly, J., Kelly, J., Bell, D., & Liu, W. (2002). Learning Bayesian networks from data: An information-theory based approach. *Artificial Intelligence*, *137*(1–2), 43–90.
Child, I. L. (1968). Personality in culture. In E. Borgatta & W. W. Lambert (Eds.), *Handbook of personality theory and research* (pp. 80–101). Chicago, IL: Rand McNally.
Choungourian, A. (1967). Introversion — extraversion and color preferences. *Journal of Projective Techniques and Personality Assessment*, *31*(4), 92–94.
Chrysafiadi, K., & Virvou, M. (2013). Student modeling approaches: A literature review for the last decade. *Expert Systems with Applications*, *40*(11), 4715–4729.
Clarkson, C., & Clarkson, S. (1996). *Educating the whole hearted child: A handbook for Christian home education*. Walnut Springs, TX: Whole Heart Ministries.
Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*(1), 37–46.
Conijn, R., Snijders, C., Kleingeld, A., & Matzat, U. (2017). Predicting student performance from LMS data: A comparison of 17 blended courses using Moodle LMS. *IEEE Transactions on Learning Technologies*, *10*(1), 17–29.
De Fruyt, F., & Mervielde, I. (1996). Personality and interests as predictors of educational streaming and achievement. *European Journal of Personality*, *10*, 405–425.
Dietz-Uhler, B., & Hurn, J. E. (2013). Using learning analytics to predict (and improve) student success: A faculty perspective. *Journal of Interactive Online Learning*, *12*(1), 17–26.
Ebben, M., & Murphy, J. S. (2014). Unpacking MOOC scholarly discourse: A review of nascent MOOC scholarship. *Learning, Media and Technology*, *39*(3), 328–345.
Emmons, R. A., Diener, E., & Larsen, R. J. (1985). Choice of situations and congruence models of interactionism. *Personality and Individual Differences*, *6*, 693–702.
Eriksson, T., Adawi, T., & Stöhr, C. (2017). "Time is the bottleneck": A qualitative study exploring why learners drop out of MOOCs. *Journal of Computing in Higher Education*, *29*(1), 133–146.
Essalmi, F., Tlili, A., Ayed, L. J. B., & Jemni, M. (2017). Toward modeling the learner's personality using educational games. *International Journal of Distance Education Technologies)*, *15*(4), 21–38.
Farsides, T., & Woodfield, R. (2003). Individual differences and undergraduate academic success: The roles of personality, intelligence, and application. *Personality and Individual Differences*, *34*(7), 1225–1243.

Franić, S., Borsboom, D., Dolan, C. V., & Boomsma, D. I. (2014). The big five personality traits: Psychological entities or statistical constructs? *Behavior Genetics*, *44*(6), 591–604.
Gao, R., Hao, B., Bai, S., Li, L., Li, A., & Zhu, T. (2013). *Improving user profile with personality traits predicted from social media content*. Proceedings of the 7th ACM conference on Recommender systems (pp. 355–358).
Heaven, P. C. L., Mak, A., Barry, J., & Ciarrochi, J. (2002). Personality and family influences on adolescent attitudes to school and self-rated academic performance. *Personality and Individual Differences*, *32*, 453–462.
Irani, T., Telg, R., Scherler, C., & Harrington, M. (2003). Personality type and its relationship to distance education students' course perceptions and performance. *Quarterly Review of Distance Education*, *4*(4), 445.
Ireland, M. E., Hepler, J., Li, H., & Albarracín, D. (2015). Neuroticism and attitudes toward action in 19 countries. *Journal of Personality*, *83*(3), 243–250.
James, L. R., & Mazerolle, M. D. (2002). *Personality in work organizations*. Thousand Oaks, CA: Sage.
John, O. P., Donahue, E. M., & Kentle, R. L. (1991). *The big five inventory–versions 4a and 54*. Berkeley, CA: University of California, Institute of Personality and Social Research.
Khodakarami, V., & Abdi, A. (2014). Project cost risk analysis: A Bayesian networks approach for modeling dependencies between cost items. *International Journal of Project Management*, *32*(7), 1233–1245.
Kim, J., Lee, A., & Ryu, H. (2013). Personality and its effects on learning performance: Design guidelines for an adaptive e-learning system based on a user model. *International Journal of Industrial Ergonomics*, *43*(5), 450–461.
Kolb, D. A. (1984). *Experiential learning: Experience as the source of learning and development*. Englewood Cliffs, NJ: Prentice-Hall.
Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159–174.
Laney, M. O. (2002). *The introvert advantage: How to thrive in an extrovert world*. New York, NY: Workman Publishing.
Mairesse, F., Walker, M. A., Mehl, M. R., & Moore, R. K. (2007). Using linguistic cues for the automatic recognition of personality in conversation and text. *Journal of Artificial Intelligence Research*, *30*, 457–500.
Manning, C. D., Raghavan, P., & Schütze, M. (2008). *Introduction to information retrieval* (pp. 260). London: Cambridge University Press.
McCrae, R. R., & John, O. P. (1992). An introduction to the five-factor model and its applications. *Journal of Personality*, *60* (2), 175–215.
National Academy of engineering. (2014). Retrieved from http://www.engineeringchallenges.org/File.aspx?id=11574&v=ba24e2ed
Okdie, B. M., Guadagno, R. E., Bernieri, F. J., Geers, A. J., & Mclarney-Vesotski, A. R. (2011). Getting to know you: Face-to-face versus online interactions. *Computers in Human Behavior*, *27*, 153–159.
Olson, D. L., & Delen, D. (2008). *Advanced data mining techniques*. New York, NY: Springer Science & Business Media.
Patrick, C. (2011). Student evaluations of teaching: Effects of the Big Five personality traits, grades and the validity hypothesis. *Assessment & Evaluation in Higher Education*, *36*(2), 239–249.
Pearl, J. (1988). *Probabilistic reasoning in expert systems: Networks of plausible inference*. San Francisco: Morgan Kaufmann Publishers.
Raja, U., & Johns, G. (2004). The impact of personality on psychological contracts. *Academy of Management Journal*, *47*, 350–367.
Ramírez-Correa, P. E., Arenas-Gaitán, J., & Rondán-Cataluña, F. J. (2015). Gender and acceptance of e-learning: A multi-group analysis based on a structural equation model among college students in Chile and Spain. *PloS one*, *10*(10), e0140460.
Scherer, K. R., & Giles, H. (Eds.). (1979). *Social markers in speech*. London: Cambridge University Press.
Schuerger, J. M., & Kuna, D. L. (1987). Adolescent personality and school performance: A follow-up study. *Psychology in the Schools*, *24*, 281–285.
Siemens, G., & Long, P. (2011). Penetrating the fog: Analytics in learning and education. *EDUCAUSE Review*, *46*(5), 30.
Tlili, A., Essalmi, F., Jemni, M., Chang, M., & Kinshuk. (2018). iMoodle: An intelligent moodle based on learning analytics. *Intelligent Tutoring Systems*, *2018*, 476–479.
Tlili, A., Essalmi, F., Jemni, M., & Chen, N. S. (2018). A complete validated learning analytics framework: Designing issues from data preparation perspective. *International Journal of Information and Communication Technology Education*, *14* (2), 1–16.
Tlili, A., Essalmi, F., Jemni, M., Kinshuk, & Chen, N. S. (2016). Role of personality in computer based learning. *Computers in Human Behavior*, *64*, 805–813.
Veletsianos, G., & Shepherdson, P. (2016). A systematic analysis and synthesis of the empirical MOOC literature published in 2013–2015. *The International Review of Research in Open and Distributed Learning*, *17*(2), 198–221.
Vinciarelli, A., & Mohammadi, G. (2014). A survey of personality computing. *IEEE Transactions on Affective Computing*, *5*(3), 273–291.
Vozniuk, A., Govaerts, S., & Gillet, D. (2013). *Towards portable learning analytics dashboards*. IEEE 13th International Conference on Advanced Learning Technologies (ICALT) (pp. 412–416).
Walsh, B. (2012). The upside of being an introvert (and why extroverts are overrated). *Time Magazine*.

Wankat, P., & Oreovicz, F. (2004). Assessing one and all. *A SEE Prism*, *7*, 49.
Watson, D., & Clark, L. A. (1997). Extraversion and its positive emotional core. In *Handbook of personality psychology* (pp. 767–793). New York, NY: Academic Press.
Yu, C. H. (2001). *An introduction to computing and interpreting Cronbach coefficient alpha in SAS*. Proceedings of 26th SAS User Group International Conference, Cary, NC, SAS Institute.
Zafar, S., & Meenakshi, K. (2012). Review of PhD theses on the role of personality in SLA. *English Language and Literature Studies*, *2*(1), 94.