Who Do You Think I Am? Modeling Individual Differences for More Adaptive and Effective Instruction

Laura K. Allen
Arizona State University
Tempe, AZ, 85283
LauraKAllen@asu.edu

ABSTRACT
The purpose of intelligent tutoring systems is to provide students with personalized instruction and feedback. The focus of these systems typically rests in the adaptability of the feedback provided to students, which relies on automated assessments of performance in the system. A large focus of my previous work has been to determine how natural language processing (NLP)_ techniques can be used to model individual differences based on students' natural language input. My proposed research will build on this work by using NLP techniques to develop stealth assessments of students' individual differences and to provide more fine-grained information about the cognitive processes in which these students are engaged throughout the learning task. Ultimately, my aim will be to combine this linguistic data with on-line system data in order to develop more robust student models within ITSs for ill-defined domains.

Keywords
Intelligent Tutoring Systems, Natural Language Processing, Writing, Feedback, User Models

1. INTRODUCTION
The purpose of intelligent tutoring systems (ITSs) is to provide students with personalized instruction and feedback based on their performance, as well as other relevant individual characteristics [1]. The focus of these systems typically rests in the adaptability of the feedback provided to student users, which relies on automated assessments of students' performance in the system. Despite this adaptive feedback, however, many ITSs lack the ability to provide adaptive *instruction* and *higher-level feedback*, particularly when providing tutoring for ill-defined domains. This shortcoming is largely due to the increased difficulties associated with accurately and reliably assessing student characteristics and performance when the learning tasks are not "clear cut." In mathematics tutors, for instance, it can be relatively straightforward to determine when a student is struggling in specific areas; thus, these systems can provide adaptive instruction and feedback accordingly. For ITSs focused on ill-defined domains (such as writing and reading), on the other hand, this process can be more complicated. In particular, students' *open-ended* and *natural language* responses to these systems present unique assessment challenges. Rather than identifying a set of "correct" answers, the system must identify and analyze characteristics related to students' responses in order to determine the quality of their performance as well as the areas in which they are struggling.

Natural language processing (NLP) techniques have been proposed as a means to target this assessment problem in adaptive systems. In particular, NLP provides detailed information about the characteristics of students' natural language responses within these systems [2] and subsequently helps to model students' particular areas of strengths and weaknesses [3]. NLP has begun to be incorporated within ITSs more frequently [4-5] because it allows systems to automatically evaluate the quality and content of students' responses [6-7]. Additionally, these assessments afford systems the opportunity to model students' learning throughout training and subsequently improve models of their performance [8]. Previous research suggests that these NLP techniques can increase the efficacy of computer-based learning systems. In particular, NLP helps to promote greater interactivity in the system and, consequently, leads to increased learning gains when compared to non-interactive training tasks (e.g., reading books, watching videos, listening to lectures [5, 9].

In my previous research, my colleagues and I have proposed that NLP techniques can be used to determine much more than simply the *quality* of a particular response in the system. Specifically, NLP can serve as a powerful methodology for modeling individual differences among students, as well as for examining the specific processes in which these students are engaging [3, 8]. In this overview, I suggest that, when combined with *on-line* interaction data, these NLP techniques can provide critical information that can be used to enhance the adaptability of ITSs, particular those focused on ill-defined domains. Thus, the aim of my research is to investigate how the linguistic characteristics of students' language can provide a window into their cognitive and affective processes. This information will then be combined with system data to promote more personalized learning experiences for the student users in these systems.

1.1 Writing Pal
The Writing Pal (W-Pal) is a tutoring system that was designed for the purpose of increasing students' writing proficiency through explicit strategy instruction, deliberate practice, and automated feedback [10]. In the W-Pal system, students are provided explicit

strategy instruction and deliberate practice throughout eight instructional modules, which contain strategy lesson videos and educational mini-games. The instruction in these modules covers specific topics in the three main phases of the writing process—prewriting (*Freewriting*, *Planning*), drafting (*Introduction Building*, *Body Building*, *Conclusion Building*), and revising (*Paraphrasing*, *Cohesion Building*, *Revising*).

Animated pedagogical agents narrate the W-Pal lesson videos by providing explicit descriptions of the strategies and examples of how these strategies can be used while writing (see Figure 1 for screenshots). The content covered in these videos can be practiced in one or more of the mini-games contained within each module. The purpose of these mini-games is to offer students the opportunity to practice the individual writing strategies without having to compose an entire essay.

Figure 1. Screenshots of the W-Pal Lesson Videos

W-Pal contains an AWE component in addition to the eight instructional modules, where students can practice holistic essay writing. This component of W-Pal contains a word processor where students can compose essays and automatically receive summative (i.e., holistic scores) and formative (i.e., actionable, strategy-based) feedback on these essays. The *summative* feedback in W-Pal is calculated using the *W-Pal assessment algorithm*. This algorithm employs linguistic indices from multiple NLP tools to assign essays a score from 1 to 6 (for more information, see 11). The purpose of the *formative* feedback is to teach students about high-quality writing and to provide them with actionable strategies for improving their essays. To deliver this feedback, W-Pal first identifies weaknesses in students' essays (e.g., essays are too short; essays are unorganized). It then provides students with feedback messages that designate specific strategies that can help them to work on the problems. Previous studies have demonstrated that W-Pal is effective at promoting increases in students' essay scores over the course of multiple training sessions [6; 12].

1.2 Current Work

The focus of my doctoral research will be on the use of NLP techniques to develop stealth assessments of students' individual differences and to provide more fine-grained information about the cognitive processes in which these students are engaged throughout the learning task. Ultimately, the aim of this research will be to combine this linguistic data with on-line system data in order to develop more robust student models within ITSs for ill-defined domains, such as W-Pal.

The goal of this specific research project will be to use the linguistic properties of students' essays to model individual differences related to writing performance (e.g., vocabulary knowledge). This data will then be combined with *on-line* process data, such as students' keystrokes while writing, to provide a more complete understanding of their writing processes. Ultimately, this project will aim to determine whether there are specific writing processes (as identified by the *characteristics* of the essays and students' *on-line processes*) that are more or less predictive of successful writing and revision. My final goal will then be to use this information to provide more adaptable instruction and formative feedback to students.

2. Proposed Contributions of Current Work

This proposed research project will contribute to both the W-Pal system, as well as the EDM community more generally. Regarding the W-Pal system, the development of stealth assessments and online student models will significantly enhance the adaptability and, theoretically, the efficacy of the system. The current version of W-Pal does not provide individualized instruction to students and only adapts the feedback based on single (i.e., isolated) essays that they generate. Thus, the system does not consider students' previous interactions with the system when providing feedback, nor the individual characteristics of these student users. Therefore, the proposed work will help to provide a much more robust student model, which should help W-Pal provide more personalized instruction and feedback.

More generally, the results of this project (and future projects) will contribute to the EDM community, as well as to research with natural language data more broadly. Language is pervasive and, here, we propose that it can be used to provide *unique* information about individuals' behaviors, cognitive processes, and affect. By investigating the specific characteristics of students' natural language data, we can glean important insights about their learning processes, beyond information that can be extracted from system log data. By combining NLP with other forms of data, researchers will gain a more complete picture of the students using the system, which should ultimately lead to more effective instruction.

3. Previous Work

A large focus of my previous work has been to determine how NLP techniques can be used to model individual differences based on students' natural language input. Importantly, this input has ranged from more structured language (such as essays) to naturalistic language responses (such as self-explanations). As an example, in one study, my colleagues and I investigated whether we could leverage NLP tools to develop models of students' comprehension ability based on the linguistic properties of their self-explanations [3]. Students (n = 126) interacted with a reading comprehension tutor where they self-explained target sentences from science texts. Coh-Metrix [13] was then used to calculate the linguistic properties of these aggregated self-explanations. The results of this study indicated that the linguistic indices were predictive of students' reading comprehension ability, over and above the current system algorithms (i.e., the self-explanation scores). These results are important, because they suggest that NLP techniques can inform stealth assessments and help to improve student models within ITSs.

In further research projects, we have begun to investigate how these linguistic characteristics change across time, and how these changes relate to individual differences among the students [14].

In particular, we proposed that the *flexibility* of students' writing style could provide important information about their writing proficiency. In one study, we investigated college students' (n = 45) flexibility in their use of cohesion across 16 essays and whether this flexibility related to their writing proficiency. The results suggested that more proficient writers were, indeed, more flexible in their use of cohesion across different writing prompts and that this cohesive flexibility was most strongly related to the unity, or coherence, of students' writing. The results of this study indicated that students might differentially employ specific linguistic devices in different situations in order to achieve coherence among their ideas. Overall, the results of these (and many other) studies provide preliminary evidence that NLP techniques can be used to provide unique information about students' individual differences and learning processes within ITSs.

4. Advice Sought

I am seeking advice for my proposed research regarding two primary questions. First, *what analytical methods should be used to most effectively model individual differences based on linguistic data?* In previous research, my colleagues and I have relied heavily on stepwise regression and discriminant function analysis techniques to model students' essay scores and individual differences. However, this technique can pose particular problems and is not always the most effective regarding large-scale data sets containing many variables, such as these. Thus, I would largely benefit from expert advice regarding the specific modeling techniques that can help to improve this research.

My second question relates to: *what on-line process data can be most effectively tied with this linguistic data – and how?* In previous studies, we have heavily relied on the linguistic properties of students' responses alone to model and understand the learning process. However, these models could be greatly strengthened through the addition of on-line processing data, such as keystrokes or eye tracking. We have begun to implement keystroke logging into the W-Pal system to begin to investigate this question. However, I would greatly benefit from expert advice regarding the best methods for combining this data into a reliable and accurate student model.

5. ACKNOWLEDGMENTS

Prior research was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305A080589 to Arizona State University. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education. I would also like to thank Danielle McNamara, Scott Crossley, Erica Snow, Jennifer Weston, Rod Roscoe, and Matthew Jacovina for their contributions to this line of work.

6. REFERENCES

[1] Murray, T. 1999. Authoring intelligent tutoring systems: An analysis of the state of the art. International *Journal of Artificial Intelligence in Education,* 10, (1999) 98-129.

[2] Crossley, S. A., Allen, L. K., and McNamara, D. S. 2014. Analyzing discourse processing using a simple natural language processing tool (SiNLP). *Discourse Processes,* 51, (2014) 511-534.

[3] Allen, L. K., Snow, E. L., and McNamara, D. S. in press. Are you reading my mind? Modeling students' reading comprehension skills with natural language processing techniques. *Proceedings of the 5th International Learning Analytics and Knowledge Conference*.

[4] Graesser, A. C., Lu, S., Jackson, G. T., Mitchell, H. H., Ventura, M., Olney, A. and Louwerse, M. 2004. AutoTutor: A tutor with dialogue in natural language. *Behavior Research Methods, 36*, (2004) 180-193.

[5] VanLehn, K., Graesser, A. C., Jackson, G. T., Jordan, P., Olney, A., and Rose, C. P. 2007. When are tutorial dialogues more effective than training? *Cognitive Science, 31*, (2007), 3-62.

[6] Crossley, S. A., Varner, L. K., Roscoe, R. D., and McNamara, D. S. 2013. Using automated indices of cohesion to evaluate an intelligent tutoring system and an automated writing evaluation system. In K. Yacef et al (Eds.), *Proceedings of the 16th International Conference on Artificial Intelligence in Education (AIED).* Springer, Heidelberg, Berlin, 269-278.

[7] Rus, V., McCarthy, P., Graesser, A. C., and McNamara, D. S. 2009. Identification of sentence-to-sentence relations using a textual entailer. *Research on Language and Computation, 7,* (2009) 209-229.

[8] Varner, L. K., Jackson, G. T., Snow, E. L., and McNamara, D. S. 2013. Are you committed? Investigating interactions among reading commitment, natural language input, and students' learning outcomes. In S. K. D'Mello, R. A., Calvo, & A. Olney (Eds.), *Proceedings of the 6th International Conference on Educational Data Mining*. Springer, Heidelberg, Berlin, 368-369.

[9] Graesser, A. C., McNamara, D. S., and Rus, V. 2007. Computational modeling of discourse and conversation. In M. Spivey, M. Joanisse, & K. McRae (Eds.), *Cambridge Handbook of Psycholinguistics*. Cambridge University Press, Cambridge, UK.

[10] Roscoe, R. D., Allen, L. K., Weston, J. L., Crossley, S. A., and McNamara, D. S. 2014. The Writing Pal intelligent tutoring system: Usability testing and development. *Computers and Composition*, 34, (2014) 39-59.

[11] McNamara, D. S., Crossley, S. A., Roscoe, R. D., Allen, L. K., and Dai, J. 2015. Hierarchical classification approach to automated essay scoring. *Assessing Writing*, 23, (2015), 35-59.

[12] Allen, L. K., Crossley, S. A., Snow, E. L., and McNamara, D. S. 2014. Game-based writing strategy tutoring for second language learners: Game enjoyment as a key to engagement. *Language Learning and Technology*, 18 (2014), 124-150.

[13] McNamara, D. S., Graesser, A. C., McCarthy, P., and Cai, Z. (2014). *Automated evaluation of text and discourse with Coh-Metrix.* Cambridge: Cambridge University Press.

[14] Allen, L. K., Snow, E.L., and McNamara, D. S. 2014. The long and winding road: Investigating the differential writing patterns of high and low skilled writers. In J. Stamper, S. Pardos, M. Mavrikis, & B. M. McLaren (Eds.), *Proceedings of the 7th International Conference on Educational Data Mining* (pp. 304-307). London, UK.