Choosing to Interact: Exploring the Relationship Between Learner Personality, Attitudes, and Tutorial Dialogue Participation

Aysu Ezen-Can
Department of Computer Science
North Carolina State University
aezen@ncsu.edu

Kristy Elizabeth Boyer
Department of Computer Science
North Carolina State University
keboyer@ncsu.edu

ABSTRACT
The tremendous effectiveness of intelligent tutoring systems is due in large part to their interactivity. However, when learners are free to choose the extent to which they interact with a tutoring system, not all learners do so actively. This paper examines a study with a natural language tutorial dialogue system for computer science, in which students interacted with the JavaTutor system through natural language dialogue over the course of problem solving. We explore the relationship between students' level of dialogue interaction and learner characteristics including personality profile and pre-existing attitudes toward the learning task. The results show that these learner characteristics are significant predictors of the extent to which students engage in dialogue with the tutoring system, as well as the number of task actions students make. By identifying students who may not engage with tutoring systems as readily, this work constitutes a step toward building adaptive systems that successfully support a variety of students with different attitudes and personalities.

Keywords
Learner characteristics, personality, disengagement, tutorial dialogue

1. INTRODUCTION
Tutorial dialogue systems effectively support learning through rich natural language dialogue [7,8,14,19]. However, the effectiveness of tutorial dialogue systems, like other adaptive learning environments, depends in large part on students' willingness to interact with them [18]. Interaction varies tremendously across individual students and student populations. We observe various types of *disengagement* including lack of motivation or interest for the learning task [10], as well as *gaming* an intelligent tutor by exploiting properties of the learning environment [2, 4].

In addition to these factors, individual differences such as self-reported interest in the task and confidence in learning have been found to be strong predictors of engagement [6]. Similarly, students' hidden attitudes toward learning [1] and motivation for the task [3] may be highly influential. Boredom, which is associated with reduced motivation to perform the activity [15], has been positively correlated with attention problems and negatively correlated with performance. Students' participation in tutorial dialogue has also been found to be associated with the students' expectations [11], and in human-human tutorial dialogue, student personality traits have recently been found to be significant factors [16]. However, the field is far from a full understanding of the factors that influence students' choices to engage or interact with tutorial dialogue systems.

This paper presents an investigation into the relationship between student characteristics and interactions with a tutorial dialogue system. We hypothesized that students' personality profile, for example their tendencies toward extraversion or openness, would be significantly associated with the level of natural language interaction observed within a tutorial dialogue system. We also hypothesized that students' attitudes toward the learning task would be a significant factor in their interactions with the system. We examine these hypotheses within a data set of 51 university students interacting with the JavaTutor tutorial dialogue system for introductory computer science. Regression models were built that predict both dialogue and task participation by the students, who have the choice to interact with the dialogue system as little or as much as desired over the course of the learning tasks. The models demonstrate that students' attitudes and personalities are significantly predictive of their willingness to interact with the tutorial dialogue system. The findings suggest that some learner characteristics may put students at risk of low participation with a tutorial dialogue system, and constitute a first step toward proactively adapting the systems to benefit these learners.

2. TUTORING STUDY
The JavaTutor tutorial dialogue system (Figure 1) supports students in solving introductory computer programming problems in the Java programming language while interacting in textual natural language. Students are provided with a series of learning tasks that build on each other to guide the students through creation of a simple text-based adventure game.[1]

The study reported here was conducted with the JavaTutor tutorial dialogue system in 2014. The students (12 female; 39 male; mean age = 21) were drawn from a university-level engineering class. They interacted with the tutorial dialogue system for one session lasting approximately 45 minutes.

[1]Implementation details of the system are beyond the scope of this paper but are described in [9].

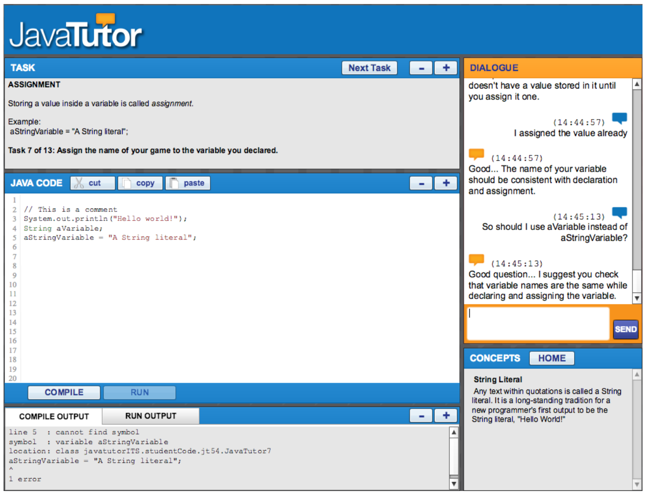

Figure 1: Screenshot from the tutorial dialogue system.

Prior to interacting with JavaTutor, students took a pre-survey that included validated items to measure goal orientation [17], general self-efficacy [5], confidence in learning computer science and programming [13], and personality profile using a concise version of the Big Five Inventory [12]. Students also completed a pre-test and posttest before and after their interaction with JavaTutor.

3. ANALYSIS

Students were instructed that they could make comments, pose questions, and request feedback at any time through textual dialogue. Overall, students interacting with JavaTutor achieved significant learning gains from pre-test to posttest (average= 12%, median= 13.4%, stdev = 32%, $p = 0.001$). However, we observed that 58.8% of students never made an utterance. For students who did engage in dialogue with the tutor, the average number of utterances was 5.1 (stdev=7.36, median= 2). Regardless of the extent to which they chose to engage in natural language dialogue, all students received some tutorial dialogue utterances based upon the system's model of feedback for task events.

Our goal is to identify the factors that may be influential in students' levels of interaction with the system. To this end, we built multiple regression models. The remainder of this section describes the analysis.

3.1 Response Variables

Based upon the logged interaction traces, we extracted dialogue and task events and used them to compute a numeric representation of the student's level of interaction with the system. For dialogue interaction we utilized the *number of utterances* written by each student. The range of number of student utterances was between 0 and 33.

We extracted four features that represent interaction of students with the system throughout tutoring. The first of these four features is *number of content changes* which refers to the changes in the student's programming code, as the code they write is referred to as content pane. We also computed *the number of compile events* and *number of run activities*. The number of compile/run events ranges from 4 to 224, whereas the number of content changes ranges from 88 to 1099 to complete the series of learning tasks.

Finally, we computed the *number of tutor messages* each student received. The tutoring systems provided students with feedback. The number of messages received is closely related to the number of actions that triggered tutor feedback, which is also a measure of participation. The minimum number of tutor messages provided to any student was 8, whereas the largest number of tutor messages to a student during a tutoring session was 121. We built separate multiple regression models to predict level of dialogue interaction and level of task interaction.

3.2 Predictor Variables

We hypothesized that several learner characteristics were significantly associated with level of interaction in the system. We provided these variables for selection within the models (see Table 1). All of the predictors were standardized to a common scale before model building.

Predictor variable	Example survey item/ Description
Computer science confidence	*I am sure that I can learn programming.*
Perceived computer science usefulness	*I'll need programming for my future work.*
Motivation toward computer science	*Programming is enjoyable and stimulating to me.*
General self-efficacy	*I will be able to achieve most of the goals that I have set for myself.*
Learning goal orientation	*I often look for opportunities to develop new skills and knowledge.*
Performance demonstration	*I like to show that I can perform better than my coworkers.*
Failure avoidance	*Avoiding a show of low ability is more important to me than learning a new skill.*
Achievement goals	*It is important for me to do better than other students.*
Gender	Male/female
Age	Age of the student
University class standing	The year that the student is in the university
Perception of student's own computer skill	*How skilled are you with computers, compared to the average person?*
Extraversion	*I see myself as someone who is talkative.*
Agreeableness	*I see myself as someone who is helpful and unselfish with others.*
Conscientiousness	*I see myself as someone who does a thorough job.*
Neuroticism	*I see myself as someone who is depressed, blue.*
Openness	*I see myself as someone who is original, comes up with new ideas.*
Pre-test score	Score showing the performance of the student before tutoring session

Table 1: Predictor variables from pre-survey and pre-test.

3.3 Modeling Level of Participation

We built separate models for each of the response variables (number of utterances, compile/run events, content changes, received tutor messages). For each response variable we used the whole dataset and selected features via stepwise linear regression. Because the goal was to investigate relationships between pre-measures (student characteristics, attitudes) and level of participation, we conducted descriptive analyses using the entire data set for model building.

The model for number of dialogue utterances (Table 2) revealed that students' failure avoidance characteristic is a significant predictor of tutorial dialogue interactivity. Students who indicated that they tend to avoid tasks in which they may have higher chance of failure wrote fewer utterances to the system.

Number of utterances =	**Coefficient**	***p***
Failure Avoidance	-0.3089	0.0274
~1 (intercept)		1
RMSE = 0.961		
R^2 = 0.0954		

Table 2: Stepwise linear regression model for the number of utterances.

The model for number of compile/run events during tutoring session showed that students' personality scores, particularly the binary agreeableness score, was a significant predictor of participation from a task-related perspective. The students who were more agreeable (indicated as a 1 for the model, rather than a 0) made more task interactions considering compile/run events as shown in Table 3. The other regression model having the number of content changes as a response variable did not produce significant results.

Number of compile/run =	**Coefficient**	***p***
Agreeableness (binarized)	0.2897	0.0392
~1 (intercept)		1
RMSE = 0.967		
R^2 = 0.0839		

Table 3: Stepwise linear regression model for number of compile/run events.

Another regression model that showed significant results was the regression model that predicted the number of tutor messages students received. Interestingly, both student perceptions (computer science confidence and motivation) and personality (openness score from Big Five Inventory) were selected by the model as shown in Table 4. There was a negative correlation between computer science confidence and tutor messages, however it was the opposite for computer science motivation. The students who were more motivated to study computer science interacted more with the system, triggering more tutor messages. Also, the students who had low confidence towards programming received less tutor feedback. Figure 2 shows the scatter plots for both computer science motivation and confidence measures.

Discussion. Understanding how student characteristics are associated with tutorial dialogue interaction holds great promise for identifying possible disengagement types and taking adaptive action during tutoring sessions to further improve learning effectiveness. The results of the models indicate that as hypothesized, student characteristics such as personality profile were significantly predictive of the student's level of interactivity with the tutorial dialogue system. We found that students' attitudes and personalities have significant relationships with their level of participation in terms of

Number of tutor messages =	**Coefficient**	***p***
Age	0.3802	0.0033
Computer science confidence * Openness	-0.5244	0.0008
Computer science motivation Openness	0.5317	0.00006
~1 (intercept)		1
RMSE = 0.739		
R^2 = 0.52		

Table 4: Stepwise linear regression model for number of tutor messages received.

both dialogue and task.

Another important finding was that although pre-test was present in all regression models as an independent variable, it was not significantly predictive of either the number of utterances or the task activities. In other words, the level of participation was more correlated to student characteristics than to their incoming knowledge. These results are important for understanding how to better foster interaction with intelligent tutoring systems. If we can identify students who tend to participate less or become disengaged, the system can automatically adapt to these students with scaffolding. For instance, when a student with low motivation toward the task is identified, the tutorial dialogue system might put particular emphasis on moves that are part of "adjacency pairs," such as asking a question and awaiting a student response. Adapting the task may also be appropriate in these cases. By utilizing information that we can glean from quick pre-measures, we may be able to significantly improve the effectiveness of the system.

4. CONCLUSION

Adapting to broader populations with varying characteristics is crucial for increasing the use of intelligent tutoring systems and making them more effective. A central challenge is determining the factors that might affect level of participation with intelligent systems. The current literature is far from totally understanding underlying relationships between student characteristics and how they affect system interactions during tutoring. The findings presented here have identified student characteristics such as level of failure avoidance which are particularly strongly associated with low interaction.

Several directions of future work are promising. First, incorporating multiple sources of information such as multimodal features (e.g., posture, gesture, eye gaze) can help us better understand students and respond in real time to engage them in more interactions. Each of these types of features has been shown to contribute to modeling student behavior. Additionally, customizing scaffolding to different learner characteristics is very promising. Modifying the realized utterances delivered to students based on their personality style, gender, and skill are likely to improve interactions with the system. It is important to devise and investigate strategies for learners of all characteristics in order to better engage students and help them learn more.

Acknowledgments The authors wish to thank the members of the LearnDialogue group at North Carolina State University for their helpful input. This work is supported in part by the Department of Computer Science at North Carolina State University and the National Science Foundation through Grants IIS-1409639, and the

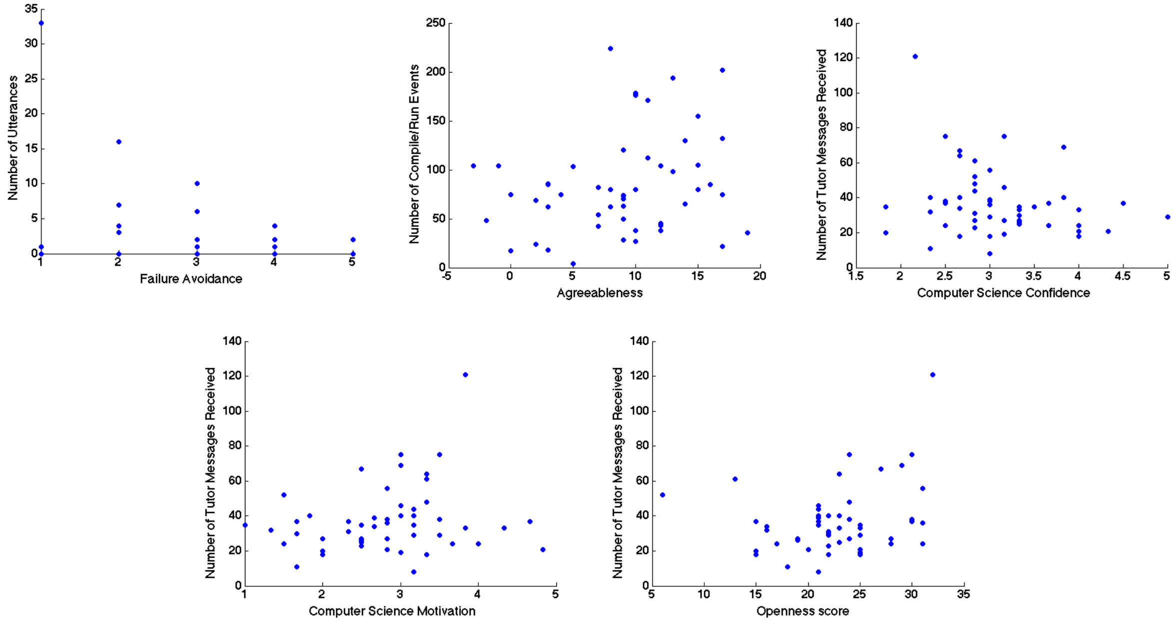

Figure 2: Scatter plots of various predictors and response variables.

STARS Alliance, CNS-1042468. Any opinions, findings, conclusions, or recommendations expressed in this report are those of the authors, and do not necessarily represent the official views, opinions, or policy of the National Science Foundation.

5. REFERENCES

[1] I. Arroyo and B. P. Woolf. Inferring learning and attitudes from a bayesian network of log file data. In *Proceedings of AIED*, pages 33–40, 2005.

[2] R. S. Baker, A. de Carvalho, J. Raspat, V. Aleven, A. T. Corbett, and K. R. Koedinger. Educational software features that encourage and discourage “gaming the system”. In *Proceedings of AIED*, pages 475–482, 2009.

[3] C. R. Beal, L. Qu, and H. Lee. Mathematics motivation and achievement as predictors of high school students’ guessing and help-seeking with instructional software. *Journal of Computer Assisted Learning*, 24(6):507–514, 2008.

[4] J. E. Beck. Engagement tracing: Using response times to model student disengagement. In *Proceedings of AIED*, pages 88–95, 2005.

[5] G. Chen, S. M. Gully, and D. Eden. Validation of a new general self-efficacy scale. *Organizational research methods*, 4(1):62–83, 2001.

[6] S. D’Mello, C. Williams, P. Hays, and A. Olney. Individual differences as predictors of learning and engagement. In *Proceedings of the Annual Meeting of the Cognitive Science Society*, pages 308–313, 2009.

[7] S. K. D’Mello, B. Lehman, and A. Graesser. A motivationally supportive affect-sensitive AutoTutor. In *New perspectives on affect and learning technologies*, pages 113–126. 2011.

[8] M. Dzikovska, N. Steinhauser, E. Farrow, J. Moore, and G. Campbell. BEETLE II: Deep natural language understanding and automatic feedback generation for intelligent tutoring in basic electricity and electronics. *IJAIED*, 24(3):284–332, 2014.

[9] A. Ezen-Can and K. E. Boyer. A tutorial dialogue system for real-time evaluation of unsupervised dialogue act classifiers: Exploring system outcomes. *To appear*.

[10] K. Forbes-Riley and D. Litman. When does disengagement correlate with performance in spoken dialog computer tutoring? *IJAIED*, 22(2):19–41, 2008.

[11] G. T. Jackson, A. C. Graesser, and D. S. McNamara. What students expect may have more impact than what they know or feel. In *Proceedings of AIED*, pages 73–80, 2009.

[12] O. P. John, L. P. Naumann, and C. J. Soto. Paradigm shift to the integrative Big Five trait taxonomy. *Handbook of personality: Theory and research*, 3:114–158, 2008.

[13] C. Lee and P. Bobko. Self-efficacy beliefs: Comparison of five measures. *Journal of Applied Psychology*, 79(3):364, 1994.

[14] D. Litman and S. Silliman. ITSPOKE : An Intelligent Tutoring Spoken Dialogue System. In *Demonstration Papers at HLT-NAACL 2004*, pages 5–8, 2004.

[15] R. Pekrun, T. Goetz, L. M. Daniels, R. H. Stupnisky, and R. P. Perry. Boredom in achievement settings: Exploring control–value antecedents and performance outcomes of a neglected emotion. *Journal of Educational Psychology*, 102(3):531, 2010.

[16] A. K. Vail and K. E. Boyer. Adapting to Personality Over Time: Examining the Effectiveness of Dialogue Policy Progressions in Task-Oriented Interaction. In *Proceedings of the Annual SIGDIAL Meeting*, pages 41–50, 2014.

[17] D. VandeWalle, W. L. Cron, and J. W. Slocum Jr. The role of goal orientation following performance feedback. *Journal of Applied Psychology*, 86(4):629, 2001.

[18] K. VanLehn, A. C. Graesser, G. T. Jackson, P. Jordan, A. Olney, and C. P. Rosé. When are tutorial dialogues more effective than reading? *Cognitive Science*, 31(1):3–62, 2007.

[19] K. VanLehn, P. W. Jordan, C. P. Rosé, D. Bhembe, M. Böttner, A. Gaydos, M. Makatchev, U. Pappuswamy, M. Ringenberg, A. Roque, et al. The architecture of Why2-Atlas: A coach for qualitative physics essay writing. In *Proceedings of ITS*, pages 158–167, 2002.