Chapter 5
Using Data Mining Techniques to Detect the Personality of Players in an Educational Game

Fazel Keshtkar, Candice Burkett, Haiying Li and Arthur C. Graesser

Abstract One of the goals of Educational Data Mining is to develop the methods for student modeling based on educational data, such as; chat conversation, class discussion, etc. On the other hand, individual behavior and personality play a major role in Intelligent Tutoring Systems (ITS) and Educational Data Mining (EDM). Thus, to develop a user adaptable system, the student's behaviors that occurring during interaction has huge impact EDM and ITS. In this chapter, we introduce a novel data mining techniques and natural language processing approaches for automated detection student's personality and behaviors in an educational game (Land Science) where students act as interns in an urban planning firm and discuss in groups their ideas. In order to apply this framework, input excerpts must be classified into one of six possible personality classes. We applied this personality classification method using machine learning algorithms, such as: Naive Bayes, Support Vector Machine (SVM) and Decision Tree.

Keywords Personality · Classification · Conversation · Larry's Rose framework · Natural language processing · Educational data

F. Keshtkar (✉) · C. Burkett · H. Li · A. C. Graesser
Southeast Missouri State University, DH 021F, Mail Stop 5950, USA
e-mail: fkeshtkar@semo.edu

C. Burkett
e-mail: cburkett@memphis.edu

H. Li
e-mail: hli5@memphis.edu

A. C. Graesser
e-mail: a-graesser@memphis.edu

A. Peña-Ayala (ed.), *Educational Data Mining*,
Studies in Computational Intelligence 524, DOI: 10.1007/978-3-319-02738-8_5,
© Springer International Publishing Switzerland 2014

Abbreviations

CBLE	Computer based learning environment
CRF	Conditional random field
EDM	Educational data mining
ITS	Intelligent tutoring system
LIWC	Linguistic inquiry and word count
NPC	Non-player characters
SVM	Support vector machine

5.1 Introduction

Interpersonal conversation is not an easy task. During conversation in educational games, ITS, or chat interaction, the students may have different ideas from the others. Because they may affect by different moods or personality when they listen or say something. On the other hand, students might have different personality characters, i.e., to be cooperative, leading, aggressive, or dependent. For all these reason, we believe personality traits should be considered in computer-based learning environments (CBLE) such as educational game and intelligent tutoring systems. For example, attitudes toward computers can be related to personality types such that those displaying higher scores on neuroticism may have greater computer related anxiety. Furthermore, it is known that it is important to take individual differences into account during learning in CBLE. For example, ITS are known for their ability to simulate effective human tutoring methods as well as take into account the individual needs of learners [1].

Although the efforts to classify personality traits can be a particularly useful endeavour, the detection of personality and/or behavior in conversation using natural language, as it turns out, is a rather difficult task. For example, in serious games in which communication occurs in chat rooms, players may discuss different ideas than others they are chatting with during conversation. Likely, they are also exposed to or affected by the different personalities or moods of other players during communication. On the other hand, players may demonstrate various personality characteristics (such as those related to helping, leading, or aggression) that may result in varied behavioral indicators within conversation.

This chapter aims to investigate, how chat interactions from student log data can be used to determine a student model to classify personality. In results, it turned out that we developed a supervised learning model based on annotated data to automatically detect the students' personality based on their chat interaction in an educational game. The purpose of this research is to identify personality traits of students in textual excerpts in an Educational Game in order to develop an automatic classier that determines the personality characteristics of a student based on their discourse in game. This automatic classier will then be implemented within the ITS module.

This research is divided into two parts; Manual Annotation, and Automatic personality detection. We also aim to answer the important questions: (a) how different types of student's behavior impact other students learning in different ways; (b) how variations (such as human computer interaction) in an ITS and educational game impact students behavior.

Moreover, in this chapter we present a dataset that we have annotated containing personality excerpts based on Leary's Rose framework (Competitive, Dependent, Leading, Helping, Aggressive, Withdrawn). By this, we have presented that the detection of personality behavior is more efficient than that of human judges. Consequently, we have presented three automated methods to personality detection, based on understanding from research in natural language processing (NLP), machine learning, and psychology. We explore that text classification based on n-gram (Unigrams and Bigrams) is the best particular detection approach. We also examined a combination method such as Linguistic Inquiry and Word Count (LIWC) and subjective lexicons features.

In the first task, we performed a coding scheme based on Leary's framework personality dimension by human judges. Therefore, we annotated the personality characteristics of students and their chat interactions from log data set. Furthermore, we have analyzed a random of 200 student's textual excerpts from the chat our annotated data set to test our automated personality detection performance. Two human judges manually annotated this subset of excerpts see Sect. 5.4.

In the second task we develop a supervised method, using data mining, NLP, and machine learning algorithm, to detect the personality of students. We have used machine learning algorithms (i.e., SVM, J48, Naive Bays) for classification. We also used Weka and other NLP tools (i.e., Standford Parser [2], and OpenNLP [3]) to develop this automated system. Our model for classification personality explained in this chapter are performed using a tenfold cross validation method under its default setting in Weka [4]. We reported: Accuracy, Precision, Recall and F-Measure.

We observed that our automated classifier approaches out performed human judges annotation with accuracy of 83 %. We analyzed our results with ANOVA method. It is used to test the difference in LIWC component scores among six types of personality: competitive, leading, between dependent, withdrawn, helping, and aggressive.

The remainder of this chapter is organized as follows. The Sect. 5.2 covers the literature review and the previous works in personality related to education data and student modeling. In the Sect. 5.3 we introduce the Leary's Rose Frame work. The Sect. 5.4 presents the annotation scheme and human annotation for data set. Section 5.5 presents our model, the main functionality of our system for automatic personality classification. Section 5.6 presents the experiences and results. In Sect. 5.7, we illustrate discussion and analysis of our results. Finally, this chapter ends with conclusion and future works in Sect. 5.8.

5.2 Literature Review

5.2.1 Personality in Computer-Based Learning Environments

There are numerous reasons personality traits should be considered in CBLE. For example, even at a very basic level, attitudes toward computers can be related to personality types such that those displaying higher scores on neuroticism may have greater computer related anxiety [1]. Also, it is useful to consider differences in students or group dynamics into account during learning in CBLE. ITS are good examples to measure the ability of students against human tutoring methods as well as needs of learners [1]. This task should not be taken lightly, however, as for both human tutor and ITS, it is difficult to accurately assess both the cognitive and emotional states of individual learners. Similarly, it is a rather complex process to categorize personality traits solely from natural language user input in CBLE.

5.2.2 Emotion Detection Using Leary's Rose Frameboard

Researchers have had some success on the deLearyous gaming project [5]. To our knowledge, this is the only research that has been done specifically on the automatic classification of sentences based on Leary's Rose for emotion detection. DeLearyous researchers described a methodology for a serious gaming project which aims at developing an environment in which users can improve their communication skills by interacting with a virtual character in written natural language (Dutch). In order to apply Leary's framework, they classified the input sentences into one of four possible "emotion" classes (above, below, opposed, together). They applied several machine learning algorithms SVM, Naive Bayes, and Conditional Random Field (CRF) to obtain the classification performance. For this, they used different features set from their dataset (unigrams, lemma trigrams and dependency structures). They obtained 52.5 % accuracy, around 25 % over the baseline. The researchers noted, however, that the manually annotated sentences used to compile their training set were labeled by one human annotator and thus may have been susceptible to issues with reliability.

5.2.3 Automatic Detection of Personality

In other research [6, 7] found that identification of personality (Big Five in speech) by automatic analysis performed better than the baseline. Their analysis confirms previous findings linking language and personality and also reveals many new linguistic and prosodic markers. However, there was a limitation in their method in that speech recognition, such as prosodic features.

In addition, there has been other research conducted in order to let a machine learner determine the appropriate sentiment/emotion class. For instance, [8] and [9] attempted to classify LiveJournal posts according to their mood using SVM trained with frequency features (word counts, POScounts), length-related features (length of posts/sentences, etc.), semantic orientation features (using WordNet to calculate the distance of each word to a set of manually classified keywords) and special symbols (emoticons).

5.2.4 Personality and Student Behavior

Gore et al. investigated the relation between personality and organizational citizenship behaviors in student populations [10]. They tested the hypothesis that conscientiousness, agreeableness, and neuroticism predict unique variance in academic citizenship attitudes. They studied 270 college students who completed an online questionnaire assessing their personality and academic citizenship attitudes.

They claimed that results confirmed the hypothesis. In another study they also found that academic citizenship attitudes mediate the association between personality and citizenship behavior. Their results showed that general conscientiousness was associated with citizenship behavior, but academic conscientiousness attitudes mediated this association.

5.2.5 The Relationship Between Personality Traits and Information Competency

Song and Kwon examined differences between Korean and American cultures in terms of the relationships between Big Five personality traits and information competency [11]. In their research, Korean ($n = 245$) and American ($n = 185$) college students completed the NEO-Five Factor Inventory and the Information Competency Scale. Their results showed both similarities and differences between the two culture groups.

They showed that Conscientiousness and openness to experience significantly predicted information competency in both Korean and American students. On the other hand, they conducted that the influence of extroversion was significant only for American students. This result happened due to the high value placed on extroversion in American culture [11].

5.2.6 Personality Traits and Learning Style in Academic Performance

In [12], Furnham conducted various tests soon after students arriving at university on the Big Five Personality Traits [13]. The first study (N = 178) showed Conscientiousness and General intelligence to be the only significant predictor of overall first year grade accounting for 11 % of the variance.

The second study (N = 93) showed that ability and non-ability factors differed in terms of their predictive validity depending on the exams taken. Individual difference factors account for around 10 % of the variance in college examination success [12].

5.2.7 A Neural Network Model for Human Personality

In this research, Read et al. [14], presented a neural network model that aims to bridge the historical gap between dynamic and structural approaches to personality. The model integrates work on the structure of the trait lexicon, the neurobiology of personality, temperament, goal-based models of personality, and an evolutionary analysis of motives. It is organized in terms of two overarching motivational systems, an approach and an avoidance system, as well as a general des-inhibition and constraint system. Each overarching motivational system influences more specific motives.

Traits are modeled in terms of differences in the sensitivities of the motivational systems, the baseline activation of specific motives, and inhibitory strength. The result is a motive-based neural network model of personality based on research about the structure and neurobiology of human personality. The model provides an account of personality dynamics and person situation interactions and suggests how dynamic processing approaches and dispositional, structural approaches can be integrated in a common framework [15].

5.2.8 Relationships Between Academic Motivation and Personality Among the Students

Relationships between personality and academic motivation were examined using 451 first-year college students [16]. In this research, multiple regressions compared three types of intrinsic motivation, three types of extrinsic motivation and motivation to five personality factors. Results indicated that those who were intrinsically motivated to attend college tended to be extroverted, agreeable, conscientious, and open to new experiences; although these trends varied depending on the specific type of intrinsic motivation.

Those who lacked motivation tended to be extroverted, agreeable, conscientious, and neurotic; depending on the type of extrinsic motivation. Those who lacked motivation tended to be disagreeable and careless. These results suggest that students with different personality characteristics have different reasons for pursuing college degrees and different academic priorities [16].

5.2.9 Relation Between Learning from Errors and Personality

This research focused on the relationship between negative emotionality and learning from errors [17]. Specifically, negative emotionality was expected to impair learning from errors by decreasing motivation to learn. Perceived managerial intolerance of errors was hypothesized to increase negative emotionality, whereas emotional stability was proposed to decrease negative emotionality. All the hypotheses were tested in a laboratory simulation.

Contrary to the prediction, a positive association was found between negative emotionality and motivation to learn. The effects of perceived managerial intolerance of errors and emotional stability on negative emotionality were as predicted. Moreover, exploratory data analysis were conducted at the level of specific negative emotions and revealed differentiated effects of specific negative emotions on learning from errors [17].

5.2.10 Academic Achievement and Big Five Model

Poropat [18] reported a meta-analysis of personality-academic performance relationships, based on the 5-factor model, in which cumulative sample sizes ranged to over 70,000. Most analyzed studies came from the tertiary level of education, but there were similar aggregate samples from secondary and tertiary education. There was a comparatively smaller sample derived from studies at the primary level. Academic performance was found to correlate significantly with Agreeableness, Consciousness, and Openness. Where tested, correlations between Conscientiousness and academic performance were largely independent of intelligence.

When secondary academic performance was controlled for, Conscientiousness added as much to the prediction of tertiary academic performance as did intelligence. Strong evidence was found for moderators of correlations. Academic level (primary, secondary, or tertiary), average age of participant, and the interaction between academic level and age significantly moderated correlations with academic performance. Possible explanations for these moderator effects are discussed, and recommendations for future research are provided [14].

5.2.11 The Big Five Personality, Learning Styles, and Academic Achievement

Personality and learning styles are both likely to play significant roles in influencing academic achievement [19]. College students (308 undergraduates) completed the Five Factor Inventory and the Inventory of Learning Processes and reported their grade point average. Two of the Big Five traits, conscientiousness and agreeableness, were positively related with all four learning styles (synthesis analysis, methodical study, fact retention, and elaborative processing), whereas neuroticism was negatively related with all four learning styles.

In addition, extraversion and openness were positively related with elaborative processing. The Big Five together explained 14 % of the variance in grade point average (GPA), and learning styles explained an additional 3 %, suggesting that both personality traits and learning styles contribute to academic performance. Further, the relationship between openness and GPA was mediated by reflective learning styles (synthesis-analysis and elaborative processing). These latter results suggest that being intellectually curious fully enhances academic performance when students combine this scholarly interest with thoughtful information processing. Implications of these results are discussed in the context of teaching techniques and curriculum design [19].

5.2.12 Using Personality and Cognitive Ability to Predict Academic Achievement

Beaujean et al. [20], conducted a study on the relationship between cognitive ability, personality, and academic achievement in post-secondary students, using latent variable models. By testing both simple and complex relationships, they found that cognitive ability and personality predicted reading achievement independently, but that they interact when predicting math achievement, at least in the Conscientiousness and Openness to Experience domains [20].

5.3 Leary's Interpersonal Frame Board

Leary's Interpersonal Circumplex (or Leary's Rose Frame Board) has been used by researchers for decades as a foundation for categorizing personality through the discourse [21]. The Circumflex defines characteristics according to two dimensions: the above-below axis represents variation from dominant (above) to submissive (below) whereas the opposed-together axis represents variations of cooperation from accommodating (together) to opposition (opposed) (See Fig. 5.1). Based on these two dimensions, the Rose can easily be separated into four quadrants and then further split into eight different categories (See Table 5.1 for examples) [6].

Fig. 5.1 Leary's interpersonal circumplex (Leary's Rose). *OPP* opposite, *TOG* together

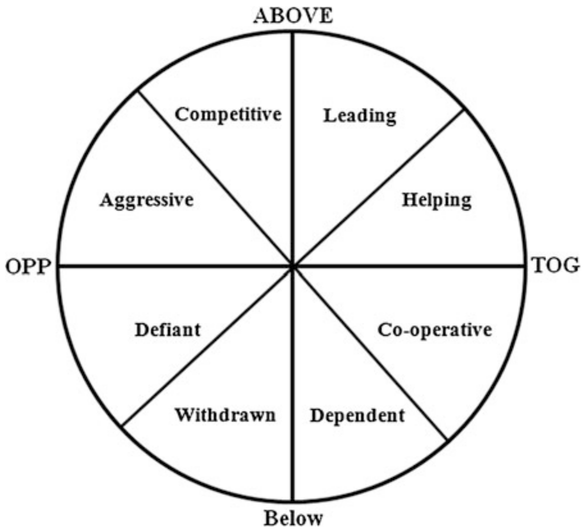

Table 5.1 Leary's Rose categories examples from land science game

Statement	Leary category	Leary quadrant
Finish your task now so we can move on	Leading	Above-together
How can I help you with that?	Helping	Above-together
My plan is better than your plan	Competitive	Above-opposed
That idea is stupid. It will never work	Aggressive	Above-opposed
Sure, we can work together on this project	Cooperative	Below-together
What should I do now?	Dependent	Below-together
Sorry, never mind, I'm not thinking	Withdrawn	Below-opposed
No. I am not going to do that	Defiant	Below-opposed

5.3.1 Land Science Game

Land Science is a serious game created by researchers at the University of Wisconsin-Madison that has been designed to simulate a regional planning practicum experience for students [11, 22–24]. During the 10-hour game, students play the role of interns at a fictitious regional planning firm (called Regional Design Associates).

Where they make land use decisions in order to meet the desires of virtual stakeholders who are represented by Non-Player Characters (NPC). Students are split into groups and progress through a total of 15 stages (all these stages are shown in Table 5.2) of the game in which they complete a variety of activities including a virtual site visit of the community of interest in which students familiarize themselves with the history and ecology of the area as well as the desires of different stakeholder groups.

Table 5.2 Land science sequence of activities

Stage #	Activity
1	Intake interview
2	Staff page
3	Request for proposals
4	Virtual site visit and site assessment
5	iPlan
6	TIM 1
7	Preference survey 1
8	Stakeholder assessment 1
9	TIM 2
10	Preference survey 2
11	Stakeholder assessment 2
12	Final plan (individual)
13	Final proposal (individual)
14	Reflection
15	Exit interview

In addition, students get feedback from the stakeholders, and use a custom designed Geographic Information System (iPlan) to create a regional design plan. Throughout the game players communicate with other members of their planning team as well as a mentor (i.e., an adult who is representing a professional planner with the fictitious planning firm) through the use of a chat feature that is embedded in the game.

5.3.2 Participants and Data Set Construction

Participants included 12 middle school students who played the epistemic game Land Science as a part of an enrichment program at the Mass Audubon Society in Massachusetts. As previously mentioned, players in the game communicated with both other players and mentors using a chat feature embedded in the interface. For the purposes of detecting the personality of players, we only analyzed the players' chat excluding mentors' chat. Annotation was done using the coding scheme (further discussed under Human Annotation in Sect. 5.4) that was developed by the researchers based on the Timothy Leary's Interpersonal Circumplex Model [21]. The researchers selected 1,000 excerpts (average = 4.8 words) to be analyzed. For our purposes, an excerpt was defined as a turn of speech that was taken by the student.

On the other words, one excerpt occurred each time a student typed something and clicked "send" or hit "enter" in the chat function. The excerpts were selected from a larger set of 3,227 excerpts, so approximately 31 % of the player excerpts were randomly used in the analyzed data set. We have used the distribution for all stages for selecting data set. Our model is illustrated in Fig. 5.2 and in the following sections we describe the components of this model.

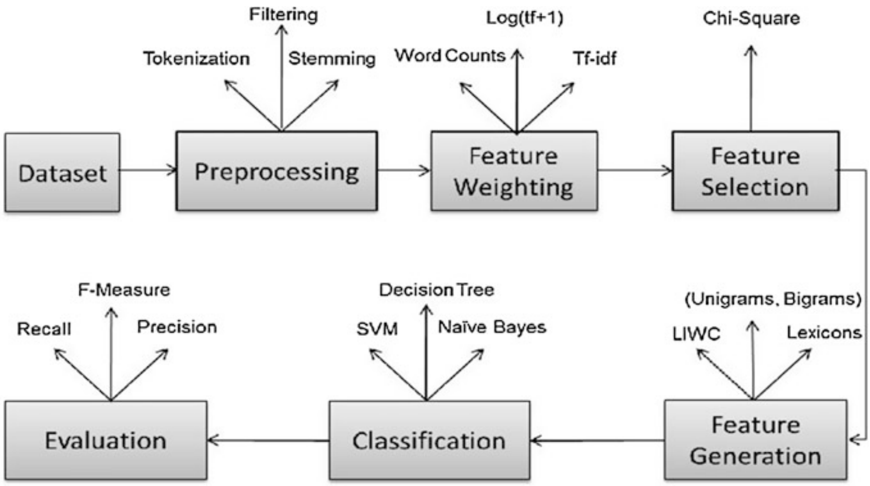

Fig. 5.2 Architecture of automated classification

5.4 Annotation Scheme

As previously described, the researchers developed a coding scheme based on Leary's Interpersonal Circumflex, which focused on 6 categories from all 4 quadrants of Leary's Rose: Competitive, Leading, Dependent, Withdrawn, Aggressive/Defiant and Helping/Cooperative.

Regarding the current annotation scheme we combined Leary's original categories of Aggressive and Defiant because in our data set there was little differentiation between these two categories. Similarly, we also combined Leary's original categories of Helping and Cooperative. Definitions and examples of each of these categories are included in Table 5.3.

5.4.1 Human Annotation

Using this coding scheme, two trained researchers annotated the data set of 1,000 excerpts. The first series of training required the human annotators to independently code 200 excerpts randomly selected from the Land Science corpus. The Kappa statistic was computed to assess inter-rater reliability on this set and agreement was fair (0.33). Following this, the annotators discussed and refined any issues regarding the coding scheme and then annotated a new set of randomly selected excerpts. The Kappa statistic was computed to assess inter-rater reliability on the second training set and agreement was substantial (0.69). Results indicated increased reliability and thus completed the training of the human annotators. Once the two annotators were trained they independently annotated a set of 1,000 excerpts.

Table 5.3 Annotation scheme; category definitions and examples from land science

Category	Leary definition	Additional information	Land science example
Competitive	Narcissistic, competing, acting confidently, boast, brag, act proud	Competitive with another orby indicating a desire to do well in the game	Beat team Eva!!
Leading	Managerial, directing, guiding, advising, teaching, ordering around, bossing	Can include explicit or indirect request	We are going to have a team meeting in about 10 min, so we need to finish our site assessments
Dependent	Asking for help, depend on, act in an over respecting manner	Seeking direction or approval	What should I say now?
Withdrawn	Acting shy or sensitively, being modest, self-condemning	Does not include lack of responses to question	Sorry, never mind. I'm not thinking
Aggressive/defiant	Rebellious actions, complaining, wariness, being skeptical	Also includes taking a strong stance and passing the blame on to someone else	No!! I'll say in a second. I'm on something else
Helping/cooperative	Takes responsibility, helping, offering, giving, agree, co-operate, compromise	Includes working together as a group or participating in group activity	If you want me to look at your plan I can.

Table 5.4 Inter-rater reliability (Kappa) for 1,000 coded excerpts

Personality category	Inter-rater reliability (kappa)
Competitive	0.82
Leading	0.65
Dependent	0.83
Withdrawn	0.77
Helping/cooperative	0.58
Aggressive/defiant	0.65
Neutral	0.60
Overall average	0.70

Overall, personality category agreement between the two annotators on the set of 1,000 excerpts was substantial (Kappa = 0.70). As shown in Table 5.4, agreement is substantial for the Competitive, Dependent and Withdrawn categories, and is moderate for the Leading, Helping/Cooperative, Aggressive/Defiant and Neutral categories. The two human annotators agreed on the personality category for a total number of 1,523 excerpts (see Table 5.5). Of those agreements, the largest percentage of excerpts is Neutral (35.78 %), indicating that the annotators agreed that there was no evidence of a personality category represented.

Table 5.5 Number of instances agreed present for each personality category

Personality category	Number of instances
Competitive	277
Leading	134
Dependent	343
Withdrawn	46
Helping/cooperative	141
Aggressive/defiant	37
Neutral	545
Total	1,523

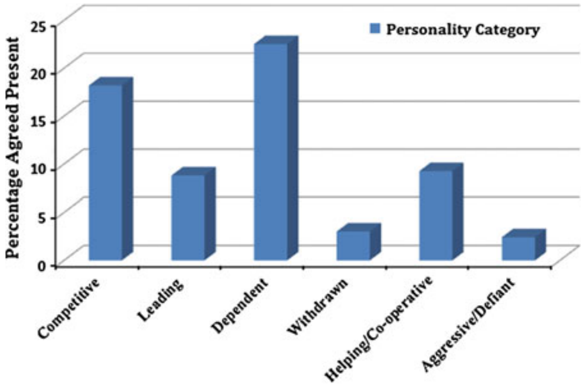

Fig. 5.3 Percentage of excerpts agreed in each personality category

Regarding excerpts for which there is an agreement that a personality category present, as seen in Fig. 5.3, the largest percentage is Dependent (22.52 %) followed by Competitive (18.19 %) and Helping/Cooperative (9.26 %). The least represented personality categories are Leading (8.80 %), Withdrawn (3.02 %) and Aggressive/Defiant (2.43 %).

5.5 Model

To our knowledge, the only research has been done specifically on the automatic classification of sentences based on Learys Rose for emotion detection is done by [25]. They described a methodology for a serious gaming project, deLearyous, which aims at developing an environment in which users can improve their communication skills by interacting with a virtual character in (Dutch) written natural language. In order to apply this framework, they classified the input sentences into one of four possible "emotion" classes (above, below, opp, tog, see Fig. 5.1).

They applied several machine learning algorithms, such as SVM, Naïve Bayes, Conditional Random field to obtain the calcification performance. For this, they used different features set from their dataset (unigrams, lemma trigrams and dependency structures). They obtained 52.5 % accuracy around 25 % over the baseline. In contrast, in our method we use Leary's Rose framework to detect personality rather than emotion.

5.5.1 Lexicon Resources

Sentiment-based lexical resources annotate words/concepts with polarity. To achieve greater coverage, we use four different sentiment-based lexical resources. They are described as follows.

1. SentiWordNet [26]. Assigns three scores to Synsets of WordNet: positive score, negative score and objective score. When a word is looked up, the label corresponding to maximum of the three scores is returned. For multiple synsets of a word, the output label returned by majority of the Synsets becomes the prediction of the resource.
2. Subjectivity lexicon [25]. Is a resource that annotates words with tags like parts-of-speech, prior polarity, magnitude of prior polarity (weak/strong), etc. The prior polarity can be positive, negative or neutral. For prediction using this resource, we use this prior polarity.
3. General Inquirer [27]. Is a list of words marked as positive, negative and neutral. We use these labels to use Inquirer resource for our prediction.
4. Taboada [28]. It is a word-list that gives a count of collocations with positive and negative seed words. A word closer to a positive seed word is predicted to be positive and vice versa.

5.5.2 Feature Extraction

From this dataset we extracted a wide range of different features. The sentences were first parsed with Stanford POS Tagger, an English language parser [2], which allowed us to extract linguistic information such as word tokens, lemmas, part-of-speech tags, syntactic functions and dependency structures.

The actual feature vectors were then generated on the basis of this linguistic information by using a "bag of n-grams" approach, i.e. by constructing n-grams (unigrams, bigrams and trigrams) of each feature type (e.g. n-grams of word tokens, n-grams of part-of-speech tags…) and by counting for each n-gram in the training data how many times it occurs in the current instance. In addition to these n-gram counts, we also included punctuation counts, average word length and average sentence length.

Sentiment Score Feature. Based on predictions of individual traits, we compute the Sentiment prediction for each trait with respect to a keyword in form of percentage of positive, negative and objective content. This is on the basis of predictions by each resource by weighting them according to their accuracies. These weights have been assigned to each resource based on experimental results. For each resource, the following scores are determined (see Eqs. 5.1, 5.2, 5.3).

$$PositiveScore(s) = \sum_{i=0}^{i=n} PiW_{Pi} \tag{5.1}$$

$$NegativeScore(s) = \sum_{i=0}^{i=n} NiW_{Ni} \tag{5.2}$$

$$ObjectiveScore(s) = \sum_{i=0}^{i=n} OiW_{Oi} \tag{5.3}$$

where, PositiveScore(s) = Positive score for each excerpt s; NegativeScore(s) = Negative score for each excerpt *s*; ObjectiveScore(s) = Objective score for each excerpt s; n = Number of resources used for prediction; Pi, Ni, Oi = Positive, Negative, and Objective count of excerpt predicted respectively using resource i; WPi,WNi,WOi = Weights for respective classes derived for each resource *i*.

5.5.3 The Linguistic Inquiry and Word Count Features

We extracted features derived from the Linguistic Inquiry and Word Count (LIWC) output. Specifically, LIWC counts and groups the number of instances of nearly 4,500 keywords into 80 psychologically meaningful dimensions. We create one feature for each of the 80 LIWC dimensions summarized under the following four categories:

- Linguistic processes: Functional aspects of text (e.g., the average number of words per sentence, the rate of misspelling, swearing, etc.)
- Psychological processes: Includes all social, emotional, cognitive, perceptual and biological processes, as well as anything related to time or space.
- Personal concerns: Any references to work, leisure, money, religion, etc.
- Spoken categories: Primarily filler and agreement words.

For each instance, we calculate the ratio of words in each category from the LIWC toolkit [18], as these features are correlated with the personality dimensions (as shown in Table 5.6). Indeed, the LIWC2007 software used in our experiments subsumes most of the features introduced in other work. Thus, we focus our psycholinguistic approach to personality detection on LIWC-based features.

Table 5.6 LIWC features [18]

Feature category	Features included
Standard counts	*Word count* Words per sentence, type/token ratio, words captured, words longer than 6 letters, negations, assents, articles, prepositions, numbers, pronouns: 1st person singular, 1st person plural, total 1st person, total 2nd person, total 3rd person
Psychological processes	*Affective or emotional processes* Positive emotions, positive feelings, optimism and energy, negative emotions, anxiety or fear, anger, sadness, cognitive processes: causation, insight, discrepancy, inhibition, tentative, certainty, sensory and perceptual processes: seeing, hearing, feeling, social processes: communication, other references to people, friends, family, humans
Relativity	*Time* Past tense verb, present tense verb, future tense verb, Space: up, down, inclusive, exclusive, motion
Personal concerns	*Occupation* School, work and job, achievement, leisure activity: home, sports, television and movies, music, money and financial issues, metaphysical issues: religion, death, physical states and functions, body states and symptoms, sexuality, eating and drinking, sleeping, grooming
Other dimensions	*Punctuation* period, comma, colon, semi-colon, question, exclamation, dash, quote, apostrophe, parenthesis, other, Swear words, non-fluencies, fillers

For each instance, we calculate the ratio of words in each category from the LIWC toolkit [18], as these features are correlated with the personality dimensions [18]. These features and their categories are shown in below.

5.5.4 Automated Approaches to Personality Classification

We explain three automated approaches to classify detecting personality behavior, each of which utilizes classifiers trained on the dataset of Sect. 5.3.2. The features employed by each strategy are described here.

Psycholinguistic Personality Detection. The Linguistic Inquiry and Word Count (LIWC) software [18] is a popular automated text analysis tool used widely in the social sciences. It has been used to detect personality traits [6], to study tutoring dynamics [29], and, most relevantly, to analyze personality detection [6].

Since LIWC software does not include a text classifier, we create features derived from the LIWC output. In particular, LIWC counts and groups the number of instances of nearly 4,500 keywords into 80 psychologically meaningful dimensions. We construct one feature for each of the 80 LIWC dimensions, which can be summarized under the four categories that explained in Sect. 5.3. Indeed, the LIWC2007 software used in our experiments subsumes most of the features introduced in other work. Thus, we focus our psycholinguistic approach to personality detection on LIWC-based features.

5.5.5 Classification Method

Naive Bayes Classifier Provides a simple approach and it is a classifier as a form of Bayesian network and it leans on two simple assumptions. First, it assumes that the predictive attributes are conditionally independent given the class. Then, it posits that no hidden or latent attributes influence the prediction process [30]. For a document X, with label class c, the Naive Bayes classifier gives us the following decision rules (see Eqs. 5.4 and 5.5) [30]:

$$P(C = c|X = x) = \frac{p(C = c)\, p(X = x|C = c)}{p(X = x)}, \tag{5.4}$$

where

$$P(X = x|C = x) = \prod_{i}^{n} P(Xi = xi|C = c) \tag{5.5}$$

We use John and Langley [30] Naïve Bayes classifier in Weka [4] to train our Naive Bayes models on all three approaches and feature sets described above, namely LIWC, lexicons, Unigrams, Bigrams. We also evaluate every combination of these features, but for brevity include only UNIGRAMS + BIGRAMS, which performs best with tenfold cross validation on the corresponding dataset.

Support Vector Machine. We also train SVM classifiers, which find a high-dimensional separating hyper-plane between two groups of data. To simplify feature analysis in Sect. 5.5, we restrict our evaluation to linear SVM, which learn a weight vector *w* and bias term *b*, such that a document x can be classified by (5.6):

$$y = sign\ (\vec{w}.\vec{x}) + b \tag{5.6}$$

We use SMO [31] to train our SVM models on all three approaches and feature sets described above: LIWC, LEXICONS, UNIGRAMS, and BIGRAMS. We also evaluate every combination of these features, but for shortness include only LIWC + BIGRAMS, and LEXICON + BIGRAMS which performs best.

Decision Trees. We use J48, an open source Java implementation of the C4.5 algorithm in Weka [4] data mining tool to train our dataset for decision trees classifier. We evaluate approaches on all combination of feature set, but we consider the features which performed best (UNIGRAMS + BIGRAMS, UNIGRAMS + LIWC). Our classification experiments are carried out with tenfold cross validation on the corresponding dataset. A sample of the results achieved by the three methods is stated in Table 5.7.

Table 5.7 Automated classifier performance for three approaches based on tenfold cross- validation experiments

Approach	Features	Acc. (%)	COM			DEP		
			P	R	F	P	R	F
LEXICAL	Lexicons$_{j48}$	61.95	0.67	0.66	0.67	0.56	0.66	0.61
LIWC	Liwc$_{j48}$	59.30	0.57	0.62	0.60	0.64	0.67	0.65
Method	Unigrams$_{svm}$	60.54	0.74	0.70	0.72	0.64	0.63	0.63
	Bigrams$_{svm}$	70.40	0.92	0.65	0.76	0.92	0.75	0.82
	Liwc + bigrams$_{\mathrm{svm}}$	77.47	0.90	0.75	0.82	0.93	0.78	0.85
	Lexicons + bigrams$_{\mathrm{svm}}$	83.71	0.96	0.80	0.87	0.96	0.84	***0.**90*
	Bigrams$_{nb}$	65.02	0.04	0.87	0.62	0.87	0.77	0.50
	Unigrams + bigrams$_{nb}$	60.53	0.77	0.60	0.67	0.72	0.68	0.50
	Unigrams + bigrams$_{j48}$	62.78	0.83	0.67	0.74	0.83	0.71	0.46
	Unigrams + liwc$_{j48}$	74.0	0.86	0.80	0.83	0.81	0.77	0.63

Approach	Features	Acc. (%)	LEA			WIT		
			P	R	F	P	R	F
LEXICAL	Lexicons$_{j48}$	61.95	0.61	0.55	0.58	0.56	0.54	0.55
LIWC	Liwc$_{j48}$	59.30	0.52	0.40	0.45	0.62	0.42	0.50
Method	Unigrams$_{svm}$	60.54	0.50	0.32	0.50	0.83	0.39	0.53
	Bigrams$_{svm}$	70.40	0.79	0.40	0.52	1	0.44	0.62
	Liwc + bigrams$_{\mathrm{svm}}$	77.47	0.95	0.64	0.77	0.93	0.54	0.68
	Lexicons + bigrams$_{\mathrm{svm}}$	83.71	0.98	0.76	0.86	1	0.74	0.85
	Bigrams$_{nb}$	65.02	0.50	0.21	0.3	0.80	0.44	0.57
	Unigrams + bigrams$_{nb}$	60.53	0.50	0.53	0.51	1	0.39	0.54
	Unigrams + bigrams$_{j48}$	62.78	0.46	0.43	0.45	0.82	0.47	0.60
	Unigrams + liwc$_{j48}$	74.0	0.63	0.64	0.63	0.85	0.78	0.81

Approach	Features	Acc. (%)	COP			AGG		
			P	R	F	P	R	F
LEXICAL	Lexicons$_{j48}$	61.95	0.66	0.60	0.63	0.25	0.21	0.22
LIWC	Liwc$_{j48}$	59.30	0.60	0.62	0.61	0.50	0.53	0.51
Method	Unigrams$_{svm}$	60.54	0.52	0.74	0.61	0.17	0.33	0.22
	Bigrams$_{svm}$	70.40	0.50	1	0.66	1	0.17	0.29
	Liwc + bigrams$_{\mathrm{svm}}$	77.47	0.54	0.95	0.69	1	0.2	0.33
	Lexicons + bigrams$_{\mathrm{svm}}$	83.71	0.98	0.76	0.86	*1*	*0.32*	*0.48*
	Bigrams$_{nb}$	65.02	0.46	0.96	0.62	1	0.16	0.28
	Unigrams + bigrams$_{nb}$	60.53	0.40	0.74	0.52	0.67	0.5	0.57
	Unigrams + bigrams$_{j48}$	62.78	0.42	0.84	0.56	0.26	0.22	0.24
	Unigrams + liwc$_{j48}$	74.0	0.63	0.75	0.69	0.50	0.68	0.57

Reported Accuracy, (*P*) precision, (*R*) recall and (*F*) measure

5.6 Experience and Results

5.6.1 Classification Results

The model for classification personality strategies explained in Sect. 5.5 are per-formed using a tenfold cross validation method under its default setting in Weka [4]. The parameters for model are chosen for each test fold based on standard cross validation experiments on the training dataset. All folds are chosen so that each includes all instances from six classes; therefore, learned classifiers are always measured on dataset from unseen instances.

Table 5.8 shows the results of the top scores that we managed to achieve with each of the three classifiers over three approaches. We also use the combination of features and learner parameters that were determined to give the best accuracy by the classifiers. "Approach" column shows the model that have been tested, the "features" column indicates the types of features that have been used, the rest of columns indicates the results based on Accuracy, Precision, Recall, and F-measure (Acc., P, R, F) for all six classes. We observe that our automated approaches outperformed human judges (Kappa) and baseline for most of feature sets. The statistical baseline for these six classes classification problem, considering the slight imbalances in the class distribution, is 30 %. However there is an exception such as Recall for "aggressive" which is not significant.

We can argue on this due to low number of instances in this class. However, this is expected given that human judges often focus on unreliable cues to aggressive utterances. We observe that our automated approaches outperformed

Table 5.8 Top 15 highest weighted features learned by BIGRAMS + LEXICONSsvm and LIWCsvm. The results show for binary classification of "helping, aggressive" and "leading, dependent"

BIGRAMS + LEXICONSsvm	LIWCsvm
Helping, aggressive	Leading, dependent
Always want	Six letters
Didn't seem	Pronoun
Don t	Personal pronoun
For me	I
Is quite	We
It is	You
Need to	She/he
No need	They
People don	Impersonal pronouns
Quite deadly	Article
Really that	Verb
Seem to	Auxiliary verbs
Slow down	Past tense
Speaking Spanish	Present tense

human judges (Kappa) and baseline for most of feature sets. The statistical baseline for these six classes classification problem, considering the slight imbalances in the class distribution, is 30 %. However there is an exception such as Recall for "aggressive" which is not significant. We can argue on this due to low number of instances in this class. However, this is expected given that human judges often focus on unreliable cues to aggressive utterances.

If we look at the confusion matrix in Table 5.9; firstly, we note that most of the aggressive instances (8) classified as "helping" personality. Many other classes considered as "helping" as well. We figured out, this happened due to human judge's evaluation, because the judges considered many small responses such as: OK, Yep, Thanks, Cool, etc. as "helping" class. Secondly, as it shown in Table 5.1 the number of instances in "aggressive" class is low. We found out that the players are not often aggressive during chat conversation. It might be due to their work environment in that they are supervised by a human mentor during the game.

Interestingly, the psycholinguistic approach (LIWCj48) performs almost 30 % more accurately than baseline rather than SVM or NB. Also J48 perform higher than SVM and NB on lexical subjective scores features. Overall, all the standard text categorization approaches proposed in Sect. 5.5 perform between 9 and 53 % more accurately than baseline. However, best performance overall is achieved by combining features from these two approaches. Particularly, the combined model LEXICONS + BIGRAMSSVMis 83.71 % accurate at personality classification.

Surprisingly, models trained only on UNIGRAMSsvm(60.54 %), the simplest n-gram feature set, outperform LIWC (non-text classification) approaches, and models trained on BIGRAMSnb(65.02 %) perform even better. This suggests that a universal set of feature such as psycholinguistic keyword personality (i.e., LIWC) cannot be the best model for personality detection, and a context-sensitive approach (e.g., BIGRAMS) might be necessary to achieve state-of-the-art personality detection performance.

To better understand the models learned by these automated approaches, we report in Table 5.8 the top 15 highest weighted features for two pair classes (Helping, Aggressive and Leading, Dependent) as learned by BIGRAMS + LEXICONSsvm and LIWCsvm. From BIGRAMS + LEXICONSsvm approach we have chosen classifier for classes "Helping" (with highest F-measure) and "Aggressive" (lowest F-measure), for LIWCsvm approach we have chosen classifier for classes "Leading, Dependent" with similar reason.

We note that player with "Helping" personality behavior tend to use somehow similar language with "Aggressive" players; in particular, "need to" and "no need", the former one can be consider as "Helping" behavior and later one can be regarded as "Aggressive" attitude. Accordingly, in term of global features such as psycholinguistic features (LIWC), "Leading" and "Dependent" players tend to use similar pronouns(personal or impersonal) (i.e.; i, we, you, she/he, they). Finally, when we look at Confusion Matrix (Table 5.9), it turns out that all misclassified instances from "Aggressive" class fall into "Helping" class and similarly almost 75 % of misclassified instances in "Leading" class are classified as "Dependent" class.

Table 5.9 The confusion matrix performed by SVM classifiers approach over BIGRAMS and subjective lexicon features

a	b	c	d	e	f	Classified as
130	1	2	1	37	0	a = competitive
2	155	1	0	47	0	b = dependent
5	4	61	0	26	0	c = leading
0	0	0	14	12	0	d = withdrawn
2	1	0	0	146	0	e = helping
0	0	0	0	8	2	f = aggressive

5.7 Discussion and Analysis

5.7.1 Personality Trait Tracking Analysis

An additional aim of the current study is to explore the consistency of personality characteristics displayed by individual participants across the various stages of the game. In order to do this we randomly selected two participants and charted their-coded personality traits throughout the game.

For the purposes of the current results we focused only on three of the most prevalent personality categories overall (Competitive, Leading and Dependent). Figs. 5.4 and 5.5 display the personalities displayed by these two players (referred to as Player A and B) for each of the 15 stages of the game (numbered 0–14). First, it is important to note that both players exhibited different personalities during different stages of the game. More specifically, Player A (see Fig. 5.4) demonstrated a variety of noticeable trend for the first few stages of the game. However, there was a drastic increase in Dependent statements in stage 6 followed by an increase in Leading statements in stage 7. Competitive statements then become the most dominant for most of the final stages of the game.

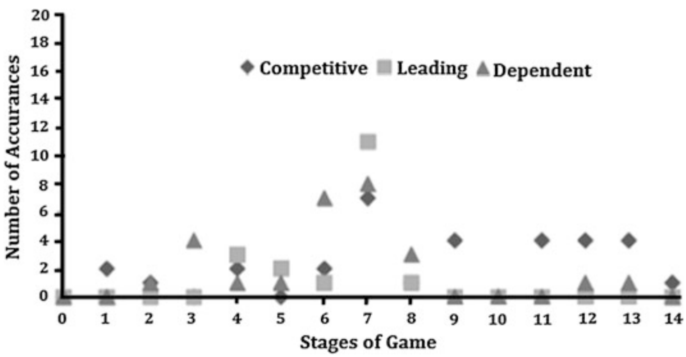

Fig. 5.4 Player A personality characteristics displayed for each stage of the game

In addition, Player B (See Fig. 5.5) exhibited a variety of personality characteristics throughout the game. For example, Dependent statements dominated 6 of the first 9 stages of the game with a drastic increase in Stage 7. However, like Player A, Competitive statements were most prevalent for the final 6 stages of the game.

Based on the above results, the changes that occur in Stage 7 of the game seem to be especially relevant. These changes highlight that players may be altering their statements based both on the demands of the game as well as the personalities exhibited by other players in the group dynamic. Specifically with the above examples, notice that during Stage 7 of the game Player A had a drastic increase in Leading statements while Player B had a drastic increase in Dependent statements. It is possible that there may be something about the task associated with Stage 7 that encourages a group dynamic in which some players become more dependent while others become more directive.

Overall, results indicate substantial agreement between two trained human annotators. Regarding coded personality categories Leading, Dependent, Helping/Cooperative and Competitive are the four most commonly present categories, whereas, Withdrawn and Aggressive/Defiant statements are less prevalent. Furthermore, players demonstrate different personality characteristics depending on the stage of the game and, likely, the dynamics of the group.

5.7.2 ANOVA Analysis

One way analysis of variance (One-way ANOVA) is used when two or more groups are compared with their mean scores on one continuous variable, also called the independent variable. A one way analysis of variance (ANOVA) will tell people whether these groups differ.

Consequently, post hoc comparisons will help to test which groups are significantly different from one another. One-way between-groups ANOVA was used

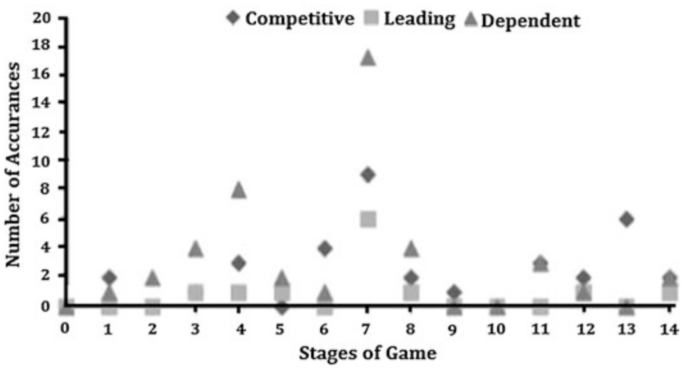

Fig. 5.5 Player B personality characteristics displayed for each stage of the game

to test the difference in LIWC component scores among six types of personality: competitive, leading, dependent, withdrawn, helping, and aggressive. The type of personality is one factor and normalized LIWC components related to psychological features are the dependent variables: Argumentation (Persuasion), Achievement, and Negative Valence [32]. Table 5.10 shows the ANOVA results that each LIWC component scores differed significantly across the three types of personality:

- Argumentation, F (5, 521) = 5.12, $p < .001$
- Achievement, F (5, 521) = 7.26, $p < .001$
- Narrative, F (5, 521) = 67.87, $p < .001$
- Negative Valence, F (5, 521) = 55.45, $p < .001$; and
- Embodiment, F (5, 521) = 12.35, $p < .001$.

Tamhane post hoc tests comparisons of the six groups indicate for LIWC component, Argumentation, the dependent personality (M = 3.05, 95 % CI [2.58, 3.52]) gave significantly higher score than leading personality (M = 1.72, 95 % CI [1.04, 2.40], $p = .025$), and helping type (M = 1.27, 95 % CI [0.57, 1.96], $p = .001$). Comparisons between the other groups were not statistically significant at $p < .05$.

The results indicated that the players with dependent personality tended to use significantly more argumentation, in other words, more cognitive words than leading and helping personality. In terms of LIWC component, Achievement, the competitive type (M = 2.04, 95 % CI [1.69, 2.40]) was significantly higher than leading personality (M = 0.29, 95 % CI [−0.19, 0.72], $p < .001$), dependent (M = 0.87, 95 % CI [0.41, 1.32], $p = .001$), and withdrawn (M = −0.45, 95 % CI [−1.75, 0.86], $p = .013$).

Moreover, helping (M = 1.51, 95 % CI [0.83, 2.20]) was significantly higher than leading personality (M = 0.29, 95 % CI [−0.19, 0.72], $p = .043$). These findings showed competitive personality tended to use significantly more achievement words compared to leading, dependent and withdrawn personality.

For LIWC component Negative Valence, withdrawn (M = 10.43, 95 % CI [4.23, 16.62]) was significantly higher than competitive (M = −0.55, 95 % CI

Table 5.10 ANOVA results of LIWC psychological features with personality type as the factor

Personality type	Groups	*df*	*F*	η	*p*
Argumentation	Between groups	5	5.116	0.047	0.000
	Within groups	521			
	Total	526			
Achievement	Between groups	5	7.261	0.065	0.000
	Within groups	521			
	Total	526			
Negative valence	Between groups	5	5.447	0.347	0.000
	Within groups	521			
	Total	526			

[−0.93, −0.17], $p = .022$), leading (M = −0.85, 95 % CI [−1.26, −0.45], $p = .018$), and dependent (M = 0.61, 95 % CI [−1.03, −0.19], $p = .021$); and helping (M = 6.54, 95 % CI [4.98, 8.09], was significantly higher than competitive (M = −0.55, 95 % CI [−0.93, −0.17], $p < .001$), leading (M = −0.85, 95 % CI [−1.26, −0.45], $p < .001$), dependent (M = −0.61, 95 % CI [−1.03, −0.19], $p < .001$), and aggressive (M = 1.32, 95 % CI −0.43, 3.07], $p < .001$).

The aforementioned findings showed that withdrawn and helping personality tended to express more negative emotions than competitive, leading, and dependent. Moreover, helping also used more negative emotion words than aggressive.

5.8 Conclusion and Future Research

In this chapter we have developed a dataset containing personality excerpts based on Leary's Rose Frameboard. By this, we have developed automatic personality detection that shows are more efficient than that of human judges. Consequently, we have presented three automated methods to personality detection, based on understanding from research in natural language processing, machine learning, and psychology characteristic.

We conducted that while text classification based on n-gram (UNIGRAMS, BIGRAMS) is the best particular detection approach, a combination-method such as LIWC and Subjective Lexicons features along with n-gram features can achieve better performance.

Eventually, we have done several notable contributions. Particularly, our results indicate to take into account both the context, such as BIGRAMS, rather than precisely using a global set of personality indications (e.g., LIWC and Subjective Lexicons). We have also reported results based on the feature weights that show the difficulties confronted by judges in annotating the dataset. Finally, we have found a possible connection between personality behavior by players, such "Helping and Aggressive" and "Dependent & Leading", based on BIGRAMSs and LIWC similarities.

For future work, we want to include an extended experiment of the methods pro-posed in current research to sentiment analysis, opinion mining, as well as emotion detection in other domains. Also, we want to extend the method in this work to apply in Big-Five personality detection. It will help us to not only detect the player's behaviors but also to detect introvert and extrovert players and a focus on approaches with POS features might be useful.

Acknowledgments This work was funded by the National Science Foundation (DRK-12-0918409). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of these funding agencies, cooperating institutions, or other individuals.

References

1. D'Mello, S., Onley, A., Person, N.: Mining collaborative patterns in tutorial dialogues. J. Educ. Data Mining. **2**(1), 1–37 (2010)
2. Toutanova, K., Klein, D., Manning, C., Singer, Y.: Feature-rich part-of-speech tagging with a cyclic dependency network. In: Proceedings of Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology, pp. 173–188. Association for Computational Linguistics, Stroudsburg (2003)
3. Wilcock, G.: Text annotation with OpenNLP and Uima. In: 17th Nordic Conference of Computational Linguistics, pp. 7–8. University of Southern Denmark, Odense (2009)
4. Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., Witten, I.H.: The WEKA data mining software: an update. SIGKDD Explor. Newsl. **11**(1), 10–18 (2009)
5. Vaassen, F., Daelemans, W.: Emotion classification in a serious game for training communication skills. In: 20th Meeting of Computational Linguistics in the Netherlands, pp. 155–168. Utrecht Institute of Linguistics, Utrecht (2010)
6. Mairesse, F., Walker, M., Mehl, M., Moore, R.: Using linguistic cues for the automatic recognition of personality in conversation and text. J. Artif. Intell. Res. **30**(1), 457–500 (2007)
7. Mairesse, F., Walker, M.: Automatic recognition of personality in conversation. In: Human Language Technology Conference of the North American Chapter of the ACL, pp. 85–88. Association for Computational Linguistics, Stroudsburg (2006)
8. Pennebaker, J.W., Chung, C.K., Ireland, M., Gonzales, A., Booth, R.J.: Linguistic Inquiry and Word Count (LIWC). Lawrence Erlbaum Associates, Mahwah (2007)
9. Keshtkar, F., Inkpen, D.: Using sentiment orientation features for mood classification in blogs. In: International Conference on Natural Language Processing and Knowledge Engineering, pp. 1–6. IEEE Press, New York (2009)
10. Gore, J., Kiefner, A., Combs, K.: Personality traits that predict academic citizenship behavior. J. Appl. Soc. Psychol. **42**(10), 2433–2456 (2012)
11. Song, H., Kwon, N.: The relationship between personality traits and information competency in Korean and American students. Int.J.Soc.Behav.Pers. **40**(7), 1153–1162 (2012)
12. Furnham, A., Chamorro-Premuzic, T., McDougall, F.: Learning style, personality traits and intelligence as predictors of college academic performance. Learn. Individ. Differ. **10**(3), 117–128 (2012)
13. Costa, P.T., McCrae, R.R.: Four ways five factors are basic. Pers. Individ. Differ. **13**(6), 653–665 (1992)
14. Read, S.J., Monroe, B.M., Brownstein, A.L., Yang, Y., Chopra, G., Miller, L.C.: A neural network model of the structure and dynamics of human personality. Psychol. Rev. **117**(1), 61–92 (2010)
15. Shaffer, D.W., Chesler, N., Arastoopour, G., D'Angelo, C.: Nephrotex: teaching first year students how to think like engineers. In: Course, Curriculum, and Laboratory Improvement PI Conference, Poster. NSF, Washington (2011)
16. Clark, M.H., Schroth, C.A.: Examining relationships between academic motivation and personality among college students. Learn Individ. Differ. **20**(1), 19–24 (2010)
17. Zhao, B.: Learning from errors: the role of context, emotion, and personality. J. Organ. Behav. **32**(3), 435–463 (2011)
18. Poropat, A.E.: A meta-analysis of the five-factor model of personality and academic performance. Psychol. Bull. **135**(2), 322–338 (2009)
19. Komarraju, M., Karau, S.J., Schmeck, R.R., Avdic, A.: The big five personality traits, learning styles, and academic achievement. Pers. Individ. Differ. **51**(4), 472–477 (2011)
20. Beaujean, A.A., Firmin, M.W., Attai, S., Johnson, C.B., Firmin, R.L., Mena, K.E.: Using personality and cognitive ability to predict academic achievement in a young adult sample. Pers. Individ. Differ. **51**(6), 709–714 (2011)
21. Leary, T.: The Interpersonal Diagnosis of Personality. Wiley, Hoboken (1957)

22. Bagley, E.A.S.: Stop talking and type: mentoring in a virtual and face-to-face environmental education environment. PhD thesis, University of Wisconsin-Madison (2011)
23. Shafer, D., Hatfield, D., Svarovsky, G., Nash, P., Nulty, A., Bagley, E., Franke, K., Rupp, A., Mislevy, R.: Epistemic network analysis: a prototype for 21st century assessment of learning. Int. J. Learn. Media **1**(2), 33–53 (2009)
24. D'Angelo, C., Arastoopour, G., Chesler, N., Shaffer, D.: Collaborating in a virtual engineering internship. In: Spada, H., Stahl, G., Miyake, N., Law, N. (eds.) Connecting Computer-Supported Collaborative Learning to Policy and Practice, pp. 626–630. International Society of the Learning Sciences, Hong Kong (2011)
25. Wiebe, J., Wilson, T.: Learning to disambiguate potentially subjective expressions. In: 6th Conference on Natural Language Learning, pp. 112–118. Association for Computational Linguistics, Stroudsburg (2002)
26. Esuli, A., Sebastian, F.: Sentiword-net: a publicly available lexical resource for opinion mining. In: 5th Conference on Language Resources and Evaluation, pp. 417–422. European Language Resources Association, Paris (2006)
27. Stone, P., Dunphy, D., Smith, M., Ogilvie, D.: The General Inquirer: A Computer Approach to Content Analysis. MIT Press, Cambridge (1966)
28. Taboada, M., Grieve, J.: Analyzing appraisal automatically. In: AAAI Spring Symposium on Exploring Attitude and Affect in Text: Theories and Applications, pp. 158–161. Association for the Advancement of Artificial Intelligence, Menlo Park (2004)
29. Cade, W.L., Lehman, B.A., Olney, A.: An exploration of off topic conversation. In: Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics, pp. 669–672. Association for Computational Linguistics, Stroudsburg (2010)
30. John, G.H., Langley, P.: Estimating continuous distributions in bayesian classifiers. In: Besnard, P., Hanks, S. (eds.) Eleventh Conference on Uncertainty in Artificial Intelligence, pp. 338–345. Morgan Kaufmann Publishers, San Francisco (1995)
31. Keerthi, S., Shevade, S., Bhattacharyya, C., Murthy, K.: Improvements to platt's SMO algorithm for SVM classifier design. Neural Comput. **13**(3), 637–649 (2001)
32. Li, H., Cai, Z., Graesser, A., Duan, Y.: A comparative study on English and chinese word uses with LIWC. In: Youngblood, G.M., McCarthy, P.M. (eds.) 25th International Florida Artificial Intelligence Research Society Conference, pp. 238–243. AAAI Press, Palo Alto (2012)