Constructing a Personality-Annotated Corpus for Educational Game Based on Leary's Rose Framework

Candice Burkett, Fazel Keshtkar, Arthur Graesser, and Haiying Li
Institute for Intelligent Systems
University of Memphis
Memphis, TN 38152
{cburkett, fkshtkar, a-graesser, hli5}@memphis.edu

Abstract

Researchers have recognized the importance of classifying personality through discourse for many years. However, this line of research tends to focus almost exclusively on the personality categories known as the Big Five factors. Though this information is certainly valuable, it may also be useful to categorize personality based on the Leary's Interpersonal Circumplex model which emphasizes a predictive function. In this paper we construct the data set for personality annotation among six dimensions (based on a coding scheme developed from Leary's Interpersonal Circumplex) for players using a chat interaction in an epistemic game, Land Science. Our results indicate that overall personality annotation is reliable (Average Kappa = 0.65) with the highest reliability for the competitive dimension and the lowest reliability for the leading dimension.

Introduction

There has been a great interest in classifying personality types in numerous fields spanning decades of research. In particular, the notion that personality characteristics are an important consideration in the field of education, has been re-emphasized in current literature. For example, personality characteristics have been shown to be related to both academic motivation and performance (Komarraju & Karau, 2005; Poropat, 2009). A considerable amount of research such as those previously mentioned is rooted in the Big Five personality factors (Norman, 1963). There is some debate, however, whether or not the Big Five factors adequately capture all dimensions of personality. For example, some research indicates that behavior may be better predicted based on specific personality traits, some of which are not accounted for in the Big Five factors (Ashton, 2001). A framework for personality classification that emphasizes the predictive value of personality traits is Leary's Interpersonal Circumplex (Leary, 1957).

This paper is organized in four sections. First, the paper includes a brief description of Leary's Rose framework and the Land Science Epistemic game. Section two explains the related work. Our model is described in section three. The results are described in the fourth section. Finally, we end the paper with Conclusion and Future work.

Leary's Interpersonal Circumplex

Leary's Interpersonal Circumplex (or Leary's Rose) has been used by researchers for decades as a foundation for

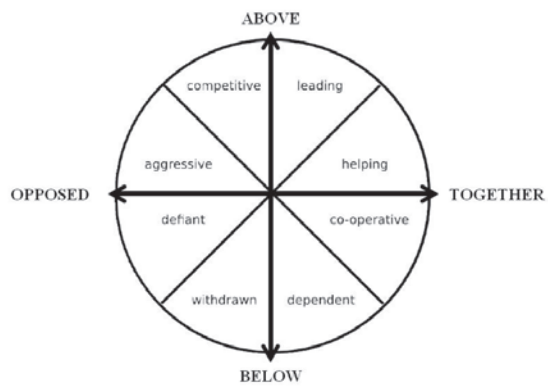

Figure 1. Leary's Interpersonal Circumplex

categorizing personality through the discourse (Leary, 1957) (See Figure 1).

The Circumplex 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 (opposed). The Rose can easily be separated into four quadrants and then further split into eight different categories (see Table 1 for examples).

An additional feature of Leary's Rose is that it allows not only for placement of a statement within the categories, but also allows for prediction of the category of the response statement. Specifically, along the above-below dimension if a statement falls within the above categories, the response will likely be in the below categories and vice-versa.

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.	Co-operative	Below-Together
What should I do now?	Dependent	Below-Together
Sorry, nevermind, I'm not thinking.	Withdrawn	Below-Opposed
No. I am not going to do that.	Defiant	Below-Opposed

Table 1. The dataset with some example statements from each of the eight personality categories of Leary's Rose.

Conversely, along the together-opposed dimension statements in the together categories are likely to evoke statements that are also within the together categories and likewise in the opposed categories.

Personality in Computer-Based Learning Environments

There are many reasons personality traits should be considered in computer-based learning environments (CBLEs). At a very basic level, attitudes toward computers can be related to personality types (Sigurdsson, 1991). Furthermore, it is well known that it is important to take individual differences into account during learning. Intelligent Tutoring Systems (ITS) are known for their ability to simulate effective human tutoring methods as well as take into account the individual needs of learners (Graesser, D'Mello & Cade, in press). In both human tutor and ITS, it is hard to accurately assess both the cognitive and emotional states of individual learners (D'Mello, Craig & Graesser, 2009; Graesser, D'Mello, & Person, 2009). However, it is a rather complex process to categorize personality traits solely from natural language user input in a CBLE.

Researchers have had some success on the *deLearyous* gaming project (Vaassen & Daelemans, 2010). However, the *deLearyous* project is a serious game intended to improve communication skills that classifies personality traits based on Leary's Rose. Vaassen and Daelemans were able to successfully classify approximately half of the sentences in their corpus into one of the four quadrants of Leary's Rose. The researchers noted, however, that the manually annotated sentences used to compile their training set were labeled by only one human annotator and thus may have been susceptible to issues with reliability. Accordingly, the goal of the current study is reliably recognize personality traits from discourse (group discussions) exchanged in the chat function of a serious epistemic game (Land Science). A secondary goal was to see which personality categories were most prevalent over the course of the game.

Land Science Epistemic Game

Land Science is an epistemic game created by researchers at the University of Wisconsin-Madison that has been designed to simulate a regional planning practicum experience for students (Bagley, 2011; D'Angelo, Arastoopour, Chesler, & Shaffer, 2011; Shaffer, C.N.A.G., & D'Angelo, 2011). During the 10 hours game, students play the role of interns at a fictitious regional planning firm where they make land use decisions in order to meet the desires of virtual stakeholders. Students are split into groups and progress through a total of 15 stages of the game. Some of the stages of the game were intended to be more task-oriented (e.g., creating a fictitious staff page) while others are intended to be more reflective (e.g., asking participants to reflect upon what they have learned). Throughout the game players communicate with other members of their planning team as well as a mentor through the use of a chat feature that is embedded within the interface.

Method

Participants and Data Set

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 communicate with both other players and mentors using a chat feature embedded in the interface. For the purposes of these analyses we only assessed the discourse of the players. The researchers selected 1,000 player excerpts to be analyzed (average length = 4.8 words). For our purposes, an excerpt is defined as a turn of speech that was taken by a student. In other words, one excerpt occurred every time a student typed something and clicked "send" or hit "enter" in the chat interface. The excerpts are selected from a larger set of 3,227. In order to proportionally represent all 15 stages of the game in the set that is analyzed, approximately 31% of the player excerpts are randomly selected from each stage.

Procedure

Annotation is done using a coding scheme that we developed based on the Timothy Leary's Interpersonal Circumplex Model (Leary, 1957). For the purposes of this coding, we combined the Helping and Co-operative categories and also the Aggressive and Defiant are categorized into one. Therefore, six categories of Leary's Rose are used for the coding scheme. Statements that do not fit into any of the categories were coded as neutral, indicating that there is no evidence of any of the six categories present.

Two graduate students, as human judges, completed two cycles of annotation training. The first cycle required the annotators to independently code 200 excerpts that were selected randomly. The *kappa* statistic was computed to assess inter-rater reliability on this set and achieve fair agreement (0.33). Following this, the annotators discuss and refine any issues regarding the coding scheme. Then, they annotated a new set of 1,000 excerpts that were randomly selected. For this set, Agreement on the second training set is substantial with 0.70. Results indicated sufficient reliability and thus completed the training of the annotators. Once the two annotators were trained they independently annotated a set of 1,000 excerpts described in the data set portion of this paper.

Results

Overall, personality category agreement between the two annotators is substantial (Average Kappa = 0.65). As shown in Table 2, agreement is substantial for the Competitive, Dependent and Withdrawn categories, and is moderate for the Leading, Helping/Co-operative, Aggressive/Defiant and Neutral categories.

Personality Category	Kappa
Competitive	0.80
Leading	0.54
Dependent	0.77
Withdrawn	0.72
Helping/Co-operative	0.59
Aggressive/Defiant	0.55
Neutral	0.57
Overall Average	0.65

Table 2. Kappa statistics for each of the six personality categories annotated as well as neutral statements for which there is no indication of any of the six categories.

The two human annotators agreed on the personality category for a total number of 737 excerpts. Of those agreements, the largest percentage of excerpts is Neutral (39.30%). Regarding excerpts for which there is an agreement that a personality category present, the largest percentage is Dependent (22.49%) followed by Competitive (17.48%) and Helping/Co-operative (7.99%). The least represented personality categories are Leading (7.05%), Withdrawn (3.25%) and Aggressive/Defiant (2.44%) (See Figure 2).

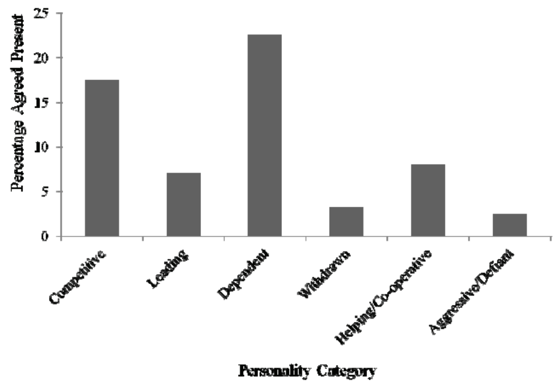

Figure 2. Percentage of statements (Agreements) for which personality categories are presented.

It is important to note that these categories are collapsed over the entirety of the game, though there are a number of different tasks and discussions that took place. However, it is likely that different personality characteristics would present themselves at different points based on what is happening in the game. In fact, there is some evidence for this based on a series of Pearson correlations (See Table 3). For example, participants competitiveness is related to which group, which phase of the game and what type of game stages they are working on (reflective vs. task-oriented).

More specifically, as the game progresses, the percentage of competitive statements increases, the Withdrawn statements and the Helping/Co-operative statements decrease. Also, the percentages of both Competitive and Leading statements are related to which group participants are in. Finally, the percentage of Competitive statements decreases during tasks that are intended to require increased reflection.

Personality Category	Group	Game Order	Stage Type
Competitive	-0.12**	0.16**	-0.07*
Leading	0.10**	-0.07	0.01
Dependent	0.04	-0.04	-0.02
Withdrawn	0.05	-0.09*	0.04
Helping/ Co-operative	0.03	-0.08*	0.02
Aggressive/ Defiant	-0.00	-0.06	0.06

* $p < .05$. ** $p < .01$.

Table 3. Correlations for six personality categories and group, game order (indicating the order of stages that were completed during the game) and type of stage (either reflective or task-oriented).

Conclusion and Future Work

The goal of this study was to reliably recognize personality traits from discourse (group discussions) in the chat function of a serious epistemic game (Land Science). The second goal is to observe which personality categories are most prevalent over the course of the game. Regarding the first goal, two independent raters annotated a corpus of 1,000 excerpts with an average *Kappa* of 0.65. Therefore, it is possible to reliably annotate discourse from an epistemic based computer game. However, it is important to note that some personality categories have greater reliability than others. These data are important to keep in mind when trying to determine the degree to which personality categories can be successfully classified automatically. Regarding the second goal, it is interesting to note that overall the Dependent and Competitive categories are the most prevalent whereas the Withdrawn and Aggressive/Defiant categories are the least prevalent. Moreover, in this research, we achieve some scientific notable factors to take into account when determining which personality traits tend to emerge among groups under what conditions (e.g., group and phase of the game).

For future work, we aim to use this data set as valuable corpus for automatic personality detection by using Natural Language Processing (NLP) and Machine Learning techniques. This future research should further refine the automatic classification of personality characteristics. In addition, we plan to compare these results to other findings obtained using a framework based on the Big Five personality factors. Ultimately, results from this line of research will be extremely useful in developing a tutoring system that is sensitive to the personality characteristics that emerge during interactions with both individuals as well as groups.

Acknowledgements

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

Bagley, E.A.S. (2011). Stop talking and type: Mentoring in a virtual and face-to-face environmental education environment. Ph.D. thesis, University of Wisconsin-Madison.

D'Angelo, C., Arastoopour, G., Chesler, N., & Shaffer, D. (2011, July) Collaborating in a virtual engineering internship. Computer Supported Collaborative Learning Conference (CSCL), Hong Kong, Hong Kong SAR.

D'Mello, S., Craig, S., & Graesser, A. (2009). Multi-method assessment of affective experience and expression during deep learning, International Journal of Learning Technology, 4(3), 165-187.

Graesser, A. C., D'Mello, S. K., & Cade, W. (in press). Instruction based on tutoring. In R. E. Mayer and P. A. Alexander (Eds.), Handbook of research on learning and instruction. New York: Routledge Press.

Graesser, A. C., D'Mello, S. K., & Person, N. (2009). Metaknowledge in tutoring. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.). Metacognition in educational theory and practice. Mahwah, NJ: Erlbaum.

Komarraju, M., & Karau, S. (2005). The relationship between the big five personality traits and academic motivation. Personality and Individual Differences, 39(3), 557-67.

Leary, T. (1957). Interpersonal diagnosis of personality. New York: Ronald Press.

Norman, W. (1963). Toward an adequate taxonomy of personality attributes: Replicated factor structure in peer nomination personality ratings. Journl of Abnormal and Social Psychology, 6, 574-583.

Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. Psychological Bulletin, 135(2), 322-38.

Shafer, D.W., C.N.A.G., & D'Angelo, C. (2011, January). Nephrotex: Teaching first year students how to think like engineers. Laboratory Improvement (CCLI) PI Conference, Washington, D.C.

Sigurdsson, J. F., (1991). Computer experience, attitudes toward computers and personality characteristics in psychology undergraduates. Personality and Individual Differences, 12(6), 617-24.

Vaassen, F., & Daelemans, W. (2010). Emotion classifications in a serious game for training communication skills.Computational Linguistics in the Netherlands. Utrecht, The Netherlands.