Using an Additive Factor Model and Performance Factor Analysis to Assess Learning Gains in a Tutoring System to Help Adults with Reading Difficulties

Genghu Shi
University of Memphis
The Institute for Intelligent Systems
365 Innovation Dr.
Memphis, TN, 38152
1 001 901 438 8934
gshi@memphis.edu

Philip Pavlik, Jr
University of Memphis
The Institute for Intelligent Systems
365 Innovation Dr.
Memphis, TN, 38152
1 001 901 678 2326
ppavlik@memphis.edu

Arthur Graesser
University of Memphis
The Institute for Intelligent Systems
365 Innovation Dr.
Memphis, TN, 38152
1 001 901 240 4795
art.graesser@gmail.com

ABSTRACT
After developing an intelligent tutoring system (ITS), or any other class of learning environments, one of the first questions that should be asked is whether the system was effective in helping students learn the targeted skills or subject matter. In this study, we employed two educational data mining models (Additive Factor Model, AFM and Performance Factor Analysis, PFA) which are available in Datashop (LearnSphere) to assess the learning gains on 5 theoretical levels of adults. With AFM, for the KC models tested, the results showed positive learning gains for the Rhetorical Structure knowledge component in contrast, for the PFA model, adults did not learn from either successes or failures.

Keywords
Learning gains, Theoretical Levels, Additive Factor Model, Performance Factor Analysis, CSAL Autotutor

1. INTRODUCTION
One of the first questions that is asked after developing an intelligent tutoring system (ITS) is whether the system was effective in helping students learn the targeted skills or subject matter. Learning gains are based on the performance of the students as they work on the system over time with many opportunities for learning. These learning gains can be assessed at a fine-grained level by tracking the learning of specific knowledge components (KCs), which are particular skills, strategies, concepts, or facts, as articulated in the Knowledge-Learning-Instruction (KLI) framework [2]. In this paper, we analyze the learning of the theoretical components (KCs) which were based on models of comprehension that adopt a multilevel framework in our dialogue-based intelligent tutoring system, called CSAL AutoTutor, that was designed to help struggling adult readers learn reading comprehension strategies. The Graesser and McNamara framework identifies 5 levels [1]: words (W), syntax (S), the explicit textbase (TB), the referential situation model (SM), the discourse genre and rhetorical structure (RS, the type of discourse and its composition). And, the computational models used in the analysis were Additive Factor Model (AFM) and Performance Factor Analysis, both of which were from Datashop (LearnSphere) [3]. 3 questions will be addressed in this paper: 1. When training the adults to read, did the performance of the adults follow the levels of text difficulty? 2. Did adults' learning gains increase after using the Autotutor which just provided some instructions on reading comprehension strategies and some practice? 3. Did adults learn from successes or failures?

2. METHODOLOGY
The adult readers were 52 adults in Atlanta and Toronto who participated in a study of 100 hours of intervention that was conducted by the CSAL team, and they completed up to 30 lessons throughout the intervention. Each lesson had between 10 and 30 multiple choice questions to assess their performance When they answered a question incorrectly, they were given a hint to see whether they selected correctly among the two remaining options. However, in this analysis we only considered performance on their first type, not the follow-up.

The original measures in the AFM model included performance, practice opportunities (the number of questions they answered in a lesson), the knowledge components (KCs were the 5 theoretical components), and subject (participant). For model fitting, pre-test scores and text difficulty (easy, medium, and hard) were entered into the original models (Table 1). Ultimately, we ran 10 models (5 AFM models and 5 PFA models) for the KC approaches, and determined which AFM and PFA models had the best performance, based on AIC, BIC, and Loglikelihood.

Table 1. Models Construction by Adding New Variables

Models	Variables
Model 1	Pre-test score
Model 2	Pre-test score, Text Difficulty
Model 3	Pre-test score, Text Difficulty: KC Model
Model 4	Pre-test score, Practice Opportunity: KC Model
Model 5	Pre-test score, Text Difficulty: Practice Opportunity: KC Model

** These models are basically logit mixed effect models. The ":" refers to interactive effect.*

3. RESULTS AND DISCUSSION
Analyses of the 10 models consistently showed that model 3 was the best model, yielding the lowest AIC BIC and Loglikelihood scores.

Both Table 2 (AFM results) and Table 3 (PFA results) confirm the obvious expectation that pretest score is a strong predictor of adults' performance. Also, only for Rhetorical Structure, performance decreased as a function of text difficulty. This is consistent with the Graesser and McNamara's multilevel

Cite as Shi, G., Pavlik Jr, P., & Graesser, A. (2017). Using an additive factor model and performance factor analysis to assess learning gains in a tutoring system to help adults with reading difficulties. In X. Hu, T. Barnes, A. Hershkovitz, L. Paquette (Eds), *Proceedings of the 10th International Conference on Educational Data Mining* (pp.376-377). Wuhan, China: EDM Society.

theoretical framework that distinguishes the deeper discourse levels of processing (such as the Situation Model and Rhetorical Structure) from the basic reading levels (such as Words and Syntax) [1]. As shown in table 2, only for Rhetorical Structure, performance significantly got better as the practice opportunity increased, but the case of the other KCs was different. As shown in table 3, although cumulative correctness had significant interactions with Syntax and Situational Model, while cumulative incorrectness had significant interactions with Syntax and Textbase, the estimates of these interactions were all negative, which indicated that the performance got worse, no matter adults experienced more successes or failures on these KCs. And, for other KCs, the coefficients drifted to 0.

Table 2. AFM Output of Model 3 – Theoretical Levels

	Estimate	SE	Z Score	P-value	Sig.
Intercept	0.675	0.25	2.66	0.01	**
Pre-test Score	0.140	0.03	4.97	0.00	***
PO : RS	0.001	0.00	2.27	0.02	*
PO : S	-0.124	0.02	-5.16	0.00	***
PO : SM	-0.003	0.00	-3.69	0.00	***
PO : TB	-0.016	0.00	-4.98	0.00	***
PO : W	-0.004	0.00	-0.95	0.34	
RS : Hard	-1.805	0.19	-9.73	0.00	***
S : Hard	0.822	0.28	2.94	0.00	**
SM : Hard	-0.111	0.18	-0.62	0.54	
TB : Hard	0.014	0.19	0.07	0.94	
W : Hard	-0.204	0.30	-0.69	0.49	
RS : Medium	-1.241	0.18	-7.07	0.00	***
S : Medium	-0.078	0.26	-0.30	0.77	
SM : Medium	-0.035	0.18	-0.20	0.84	
TB : Medium	0.133	0.19	0.71	0.48	
W : Medium	0.529	0.29	1.84	0.07	.

**PO refers to practice opportunity. RS refers to Rhetorical Structure. S refers to Syntax. SM refers to Situational Model. TB refers to Textbase. W refers to Word. Easy, Medium, Hard are three levels of text difficulty.*

Table 3. PFA Output of Model 3 – Theoretical Levels

	Estimate	SE	Z Score	P-value	Sig.
Intercept	0.671	0.26	2.60	0.01	**
pretest	0.145	0.03	4.87	0.00	***
CC : RS	0.000	0.00	-0.12	0.91	
CC : S	-0.127	0.04	-3.47	0.00	***
CC : SM	-0.005	0.00	-2.32	0.02	*
CC : TB	-0.008	0.01	-1.30	0.19	
CC : W	-0.004	0.01	-0.69	0.49	
CI : RS	0.005	0.00	1.37	0.17	
CI : S	-0.123	0.04	-3.14	0.00	**
CI : SM	0.001	0.00	0.41	0.68	
CI : TB	-0.031	0.01	-2.77	0.01	**
CI : W	-0.002	0.02	-0.13	0.90	
RS : Hard	-1.808	0.19	-9.74	0.00	***
S : Hard	0.828	0.37	2.22	0.03	*
SM : Hard	-0.099	0.18	-0.55	0.58	
TB : Hard	-0.069	0.20	-0.35	0.73	
W : Hard	-0.209	0.30	-0.69	0.49	
RS : Medium	-1.248	0.18	-7.10	0.00	***
S : Medium	-0.079	0.27	-0.29	0.77	
SM : Medium	-0.023	0.18	-0.13	0.90	
TB : Medium	0.068	0.19	0.35	0.72	
W : Medium	0.524	0.30	1.77	0.08	.

**CC and CI refer to cumulative correctness and cumulative Incorrectness. Others are the same as Table 2.*

4. CONCLUSIONS

The model comparison revealed that practice opportunity, adults' prior literacy skills, KC model (theoretical levels) and text difficulty were factors influencing adults' performance. From the interactions between theoretical levels and text difficulty, we can draw the conclusion that adults' performance on Rhetorical Structure and Situational Model matched the difficulty levels of the texts used in the lessons of the two KCs, that is, they did better on easy texts and worse on medium and hard texts. But for the basic reading levels (Word, Syntax, and Textbase), situations were different. According to the results of AFM model, the learning gains on deeper discourse levels of processing (Rhetorical Structure) increased, because adults' performance became better when they continuously got practice opportunities. There were no learning gains observed on KCs like Situational Model, Syntax, Textbase, and Word. From results of PFA model, we didn't observe significant learning gains from either successes or failures.

5. ACKNOWLEDGMENTS

This research was supported by the National Center of Education Research (NCER) in the Institute of Education Sciences (IES) (R305C120001) and the National Science Foundation Data Infrastructure Building Blocks program under Grant No. (ACI-1443068).

6. REFERENCES

[1] Graesser AC, Mcnamara DS, Kulikowich JM (2011) Coh-Metrix providing multilevel analyses of text characteristics. Educational researcher 40:223-234

[2] Koedinger KR, Corbett AT, Perfetti C (2010) The knowledge-learning-instruction (KLI) framework: Toward bridging the science-practice chasm to enhance robust student learning. Cognitive Science

[3] Pavlik Jr PI, Cen H, Koedinger KR (2009) Performance Factors Analysis--A New Alternative to Knowledge Tracing. Online Submission