



E-learners' personality identifying using their network behaviors



Fatemeh Ghorbani, Gholam Ali Montazer*

Information Technology Department, School of Engineering, Tarbiat Modares University, P.O. Box: 14115-179, Tehran, Iran

ARTICLE INFO

Article history:

Keywords:

E-learner
Personality model
Network behaviors
Fuzzy set theory
ALPIS

ABSTRACT

Automatic identification of learners attributes is the main trend in development of e-learning environments and providing personalized learning path. Many learners' characteristics may be considered to provide personalized learning that fit into individual's needs. Understanding individual differences in learning is an important issue in educational psychology. Many studies attest that learning process can be more effective by considering personality traits. In this paper, we propose a fuzzy inference system, ALPIS, to identify the Big-Five Personality from learner's network behaviors. The proposed system had been implemented through MATLAB fuzzy tools. We conducted an experimental study in real e-learning environment with 53 students (29 male and 24 female) to evaluate the ALPIS performance. Comparing to the conventional inventory-based psychological analysis (NEO-FFI) we demonstrate that learners' personality can be predicted with reasonable precision based on their online behaviors.

Experiments have shown that ALPIS is able to automatically detect learners' personality with accuracy at least 78% on medium Agreeableness category and 97% on low extroversion category. Because of using learners' behaviors ALPIS provides a reasonable precision in personality detection.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

Many e-learning systems have been developed to adapt learning content to an individual's attributes. Personality is a learner attribute that affects on learning achievement. Providing more dedicated and personalized training is a desired property for e-learning systems. However, little researches have been done to conduct proper analysis for explaining the learner's network behaviors from their personality. Many educational psychologists have claimed that personality influences the way learners actively involvement in their learning processes as well as whether they take responsibility for self-direction and discipline (Kim, Lee, & Ryu, 2013). The relationship between personality and academic performance has been examined in the context of distance and online education (Butler & Pinto-Zipp, 2006; Irani, Telg, Scherler, & Harrington, 2003; Kanuka & Nocente, 2003; Kim & Schniederjans, 2004; Lee & Lee, 2006). Although results have been mixed, but generally support a significant relationship between personality and performance. Kim and Schniederjans (2004) found a significant relationship between performance and each of the Big Five dimensions except extroversion; Lee and Lee (2006) found a relationship

between personality and interaction in web-based threaded discussion; and Butler and Pinto-Zipp (2006) found several significant relationships which were consistent with personality traits. Pavalache-Ilie and Cocorada (2014) investigated the relations between some of the student's personality characteristics and some dimensions of online learning (preferred learning methods, student's needs, relations with the teacher and classmates). The obtained data suggest that the introverts prefer to learn and contribute to an online discussion forum while the extroverts would rather take part in the face to face classroom environment. In Kim et al. (2013), the relation between a learner's personality traits has been tested and the effects of these traits on learning preferences have been analyzed. Also, Feldman, Smart, and Ethington's (1999) noted that achievement is a function of the fit between personality type and environment. Vermunt, Bronkhorst, and Martín ez-Fernández (2014) emphasizes individual differences in the use of learning strategies in higher education. The relationship between academic performance and personality traits has been described by Ackerman and his co-worker (Ackerman, 1999; Ackerman & Beier, 2002; Ackerman & Heggstad, 1997; Goff & Ackerman, 1992).

According to above brief discussion, identifying learners' personality will help the designer create customized educational materials tailored to each individual. The system's information about learners' characteristics has implications on selecting the teaching strategies. Knowing the personality type of each learner

* Corresponding author at: Information Technology Department, School of Engineering, Tarbiat Modares University, P.O. Box: 14115-179, Nasr Bridge, Jalal Al Ahmad Highway, Tehran, Iran. Tel.: +98 (21)82883990.

E-mail address: montazer@modares.ac.ir (G.A. Montazer).

can help to identify their learning preferences and strengths, which can be utilized in instructional designs and making the learning process more effective and improving students' performances.

Psychometric instruments, such as NEO-FFI, are the most used instruments to identify people personality but some researches (Castillo, Gama, & e Breda, 2005) explain that the information about the learners acquired by psychometric instruments encloses some uncertainty because this results acquired from self-reports (Castillo et al., 2005).

On the other hand, peoples' values and preferences are often reflected in their personality traits. They could explain the occurrences of certain individual behaviors. Personality is hypothesized to lead to beliefs related to the behavior (Devaraj, Easley, & Crant, 2008). In the other words, analyzing outer behaviors can be used to personality analysis since behavior is the expression of personality.

Hence, identifying learners' personality based on their outer behaviors, using automatic approaches, tend to be more accurate since they analyze data derived from an interval of time, instead of data collected at a particular point in time and by self-reports (Bai, Zhu, & Cheng, 2012).

According to our knowledge, there is no research in literature to identify learners' personality based on their network behaviors in e-learning environment. But some researches have been done in social networks environment context. Gosling experimented on the description of personality in social networks (Gosling, Augustine, Vazire, Holtzman, & Gaddis, 2011). He introduced a mapping between personality and network behaviors. They used self-reported Facebook usage and observable profile information and finally gave the correlation factor between personality and online behavior. Orr et al. (2009) discussed the relation between shyness and use of social networks. They discovered that shyness was significantly positively correlated with the time spent on these networks and negatively correlated with the number of friends. Bai et al. proposed an automatic and objective personality prediction system based on user's behaviors on social networks (Bai et al., 2012).

Acquired data from learner behaviors may contain uncertainty and vague. Fuzzy sets theory is one of the most appropriate mathematical theories to deal with uncertainty. Fuzzy models can model both quantitative and qualitative information (Georgiou & Makry, 2004). In this context, we believe that fuzzy approach eliminates many uncertainties and vague for automatic identification of personality based on learner behavior. This approach uses expert knowledge about relation between learner's personality and his/her behavioral patterns. This makes the method more general and more easily reusable.

In this paper, a fuzzy system, that named ALPIS, is designed to identify learners' personality based on their behaviors. The Big five personality model (NEO personality model) that supports this approach is the personality model proposed by Costa and McCrae (1992).

The paper is organized as follows: in Section 2 the Big five personality model and NEO-FFI personality inventory are described. The designed fuzzy system is introduced in Section 3 and the experimental results are discussed in Section 4. Finally, the conclusion is given in Section 5.

2. The Big five personality model

Individual differences include some traits such as personality and demographic that account for differences attributable to circumstances such as experience and training (Agarwal & Prasad, 1999). Personality is a variable to link external variables to behaviors and lead to beliefs related to the behavior (Devaraj et al., 2008). Human personality may effects on many social activities such as

academic performance (Lounsbury, Sundstrom, Loveland, & Gibson, 2003), learning strategy preference (Conti & McNeil, 2011) and academic motivation (Komarraju & Karau, 2005). In the literature of personality and individual differences, Big Five Personality (BFP) factors is the best description of personality that referred as the "Big 5" or "NEO" including neuroticism, extroversion-introversion, openness, agreeableness, conscientiousness (Costa & McCrae, 1992; Petrides et al., 2010). Therefore this research employs the Big Five Personality factors to measure learner personality.

Neuroticism is a dimension of personality defined by stability and low anxiety at one end as opposed to instability and high anxiety at the other end. The term neuroticism has an inherent negative denotation (Costa & McCrae, 1992). Extroversion defined as a trait characterized by a keen interest in other people and external events, and venturing forth with confidence into the unknown and denotes activity, energy, and vigorousness (Zuckerman, 2005). Openness to experience is associated with creativity, intelligence, imagination, and autonomy. Openness refers to how willing people are to make adjustments in notions and activities in accordance with new ideas or situations (De Raad & Schouwenburg, 1996). Agreeableness refers to the tendency to be friendly, flexible, and cooperative. Agreeableness measures how compatible people are with other people, or basically how able they are to get along with others (Ewen, 2009). Conscientiousness refers to how much a person considers others when making decisions. Conscientiousness represents characteristics such as being thorough, in a task-oriented way, systematically and carefully (Ewen, 2009).

The NEO Five-Factor Inventory (NEO-FFI) developed by Costa and McCrae (1992), consists of 60 items, 12 for the factors of Conscientiousness, Neuroticism, Extroversion, Agreeableness and Openness to Experience. It was the most widely used measure of personality traits and has demonstrated good internal and external validity (Costa & McCrae, 1992). The NEO-FFI items rated on a 5-point scale that provides a quick, reliable, and accurate measure of the five domains of personality. The score of each item is 1 for strongly disagree, 2 for disagree, 3 neither agree nor disagree, 4 agree and 5 for strongly agree. A number of the items such as 1, 3, 8 and 12 are reversely scored. The score of each dimension is in [12,60] interval. Scores between 12 and 24 are low, between 24 and 48 are normal and between 48 and 60 are high on each dimension. For example, people who score high on Conscientiousness tend to be conscientious and well-organized; have high standards and always strive to achieve their goals.

3. Automatic learner's personality identifier system (ALPIS)

A common methodology in behavioral science is to use self-report questionnaires to gather data about people such as using NEO-FFI to measure personality. Most of these instruments depend on the use of self-report because it offers an efficient way to collect information. These instruments are a cheap and quick way, in terms of time and cost, to obtain data and can be easily implemented (Bersoff & Bersoff, 2000). However some disadvantages are associated with the use of self-report instruments. For example, they affect by social desirability biases. People tend to present themselves in a favorable manner, especially when they make judgments about attitudes that are negatively valued such as irresponsibility. Some people often believe that problems that they encounter, especially interpersonal conflicts, are not due to any fault of their own. This lack of insight that is characteristic of individuals with personality disorders makes it even more difficult to obtain accurate personality information using self-report measures (Oltmanns & Turkheimer, 2006).

In order to obtain a more comprehensive perspective on an individual's interpersonal difficulties, several researchers have

turned towards gathering additional information from other informants (Klein, 2003; McDermut & Zimmerman, 2005; Westen, 1997).

In fact there are many reasons why questionnaire may not be entirely valid, such as (Hoskin, 2012):

- **Honesty:** self-report questionnaires are relying on the honesty of their participants. The degree of honesty varies with the topic of questions, for example participants are less honest about measures relating to their conscientiousness. The NEO-FFI results in our research shows that many learners think they are dutiful and do their best to deliver their assignments whereas they really do their duties with delay or not to deliver their assignments.
- **Introspective ability:** if we suppose participants are hones, they may lack the introspective ability to answer an accurate response. People appear to view themselves in a completely different light to how others see them. Therefore any self-report information may be incorrect despite the participant's trying to be honest and accurate.
- **Understanding** – the participants understanding may also vary-ing regarding their interpretation of particular questions. This is a very big problem when measuring more abstract concepts such as personality. It would be impossible to ensure that everyone who completed the questionnaire interpreted that question in the same way.
- **Most self-report measures in behavioral science** relate to con-structs, such a personality measures, that cannot be measured in interval units and are therefore always ordinal. The problem with ordinal data is using parametric statistical techniques, because these tests make assumptions about the distribution of the data that cannot be met when said data is ordinal.

Although there are many problems with using self-report ques-tionnaires they will continue to be a popular methodology in behavioral science because of their utility. Furthermore some of the problems listed above can be countered through the careful design and using other information sources. In this paper, we use network behaviors as another source to identify personality. Analyzing outer behaviors is the principle of personality analysis since behavior is the manifestation of personality (Bai et al., 2012). We propose an automatic identification system based on online learner's behaviors on Learning Management System (LMS). Learner's behaviors can be influenced by a wide variety of variables such as cultural environment and their measuring includes some uncertainty, too. For example when learner login to system and spend some times in it, we do not conclude that how much time he/she spends in course. Therefore measuring the learner's behaviors includes some vague and uncertainty. Therefore, the description of e-learner's personality based on their network's behaviors includes a great deal of uncertainty.

Fuzzy sets theory is a framework to handle vague and uncertain data. It has the capability of handling the input that is approximate rather than fixed and exact. Fuzzy logic can be considered as the simple way to arrive at a definite output based on vague, ambigu-ous, imprecise or noisy input information. In fuzzy system, “if-then rules” is used in order to inference output variables based on fuzzy input variables. Using fuzzy legal and linguistic variables in fuzzy system decision making is possible when data is incomplete and uncertain. Each fuzzy system contains four major components (Huang & Chang, 2007):

1. **Fuzzification:** using linguistic variables each interval of vari-able's values is corresponded to a linguistic variable. The rela-tionships between input variables and linguistic variables are defined by using membership functions (MFs).

2. **Knowledge Base:** included fuzzy rules i.e. application domain expert knowledge using linguistic variables.
3. **Decision Making Logic,** i.e., the inference engine has the capacity of inferring rules using fuzzy rules and fuzzy implication, and so simulates the human decision-making ability.
4. **Defuzzification,** i.e. the reverse process of fuzzificaiton. It pro-duces a crisp output value from the fuzzy outputs of the infer-ence engine.

For more details see (Wang, 1997, chap. 2–3).

In our work, we try to build a fuzzy system based on learner's online behaviors. To this end, we will answer the following questions:

1. What are the related behaviors to personality dimensions (input variables)?
2. What are the relations between network behaviors and person-ality dimensions (fuzzy rules)?

In the next subsections, we answer these questions.

3.1. Input and output variables of ALPIS

As the first step to design fuzzy system for automatic learner personality identification, some learner's observable behaviors should be related to personality dimensions. We use NEO-FFI ques-tions and relate its questions to some learner's network behaviors. These behaviors have been shown in Table 1. For example, the cor-responding items to extroversion dimension in NEO-FFI show that if a person communicate with others easily, has many friends, like to speak with others and is an active person then he/she will be more extroversion and his/her score on this dimension will be high. So, we infer participation in chat rooms and forums, number of friends and adding post in forums as extroversion indicators. The same analysis has been done for other dimension and the behavioral fac-tors for personality dimensions have been selected as shown in Table 1. The corresponded behaviors to personality dimensions have been validated by educational psychology experts.

These behaviors should be expressed as linguistic variables. Linguistic variables correspond to a range of values of the network behaviors. Based on logged data from the underlying LMS, we find the minimum and maximum of each behavior. Using these data

Table 1
The related network behaviors to personality dimensions.

Personality dimension	Mapped network behaviors
Neuroticism	Number of friends (NF) Number of entrance to system in a week (NES) Delay in assignment delivering (DAD)
Extroversion	Number of friends (NF) Participation in chat rooms and forums (PICF) Adding posts in forums (APF)
Openness	Number of friends (NF) Difficulty level of exercises and examples (DLAE) Dedicating time for reading concepts and theories (DTCM) Dedicating time for reading examples and cases (DTEC)
Agreeableness	Number of group exercises (NGE) Participation in chat rooms and forum (PICF) Participation in troubleshooting groups (PTG)
Conscientiousness	Delay in assignment delivering (DAD) Number of entrance to system in a week (NES) Scores (SR) View of lesson news (VLN)

and lesson's teacher assistances opinion we define three linguistic variables named "low", "medium", and "high" for each behavior. For example, teacher assistances expressed that if students log in to system less than 3 times in a week then the Number of entrance to system in a week is low, between 3 and 5 times is medium and more than 6 is high. We use their opinion to define corresponded network behaviors as linguistic variables. Table 2 shows the definition of these variables using trapezoidal fuzzy numbers. Trapezoidal numbers are shown using $A = (a, b, c, d)$ and define as Eq. (1).

$$\mu_A(x) = \begin{cases} \frac{x-a}{b-a} & a \leq x \leq b \\ 1 & b \leq x \leq c \\ \frac{c-d}{c-d} & c \leq x \leq d \\ 0 & o.w. \end{cases} \quad (1)$$

Fig. 1(a) shows "Participation in chat rooms and forums" as a linguistic variable which can take values in the interval [0, 15].

The output variables of system are the learner's score on different personality dimensions. These variables can take values in the interval [12, 60] as expressed in Section 2. In order to more accurate inference, the output linguistic variables are defined as "very low", "low", "medium", "high" and "very high" as shown in Fig. 1(b).

3.2. Fuzzy rules of ALPIS

It is an important issue to relate learner's personality to his/her behavioral patterns, to identify his/her personality. This issue can be solved using fuzzy rules and the relations between learner behaviors (input variables) and score of personality dimensions (output variables). These rules are extracted based on NEO-FFI analysis and using experts' knowledge that was gathered in interview. For example, NEO-FFI questions show that if a person has

many friends, communicates with others easily, be an active person and likes to speak with others then his/her score on extroversion dimension is high. According to related network behaviors to extroversion dimension, we can say if learner has many friends, participates in chat rooms and forums frequently and adds many posts then his/her score on extroversion dimension is high. Using linguistic variables, this analysis can be expressed as following fuzzy rule:

"IF participation in forums (PICF) is high AND number of friends (NF) is high AND adding post in forums (APF) is high, THEN extroversion is high". The same analysis has been done for other dimensions and fuzzy rules have been extracted. Table 3 shows some of the rules of the system.

3.3. System design

According above discussion, our system comes from a combination of expert systems from Computer Science and Big Five Inventory from Psychology. The system architecture has been shown in Fig. 2.

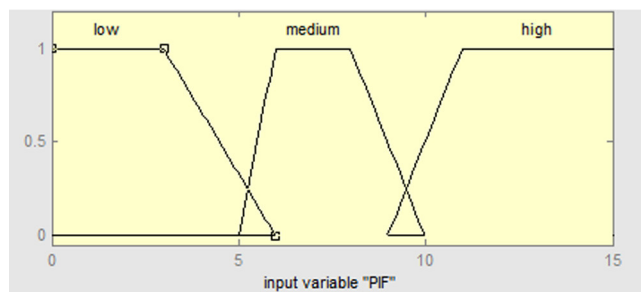
In first log in, learner finishes NEO-FFI and the results of this inventory is saved in his/her profile using a five dimension vector. During the course, we use LMS log files to gather behaviors data and then process these data to extract the behavior indicators that related to personality dimensions. These indicators are present by behavior features vectors. ALPIS use these vectors to predict learner personality and finally ALPIS outputs compare with NEO-FFI and learner profile is updated.

4. Experimental results

The proposed fuzzy system was evaluated in a real e-learning environment. We considered 53 learners in a graduate online

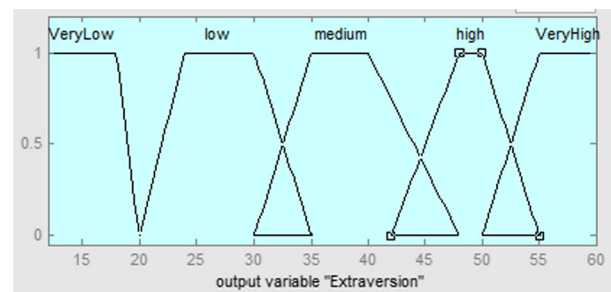
Table 2
The definition of linguistic variables as fuzzy trapezoidal numbers.

Behaviors	Linguistic variables		
	Low	Medium	High
Number of friends (NF)	(0,2,5)	(4,5,6,8)	(7,9,20)
Scores (SR)	(0,10,14)	(13,15,16,18)	(17,18,20)
Participation in chat and forums (PICF)	(0,3,6)	(5,6,8,10)	(9,11,15)
Adding post in forums (APF)	(0,1,2)	(1,3,7,10)	(8,10,15)
Difficulty level of exercises and examples (DLAE)	Low	Medium	High
Dedicating time for reading concepts and theories (DTCM)	(0,60,80)	(70,80,90,110)	(90,110,250)
Dedicating time for reading examples and exercises (DTE)	(0,3,5)	(3,6,8,10)	(9,12,30)
Participation in Troubleshooting groups (PTG)	(0,3,5)	(3,6,8,9)	(8,10,15)
Delay in assignment delivering (days) (DAD)	(0,1,3)	(2,3,5,6)	(5,7,10)
Number of entrance to system in a week (NES)	(0,2,3)	(2,3,4,5)	(4,5,7)
View of lesson news (VLN)	(0,5,7)	(6,9,12,16)	(14,16,25)



"Participation in chat rooms and forums" as a linguistic variable

(a)



Personality dimensions as a linguistic variable

(b)

Fig. 1. Linguistic variables of system.

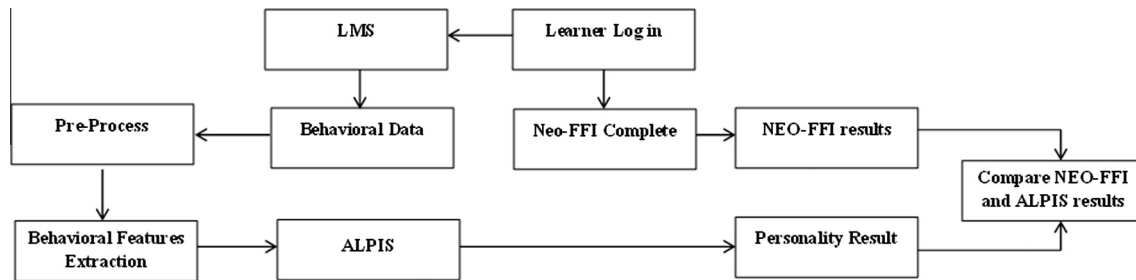


Fig. 2. ALPIS architecture.

Table 3
Some extracted fuzzy rules.

Dimension	Extracted rules
Neuroticism	If NF is low AND NES is low AND DAD is high THEN neuroticism is very high If NF is high AND NES is high AND DAD is low THEN neuroticism is very low If NF is medium AND NES is high AND DAD is medium THEN neuroticism is medium If NF is low AND NES is medium AND DAD is low THEN neuroticism is low
Extroversion	If NF is high AND PICF is high AND APF is high THEN extroversion is high If NF is medium AND PICF is low AND APF is low THEN extroversion is low If NF is medium AND PICF is medium AND APF is high THEN extroversion is medium If NF is low AND PICF is low AND APF is low THEN extroversion is very low
Openness	If NF is high AND LAE is high AND DTCM is medium OR DTEC is high THEN openness is very high If NF is low AND LAE is low AND DTCM is medium AND DTEC is low THEN openness is very low If NF is medium AND LAE is low AND DTCM is medium OR DTEC is medium THEN openness is medium If NF is low AND LAE is medium AND DTCM is low OR DTEC is low THEN openness is low
Agreeableness	If NGE is low AND PICF is low AND PTG is low THEN agreeableness is very low If NGE is medium AND PICF low AND PTG is low THEN agreeableness is low If NGE is medium AND PICF medium AND PTG is low THEN agreeableness is medium If NGE is high AND PICF high AND PTG is high THEN agreeableness is very high
Conscientiousness	If DAD is high AND NES is low AND SR is low AND VLN is low THEN conscientiousness is very low If DAD is high AND NES is low AND SR is medium AND VLN is medium THEN conscientiousness is low If DAD is medium AND NES is low AND SR is medium AND VLN is medium THEN conscientiousness is medium If DAD is low AND NES is high AND SR is medium AND VLN is high THEN conscientiousness is very high

course. We used logged files from Tarbiat Modares University's LMS (moodle2).

4.1. Data collection

At the beginning of course, learners completed NEO-FFI through Survey Monkey website. In order to keep the quality of samples, we asked teacher assistance to encourage learners about finishing the NEO-FFI. We collected the behavior data of 8 sessions in one month. Table 4 shows some information about learners and course. According to collected data, we did not have the Number of group exercises and it was removed from associated behaviors to Agreeableness dimension.

Table 4
Learners and course attributes.

	Attribute	Value
Learners	Total number of learner	95 (51 male and 44 female)
	Age average	24.37
	Age variance	2.85
Course	Sessions	8
	Duration	1 month

4.2. Data analysis

After collecting the behavior data, 4 learners removed from our experiment because they did not participate actively in course. The data have been normalized between [0,1]. Table 5 shows the Descriptive Statistics on data.

Data normality has been checked using Kolmogorov–Smirnov test. Table 6 shows the results of Kolmogorov–Smirnov test. As shown in Tables 5 and 6, Adding post in forums (APF) and View of lesson news (VLN) have not normal distribution.

Learners complete the NEO-FFI to identify their personality scores. The scores distribution has been shown in Fig. 3.

Table 7 shows the descriptive statistics of personality scores.

The normality of scores on each dimension has been checked using Kolmogorov–Smirnov and Shapiro–Wilcoxon test. According to these results, personality scores have normal distribution in conscientiousness dimension at 0.05 significant level. Neuroticism, extroversion, openness and agreeableness have not normal distribution. Table 8 shows these results.

Because of abnormal distribution of some behaviors and personality dimension scores, we use Spearman correlation coefficient to evaluate the relation between behaviors and personality dimensions. Table 9 shows the correlation between network behaviors and personality dimensions using Spearman correlation coefficient. As shown in this table there is a reasonable correlation between personality dimensions and corresponded behaviors.

4.3. ALPIS implementation and evaluation

We implemented ALPIS system using MATLAB 2010 fuzzy tool and used Mamdani type of fuzzy system with following configuration that set by try and error:

- AND method: prod.
- OR method: max.
- Defuzzification method: LOM (Large Of Maximum).

The proposed system includes different rules and input variables (different behaviors) for each dimension therefore we implemented it in 5 separated fuzzy systems.

To evaluate the performance of fuzzy system, flexibility, completeness and consistency are essential. These criterions influence

Table 5

The descriptive statistics of behavior data.

Behaviors	Mean	Std. deviation	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Std. error	Statistic	Std. error
Number of friends (NF)	0.54	0.28	−0.221	0.285	−0.726	0.511
Scores (SR)	17.27	6.80	0.673	0.258	−0.675	0.511
Participation in chat and forums (PICF)	0.61	0.25	−0.476	0.258	−0.539	0.511
Adding post in forums (APF)	0.55	0.24	−0.302	0.258	−0.456	0.511
Difficulty level of exercises and examples (DLAE)	0.39	0.27	0.673	0.258	−0.675	0.511
Dedicating time for reading course materials (DTCM)	0.51	0.22	0.707	0.258	1.794	0.511
Dedicating time for reading examples and exercises (DTE)	0.59	0.27	−0.367	0.258	−0.992	0.511
Participation in Troubleshooting groups (PTG)	0.42	0.25	0.530	0.258	−0.419	0.511
Delay in assignment delivering (days)	0.23	0.24	0.975	0.258	0.282	0.511
Number of entrance to system in a week (NES)	0.6	0.24	−0.124	0.258	−0.311	0.511
View of lesson news (VLN)	0.49	0.24	0.429	0.258	−0.371	0.511

Table 6

The Kolmogorov–Smirnov test results on behavior data.

Behavior	Kolmogorov–Smirnov			Shapiro–Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Number of friends (NF)	.079	95	.167	.963	95	.009
Scores (SR)	.130	95	.200 ^A	.893	27	.009
Participation in chat and forums (PICF)	.102	95	.016	.944	95	.001
Adding post in forums (APF)	.069	95	.200 ^A	.983	95	.243 ^A
Difficulty level of exercises and examples (DLAE)	.154	95	.000	.906	95	.000
Dedicating time for reading course materials (DTCM)	.124	95	.001	.955	95	.003
Dedicating time for reading examples and exercises (DTE)	.087	95	.074	.948	95	.001
Participation in Troubleshooting groups (PTG)	.107	95	.010	.966	95	.015
Delay in assignment delivering (days)	.167	95	.000	.843	95	.000
Number of entrance to system in a week (NES)	.066	95	.200 ^A	.971	95	.035
View of lesson news (VLN)	.059	95	.200 ^A	.981	95	.198 ^A

^A Variable distribution is not normal.

the suitability of a fuzzy system. Fuzzy rules are regarded as inconsistent, if

- They conflict with the expert knowledge or heuristics.
- They have very similar premise parts, but possess rather different consequents.

To identify confliction between the results of ALPIS with expert knowledge, its results have been compared with the obtained results from the NEO-FFI filled out by the learners. NEO-FFI categorizes learners in 3 classes based on their scores as follow:

- Low: scores between 12 and 24;
- Medium: scores between 24 and 48;
- High: scores between 48 and 60.

We categorized learners in same classes based on ALPIS scores and then compared NEO-FFI and ALPIS results to evaluate the ALPIS accuracy in personality prediction. In this point of view, personality prediction can be seen as a classification problem that classifies learners in 3 classes (low, medium and high). Precision, recall, *F*-value and accuracy are some useful measures to evaluate the classifier efficacy. In a classification task, the terms true positives, true negatives, false positives and false negatives compare the results of the classifier under test with trusted external judgments. The number of true positives (TP) is the number of items correctly labeled as belonging to the class and false positive (FP) is the number of items incorrectly labeled as belonging to the class. False negative (FN) are items which were not labeled as belonging to the class but should have been. True negative (TN) are items that were not labeled as belonging to the class and should not have been. The precision, recall, *F*-value and accuracy are defined as (Olson & Delen, 2008):

$$\begin{aligned}
 \text{Precision} &= \frac{TP}{TP + FP} \\
 \text{Recall} &= \frac{TP}{TP + FN} \\
 F_Value &= 2 * \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}} \\
 \text{Accuracy} &= \frac{TN + TP}{\text{total number of items}}
 \end{aligned} \tag{2}$$

Using 10-fold cross validation, the results (precision, recall, *F*-value and accuracy) of three-class classification problem for five personality dimensions are shown in Table 10. As shown in Table 10, the ALPIS accuracy is at least 0.78 in learner's personality identification.

To evaluate ALPIS performance from statically view, we compare the difference between ALPIS results and NEO-FFI ones. Since some personality scores have not normal distribution, the Wilcoxon test has been used to evaluate the difference between ALPIS results and NEO-FFI results. Table 11 shows these results. According to these results, there is significant difference between ALPIS scores and NEO-FFI scores in Neuroticism and Conscientiousness dimensions.

The main aim in personality identification is personality category identification; hence we use Chi-square test as another evaluation criterion to compare between achieved personality category from ALPIS results and that achieved from NEO-FFI results. Table 12 shows this comparison. With 2 degrees of freedom and alpha set at 0.05 the tabled value is 5.99. As shown in Table 12, all calculated chi-squared values are smaller than tabled value, expect in Conscientiousness dimension. Hence, there is no significant difference between ALPIS identified categories and NEO-FFI categories in neuroticism, extroversion, openness and agreeableness.

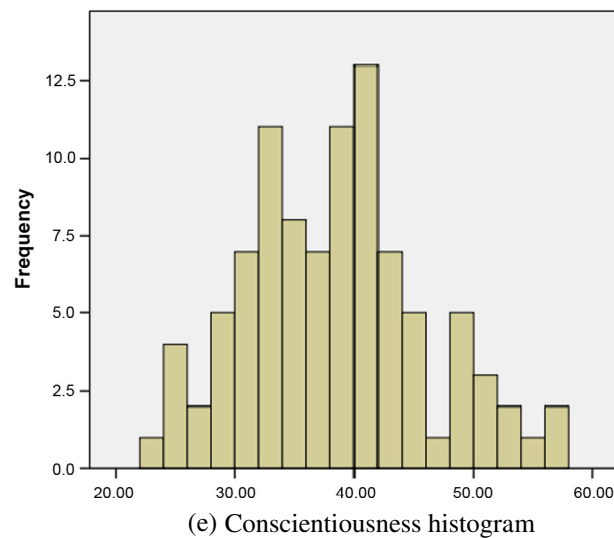
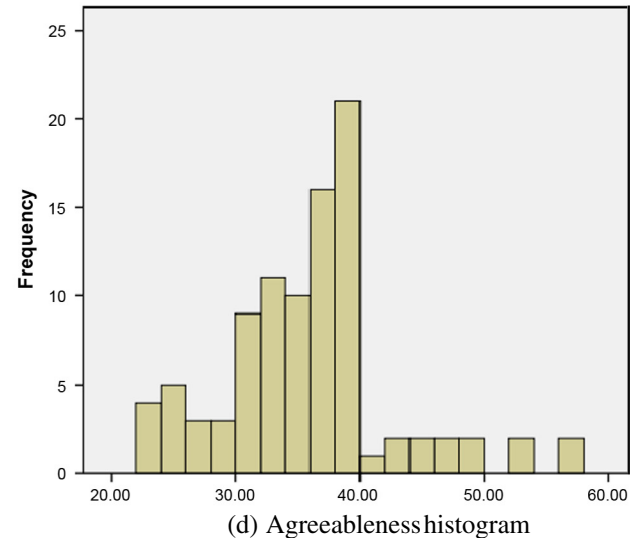
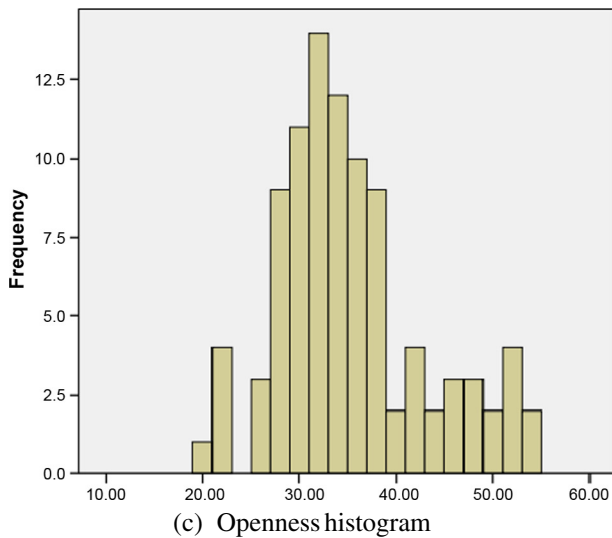
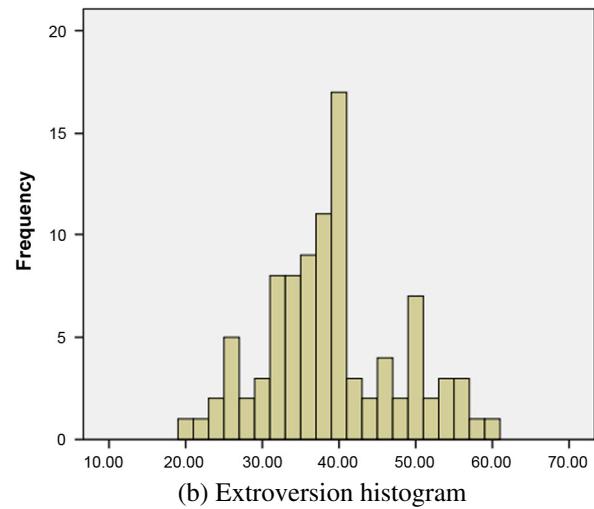
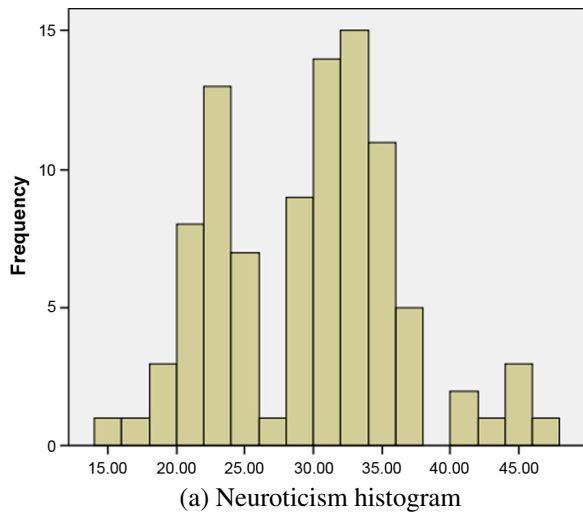


Fig. 3. Personality score distribution.

On the other hand, the discussion of consistency is sensible only if the premise parts of the rules are very similar, if not necessarily the same. Therefore, it is necessary to provide a proper definition of consistency. We use the definition of

the consistency that provided by Jin, Seelen, and Sendhoff (1999). They provided the definition of the similarity of rule premise (SRP) and the similarity of rule consequent (SRC) as follow:

Table 7

The descriptive statistics of personality scores.

Personality dimension	Minimum	Maximum	Mean	Std. deviation	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Statistic	Std. error	Statistic	Std. error
Neuroticism	15.00	49.00	29.27	6.73	.196	.247	−.197	.490
Extroversion	20.00	59.00	38.76	8.62	.323	.247	−.222	.490
Openness	20.00	54.00	35.01	7.80	.693	.247	.101	.490
Agreeableness	23.00	56.00	35.23	6.84	.611	.247	1.181	.490
Conscientiousness	23.00	57.00	37.82	7.59	.329	.247	−.184	.490

Table 8

The Kolmogorov–Smirnov test results on personality dimensions.

	Kolmogorov–Smirnov			Shapiro–Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Neuroticism	.098	95	.025	.971	95	.034
Extroversion	.148	95	.000	.972	95	.042
Openness	.132	95	.000	.940	95	.000
Agreeableness	.154	95	.000	.944	95	.000
Conscientiousness	.080	95	.153	.982	95	.215

Consider two rules in the rule base

 R_i : If x_1 is $A_{i1}(x_1)$ and x_2 is $A_{i2}(x_2)$ and ... x_n is $A_{in}(x_n)$, then y is $B_i(y)$ R_k : If x_1 is $A_{k1}(x_1)$ and x_2 is $A_{k2}(x_2)$ and ... x_n is $A_{kn}(x_n)$, then y is $B_k(y)$

then SRP and SRC of these two rules are defined as follow:

$$SRP(i, k) = \min_{j=1, \dots, n} S(A_{ij}, A_{kj}) \quad (3)$$

$$SRC(i, k) = S(B_i, B_k) \quad (4)$$

where n is the total number of input variables and $S(A, B)$ is the fuzzy similarity measure of fuzzy sets A and B . Then the consistency of rule R_i and R_k is defined by:

$$Cons(R_i, R_k) = \exp \left\{ - \frac{\left(\frac{SRP(i, k)}{SRC(i, k)} - 1 \right)^2}{\left(\frac{1}{SRP(i, k)} \right)^2} \right\} \quad (5)$$

Jin has pointed that the consistency definition is generally suitable for Mamdani fuzzy systems. For more details see Jin et al. (1999). We extend this definition to measure the consistency of a fuzzy system included m rules and define three consistency measures as follow:

Table 9

The spearman correlation coefficient.

Personality dimension	Mapped network behaviors	Spearman correlation	Sig. (2-tailed)
Neuroticism	Number of friends (NF)	−0.726**	0.000
	Number of entrance to system in a week (NES)	−0.774**	0.000
	Delay in assignment delivering (DAD)	−0.571**	0.000
Extroversion	Number of friends (NF)	0.786**	0.000
	Participation in chat rooms and forums (PICF)	0.889**	0.000
	Adding posts in forums (APF)	0.765**	0.000
Openness	Number of friends (NF)	0.314**	0.002
	Difficulty level of exercises and examples (DLAE)	0.805**	0.000
	Dedicating time for reading concepts and theories (DTCM)	0.502**	0.000
	Dedicating time for reading examples and cases (DTEC)	0.621**	0.000
Agreeableness	Participation in chat rooms and forum (PICF)	0.817**	0.000
	Participation in troubleshooting groups (PTG)	0.671**	0.000
Conscientiousness	Delay in assignment delivering (DAD)	−0.521**	0.000
	Number of entrance to system in a week (NES)	0.504**	0.000
	Scores (SR)	0.751**	0.000
	View of lesson news (VLN)	.856**	0.000

** Correlation is significant at the 0.01 level (2-tailed).

Table 10

ALPIS evaluation results.

Dimensions	Classes	Precision	Recall	F-value	Accuracy
Neuroticism	Low	0.63	0.875	0.73	0.89
	Medium	0.94	0.77	0.84	0.85
	High	0.17	0.5	0.253	0.90
Extroversion	Low	0.83	1	0.91	0.97
	Medium	0.94	0.89	0.91	0.88
	High	0.67	0.75	0.71	0.88
Openness	Low	0.83	0.83	0.83	0.96
	Medium	0.88	0.91	0.89	0.86
	High	0.82	0.75	0.78	0.90
Agreeableness	Low	0.4	1	0.57	0.93
	Medium	0.86	0.84	0.85	0.78
	High	0.5	0.38	0.43	0.84
Conscientiousness	Low	0.5	1	0.67	0.94
	Medium	0.82	0.90	0.86	0.81
	High	0.9	0.56	0.69	0.84

$$Cons1 = \min_{i=1, \dots, m} \min_{k \neq i} Cons(R_i, R_k) \quad (6)$$

$$Cons2 = \frac{2}{m(m-1)} \sum_{i=1}^{m-1} \sum_{k=i+1}^m Cons(R_i, R_k) \quad (7)$$

$$Cons3 = \max_{i=1, \dots, m} \min_{k \neq i} Cons(R_i, R_k) \quad (8)$$

where m is the total number of rules of system and $Cons(R_i, R_k)$ is defined by Eq. (3). Table 13 shows the consistency of ALPIS by applying these criterions.

We notice from Table 11 that the total consistency is at least 0.8 for neuroticism dimension, 0.91 for extroversion, 0.90 for openness

Table 11

The Wilcoxon test results.

		N	Mean rank	Sum of ranks	Z	Asymp. sig. (2-tailed)
Neuroticism	Negative ranks	31 ^a	45.81	1420.00	−2.528 ^d	0.011 ^e
	Positive ranks	59 ^b	45.34	2675.00		
	Ties	5 ^c				
Extroversion	Negative ranks	35 ^a	45.80	1603.00	−1.481 ^d	0.139
	Positive ranks	53 ^b	43.64	2313.00		
	Ties	7 ^c				
Openness	Negative ranks	43 ^a	40.49	1741.00	−0.559 ^d	0.576
	Positive ranks	43 ^b	46.51	2000.00		
	Ties	9 ^c				
Agreeableness	Negative ranks	39 ^a	35.32	1377.50	−1.666 ^d	0.096
	Positive ranks	44 ^b	47.92	2108.50		
	Ties	12 ^c				
Conscientiousness	Negative ranks	62 ^a	51.98	3222.50	−5.549 ^d	0.000 ^e
	Positive ranks	25 ^b	24.22	605.50		
	Ties	8 ^c				

^a ALPIS score < NEO-FFI scorera van1.^b ALPIS score > NEO-FFI score.^c ALPIS score = NEO-FFI score.^d Based on negative ranks.^e p-Value < 0.05 (significant difference).**Table 12**

Chi-square test.

Dimensions	Classes	Frequency in NEO-FFI	Frequency in ALPIS	Chi-square value
Neuroticism	Low	27	24	4.8
	Medium	66	66	
	High	2	5	
Extroversion	Low	4	5	0.83
	Medium	74	76	
	High	17	14	
Openness	Low	5	6	1.52
	Medium	82	78	
	High	8	11	
Agreeableness	Low	8	10	2.31
	Medium	81	76	
	High	6	9	
Conscientiousness	Low	4	10	10.01
	Medium	81	78	
	High	10	7	

Table 13

Consistency results of ALPIS for personality dimensions.

Dimension	Number of rules	Cons 1	Cons 2	Cons 3
Neuroticism	27	0.8	0.84	0.86
Extroversion	27	0.91	0.93	0.95
Openness	81	0.90	0.93	0.94
Agreeableness	27	0.87	0.89	0.91
Conscientiousness	81	0.89	0.91	0.92

dimension, 0.87 for agreeableness and 0.89 for conscientiousness dimension that means the consistency of ALPIS is being in meaningful range.

5. Discussion

From the above results, it is seen that ALPIS can identify the learner's personality in acceptable accuracy. As shown in Table 10, recall, precision, *F*-value and accuracy are in reasonable range except in some cases that explain following.

Where neuroticism score is high, precision, recall and *F*-value are low but accuracy is high. There are few people in our experiment,

and in nature world generally, with high neuroticism scores. ALPIS can predict high neuroticism category, so its accuracy is high on this category (0.9) but since some learners with medium scores classify in high category according to their low participation in course, precision and recall are low for high neuroticism category.

In extroversion dimension, ALPIS identifies low category exactly (*recall* = 1) but some learners with medium scores classify in high category because of their active participation in course. It causes precision in high extroversion category is lower than low and medium categories.

In openness dimension, all measures are in acceptable range. In the other words, ALPIS accuracy prediction is reasonable in openness dimension.

Low agreeableness has been predicted exactly using ALPIS and therefore *recall* = 1 for this category but some learners with medium agreeableness scores classify in low category and causes fewer precision measures in this category.

Finally in conscientiousness dimension, ALPIS predicts low scores exactly (*recall* = 1) while some learners having medium scores in this dimension (based on NEO-FFI results) classify in low category based on ALPIS results and according to their behaviors. So, precision in low conscientiousness category is lower than medium and high categories. On the other hand, some learners classify in high conscientiousness category based on NEO-FFI results while using ALPIS results their conscientiousness scores are medium, hence the recall measure in high conscientiousness category is lower than medium and low scores.

Using Wilcoxon test there is significant difference between NEO-FFI personality scores and ALPIS ones in neuroticism and conscientiousness dimensions and there is no significant difference between NEO-FFI personality scores and ALPIS ones in extroversion, openness and agreeableness dimensions. In the other words, ALPIS identifies learners' personality scores as NEO-FFI in extroversion, openness and agreeableness dimensions.

To compare the achieved personality categories from ALPIS and NEO-FFI, the Chi-square test is used. The results of this test show no significant difference between ALPIS and NEO-FFI categories expect in conscientiousness dimension. Because of different results in high category, there is significant difference between ALPIS and NEO-FFI categories. This means some learners report high conscientiousness but their behaviors show different result.

It should be considered that the above results show the difference between ALPIS and NEO-FFI results and do not discuss about the correctness of ALPIS results or NEO-FFI ones. In the other words, this comparison shows difference between observable learner behaviors and their report about theirs elf.

6. Conclusion and future works

The automatic personality identification is a new approach in computer science and psychology especially in e-learning systems. If intelligent tutoring systems have some information about learner's personality then they can adapt teaching strategies based on it. In more psychology researches, personality identifies using self-reports such as through questionnaire. These results include some uncertainty because of hypocrisy, having incorrect perception about theirs self and so on. On the other hand, people behaviors are the reflection of their habits and believes. Therefore, observable behaviors show people personality and can be used to inference about their personality. In this paper, we proposed a fuzzy inference system, named ALPIS, for automatically detecting learner personality based on their observable behaviors in an e-learning environment. The proposed system uses learners' behaviors as fuzzy input variables and then detects their personality based on outer behaviors not based on their self report. Results have shown the efficiency and effectiveness of the proposed approach. Finally, the proposed system solves some important problems ignored by most of the analyzed approaches such as uncertainty and vague in data using fuzzy linguistic variables. The proposed system extract relation between network behaviors and personality traits using expert knowledge and through fuzzy rules while some analyzed approach need training data to extract these relations. Because of using outer behaviors to identify learner's personality proposed system brings advantages to solve self-report disadvantages such as hypocrisy. Having compared ALPIS results and those obtained by the NEO-FFI, the ALPIS can predict learner's personality based on his/her network behaviors with reasonable precision.

We will continue our work on this cross discipline topic to make the whole system better using other intelligent methods. We are going to propose other systems to identify some learners' attributes such as learning style.

It is being noted that according to some limitations we have small sample and will try to test the system using larger sample in the future.

Acknowledgment

This work was partially supported by the Iran Telecommunications Research Center (ITRC), Tehran, Iran, under Grant No. 90-06-82. Authors would like to thank this center.

References

- Ackerman, P. L. (1999). Traits and knowledge as determinants of learning and individual differences: Putting it all together. In P. L. Ackerman, P. C. Kyllonen, & R. D. Roberts (Eds.), *Learning and individual differences: Process, trait, and content determinants* (pp. 437–462).
- Ackerman, P. L., & Beier, M. E. (2002). Individual differences in working memory within a nomological network of cognitive and perceptual speed abilities. *Journal of Experimental Psychology*, 131, 567–589.
- Ackerman, P. L., & Heggestad, E. (1997). Intelligence, personality, and interests: Evidence for overlapping traits. *Psychological Bulletin*, 121, 219–245.
- Agarwal, R., & Prasad, J. (1999). Are individual differences germane to the acceptance of new information technology? *Decision Science*, 30(2), 361–391.
- Bai, Sh., Zhu, T., Cheng, L., (2012), *Big-five personality prediction based on user behaviors at social network sites*, arXiv:1204.4809v1. <<http://arxiv.org>> (at 7.17.13).
- Bersoff, D. M., & Bersoff, D. N. (2000). Ethical issues in the collection of self-report data. In A. A. Stone et al. (Eds.), *The science of self-report: Implications for research and practice* (pp. 9–24). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.
- Butler, T. J., & Pinto-Zipp, G. (2006). Students' learning styles and their preferences for online instructional methods. *Journal of Educational Technology Systems*, 34, 199–221.
- Castillo, G., Gama, J., & e Breda, A. (2005). An adaptive predictive model for student modeling. *Advances in Web-Based Education: Personalized Learning Environments*, 70–92.
- Conti, G. J., & McNeil, R. C. (2011). Learning strategy preference and personality type: Are they related? *Journal of Adult Education*, 40(2), 1–8.
- Costa, P. T., & McCrae, R. R. (1992). *Revised NEO personality inventory (NEO PI-R) and NEO five-factor inventory (NEO-FFI) professional manual*. Odessa, FL: Psychological Assessment Resources.
- De Raad, B., & Schouwenburg, H. C. (1996). Personality in learning and education: A review. *European Journal of Personality*, 10, 303–336.
- Devaraj, S., Easley, R. F., & Crant, J. M. (2008). Research note-how does personality matter? Relating the five-factor model to technology acceptance and use. *Information Systems Research*, 19(1), 93–105.
- Ewen, R. B. (2009). *An introduction to theories of personality* (7th ed.). Mahwah, NJ: Lawrence Erlbaum Associates.
- Feldman, K., Smart, J. C., & Ethington, C. A. (1999). Major field and person environment fit: Using Holland's theory to study change and stability in college students. *The Journal of Higher Education*, 70(6), 642–669.
- Georgiou, D. A., & Makry, D. (2004). A learner's style and profile recognition via fuzzy cognitive map. In *4th IEEE international conference on advanced learning technologies* (ICALT'04).
- Goff, M., & Ackerman, P. L. (1992). Personality-intelligence relations: Assessment of typical intellectual engagement. *Journal of Educational Psychology*, 84, 537–553.
- Gosling, S. D., Augustine, A. A., Vazire, S., Holtzman, N., & Gaddis, S. (2011). Manifestations of personality in online social networks: Self-reported Facebook-related behaviors and observable profile information. *Cyberpsychology, Behavior, and Social Networking*, 14(9), 483–488.
- Hoskin, R., (2012), *The dangers of self-report*, <<http://www.sciencebrainwaves.com>> (12.12.14).
- Huang, Y., -P. C. -C. Lu, & Chang, T. -W. (2007). An intelligent approach to detecting the bad credit card accounts. In *25th IASTED International Multi-Conference Artificial Intelligence and Applications* (pp. 1–6). Innsbruck, Austria.
- Irani, T., Telg, R., Scherler, C., & Harrington, M. (2003). Personality and its relationship to distance education students' course perceptions and performance. *The Quarterly Review of Distance Education*, 4, 445–453.
- Jin, Y., Seelen, W., & Sendhoff, B. (1999). On generating FC fuzzy rule systems from data using evolution strategies. *IEEE Transactions on Systems, Man, and Cybernetics – Part B: Cybernetics*, 29(6), 829–845.
- Kanuka, H., & Nocente, N. (2003). Exploring the effects of personality type on perceived satisfaction with web-based learning in continuing professional development. *Distance Education*, 24(2), 228–245.
- 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.
- Kim, E. B., & Schniederjans, M. J. (2004). The role of personality in web-based distance education courses. *Communications of the ACM*, 47(3), 95–98.
- Klein, D. N. (2003). Patients' versus informants' reports of personality disorders in predicting 7½ year outcome in outpatients with depressive disorders. *Psychological Assessment*, 15, 216–222.
- Komarraju, M., & Karau, S. J. (2005). The relationship between the big five personality traits and academic motivation. *Personality and Individual Differences*, 39, 557–567.
- Lee, J. M., & Lee, Y. (2006). Personality types and learners' interaction in web-based threaded discussion. *Quarterly Review of Distance Education*, 7(1), 83–94.
- Lounsbury, J. W., Sundstrom, E., Loveland, G. M., & Gibson, L. W. (2003). Intelligence, "Big Five" personality traits, and work drive as predictors of course grade. *Personality and Individual Differences*, 35, 1231–1239.
- McDermut, W., & Zimmerman, M. (2005). Assessment instruments and standardized evaluation. In J. M. Oldham, A. E. Skodol, & D. S. Bender (Eds.), *Textbook of personality disorders*. Washington, DC: American Psychiatric Press.
- Olson, D. L., & Delen, D. (2008). *Advanced data mining techniques* p. 138 (1st ed.). Springer.
- Oltmanns, T. F., & Turkheimer, E. (2006). Perceptions of self and others regarding pathological personality traits. In R. F. Krueger & J. L. Tackett (Eds.), *Personality and psychopathology* (pp. 71–111). New York, NY, US: Guilford Press.
- Orr, E. S., Sisic, M., Ross, C., Simmering, M. G., Arseneault, J. M., & Orr, R. R. (2009). The influence of shyness on the use of Facebook in an undergraduate sample. *CyberPsychology & Behavior*, 12(3), 337–340.
- Pavalache-Ilie, M., & Cocorada, S. (2014). Interactions of students' personality in the online learning environment. *Procedia – Social and Behavioral Sciences*, 128, 117–122.
- Petrides, K. V., Vernon, P. A., Schermer, J. A., Lighthart, L., Boomsma, D. I., & Veselka, L. (2010). Relationships between trait emotional intelligence and the Big Five in the Netherlands. *Personality and Individual Differences*, 48, 906–910.
- Vermunt, J. D., Bronkhorst, L. H., & Martínez-Fernández, J. R. (2014). The dimensionality of student learning patterns in different cultures. In D. Gijbels,

- V. Donche, J. T. E. Richardson, & J. D. Vermunt (Eds.), *Learning patterns in higher education: Dimensions and research perspectives* (pp. 33–55). New York: Routledge.
- Wang, L. X. (1997). *A course in fuzzy systems and controls*. Prentice-Hall International, Inc..
- Westen, D. (1997). Divergences between clinical and research methods for assessing personality disorders: Implications for research and the evolution of Axis II. *American Journal of Psychiatry*, 154, 895–903.
- Zuckerman, M. (2005). *Psychobiology of personality* (Second Edition, Revised and Updated). New York: Cambridge University Press.