Comprehension Factor Analysis: Modeling student's reading behaviour*

Accounting for reading practice in predicting students' learning in MOOCs†

Khushboo Thaker
University of Pittsburgh
Pittsburgh, PA
k.thaker@pitt.edu

Paulo Carvalho
Carnegie Mellon University
Pittsburgh, PA
pcarvalh@cs.cmu.edu

Kenneth Koedinger
Carnegie Mellon University
Pittsburgh, PA
koedinger@cmu.edu

ABSTRACT

Massive Open Online Courses (MOOCs) often incorporate lecture-based learning along with lecture notes, textbooks, and videos to students. Moreover, MOOCs also incorporate practice activities and quizzes. Student learning in MOOCs can be tracked and improved using state-of-the-art student modeling. Currently, this means employing conventional student models that are constructed around Intelligent Tutoring Systems (*ITS*). Traditional ITS systems only utilize students performance interactions (quiz, problem-solving or practice activities). Therefore, text interactions are entirely ignored while modeling students performance in MOOCs using these cognitive models. In this work, we propose a Comprehension Factor Analysis model (*CFM*) for online courses, which integrates student reading interactions in student models to track and predict learning outcomes. Our model evaluation shows that *CFM* outperforms state-of-the-art models in predicting students' performance in a MOOC. These models can help better student-wise adaptation in the context of MOOCs.

KEYWORDS

Student modeling, Education Data Mining, MOOCs, Reading Behaviour

ACM Reference Format:
Khushboo Thaker, Paulo Carvalho, and Kenneth Koedinger. 2019. Comprehension Factor Analysis: Modeling student's reading behaviour: Accounting for reading practice in predicting students' learning in MOOCs. In *The 9th International Learning Analytics & Knowledge Conference (LAK19), March 4–8, 2019, Tempe, AZ, USA.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3303772.3303817

1 INTRODUCTION

Reading is ubiquitous in education, from textbooks to online courses. The truth is, students read to learn, and furthermore, they often believe that reading is the best way to do it. For example, students indicate focusing on reading in online courses is optimal, even when that was not the case [17]. Reading helps a learner to comprehend core ideas for dealing with the practical or analytical aspect of the subject. Although completing optional activities has been shown to be more strongly related to final course performance than reading more [17], it has also been shown that sometimes students who complete more optional readings provided in the course are more successful than those who did not [3]. Thus, reading is an undeniable part of learning in educational contexts. Yet, most student model approaches used in online learning contexts do not take reading into account, instead adopting the student modeling framework from Intelligent Tutoring Systems (ITS) [21]. State-of-the-art student modeling frameworks are based on traditional ITS. The main focus of ITS is to increase students performance through practice activities (which involve quizzes, practice activities or programming steps)[7]. Further ITS systems use student modeling to provide personalized or adaptive content to students for enhance student learning. However, the current online course platforms like Massive Open Online Courses (MOOCs) - for better or for worse [17] - rely on reading as a major learning vehicle. Thus, if we want to employ the power of student models to improve learning in online courses, we must devise student models that take reading into account as a fundamental component of learning in that environment.

To address the challenge of incorporating students' interaction with text in student modeling for online course platforms, in this work we propose to investigate the integration of reading activities in modeling student learning. Specifically, with this work we address two key research questions:

- Do problem solving skills improve with time spent on reading the course content ?
- How can reading behavior be incorporated in student modeling framework ?

To address these research questions we have proposed and investigated Comprehensive Factor Analysis Model (*CFM*), a factor analysis model that incorporates student reading behavior when modeling students' learning and performance. Benefit is *CFM* model will predict better student performance than state-of-the-art student models that completely ignore students' interactions with text. Moreover, *CFM* model can recommend adaptive text in addition to adaptive practice content to students.

*Produces the permission block, and copyright information
†The full version of the author's guide is available as `acmart.pdf` document

LAK19, March 4–8, 2019, Tempe, AZ, USA

ACM ISBN 978-1-4503-6256-6/19/03…$15.00
https://doi.org/10.1145/3303772.3303817

2 RELATED WORK

Approaches to student modeling in ITS could be classified into two major groups: Logistic Regression models [11, 13, 26] and Knowledge Tracing models [6, 24]. Both groups of models rely on expert annotated *Skills* (also known as Knowledge Components or Concepts). *Skills* are knowledge units associated with student activities or steps on which students knowledge or performance is tested [16]. Logistic regression models are motivated by the power law of learning [20], which states that the probability of applying a skill correctly increases by a power function of opportunity. These models utilize student observation logs as the inputs, and predict student performance in a given learning activity based on the *Skills* associated with that activity. One of the basic models in this group is known as Additive Factor Model (*AFM*) [4, 5], which computes the odds of a student's success on a particular problem step based on the number of previous attempts the student had on that step. Performance Factor Analysis [23] extends *AFM* by separately modeling the student's previous successes and failures on a particular skill. In contrast to regression based models, Knowledge Tracing (KT) models [6] directly represent *Skill* level knowledge estimation and allow dynamic knowledge update. KT uses Hidden Markov Models (HMM) to model student knowledge as binary latent variables. Both group of student models are further extended for personalization and adaptation [9, 19, 22, 25].

Reading is a cognitive process whereby the reader builds a situation model of text to comprehend[14] the text. Several computational models are being studied to understand reading behavior[8, 14], which try to infer readers comprehension. A recent trend in student modeling research is to incorporate student reading behavior [10, 12, 25] to incorporate student comprehension. Eagle et al. [10] were among the first to incorporate student reading rate in a knowledge tracing model. Their study depicted the positive effect of integrating students' reading rate to provide individualization. Huang et al. [12] also modeled student reading behavior using a knowledge tracing model for online adaptive textbooks, by learning students skimming and reading behavior. Across these efforts, the key idea is to provide content adaptation based on the student's knowledge state. The model has a strict assumption that students' reading rate is positively correlated with their performance. However, this assumption does not hold for all students [1]. Thaker et al. [25] addressed this limitation by integrating both practice activities and reading interactions to deal with students' noisy reading behavior. Furthermore, recently, Carvalho et al. [3] investigated the effect of attempting optional reading exercises in MOOCs. Their study suggested that attempting optional reading activities helps to boost students' performance and learning[3]. As can be seen, to date, most of the work related to student reading behavior has been done using knowledge tracing models and with the purpose of providing adaptation. In this work we have incorporated student reading behavior in regression based models [4, 23] to model student activity performance. One benefit of our approach is that regression based models outperform knowledge tracing models [23] and can be used in online adaptive environment.

3 MODELING READING BEHAVIOR IN FACTOR ANALYSIS MODEL

Our work attempts to improve student modeling for online course systems that have reading as a significant part of the learning process. To achieve this we will work with traditional factor analysis models *AFM* and *PFA*. Both *AFM* and *PFA* are logistic regression based models. At the base of both models is a Qmatrix. A Qmatrix is a binary matrix where columns represent *Skills* and rows are individual steps of activities. Each cell is a binary value, where 1 in the cell with row r and column c represents that step r is an application of skill c.

As shown in Equation 1, *AFM* represents student probability of success on a step as a function of step difficulty and number of practice attempts the student received on the *Skill*.

$$\textbf{AFM}: \ln \frac{p_{ij}}{1-p_{ij}} = \alpha_i + \sum_k \beta_k Q_{kj} + \sum_k Q_{kj}(\gamma_k N_{ik}) \tag{1}$$

where, i is a student, j is a step. k is a *Skill*. α_i is a coefficient associated with student i (regression intercept) and represents the proficiency of student i. Q is a Qmatrix and Q_{kj} is Qmatrix cell associated with item j and *Skill* k. β_k and γ_k are coefficients associated with skill k. β_k represents the difficulty of skill k, whereas γ_k represents learning rate of skill k. N_{ik} is the number of practice opportunities student i received on skill k.

Because *AFM* relies on practice opportunities, it assumes that the number of practice attempts students get on a particular skill are directly associated with their success on problems targeting that skill. *PFA*, takes *AFM* one step further by incorporating outcomes on previous attempts as shown in Equation 2

$$\textbf{PFA:} \ln \frac{p_{ij}}{1-p_{ij}} = \alpha_i + \sum_k \beta_k Q_{kj} + \sum_k Q_{kj}(\mu_k S_{ik} +_k F_{ik}) \tag{2}$$

where *PFA*, introduces S_{ik} and F_{ik} as number of success and failure attempts respectively of student i on skill k. Thus *PFA* breaks *AFM*'s assumption that all students have similar learning rates γ_{ik} and provides granular evaluation based on individual students' prior success and failure on a particular skill. *PFA* is shown to outperform both *AFM* and Bayesian Knowledge Tracing [23].

Our proposed model, *CFM* is an extension of *PFA*, with the addition of student reading activities as a predictor of student's success in the step. We have two variations of including CFM that vary on how we account for reading behavior.

Reading Opportunities : We will refer to this model as *CFM-RO*. The reading opportunity parameter assumes that students skill mastery improves with the opportunities the student have to read materials associated with the skill. One reading opportunity is a duration for which a student has the text page opened. Thus reading opportunity starts when the student visits a particular page and it ends when the student starts performing practice activities on that page or leaves the page to visit another page. The Below equation defines *CFM-RO* model.

$$\begin{aligned}\textbf{CFM-RO:} \ln \frac{p_{ij}}{1-p_{ij}} &= \alpha_i + \sum_k \beta_k Q_{kj} \\ &+ \sum_k Q_{kj}(\mu_k S_{ik} +_k F_{ik} + \zeta_k RO_{ik})\end{aligned} \tag{3}$$

where, ζ_k is the coefficient which measures the learning rate of a skill from reading opportunities and RO_{ik} is the number of reading opportunity student i has on skill k.

Reading Rate : Reading rate is defined as the speed with which the student is reading through the material.

$$\textbf{Reading Rate} = \frac{\text{Time Spent on a page}}{\text{Number of words on the page}} \tag{4}$$

Reading rate provides us a more granular evaluation of reading based on how long a student remains in a page as a function of how many words the page contains. The equation below is the logistic regression form of *CFM* model with reading rate. We would refer this model as *CFM-RR*.

$$\begin{aligned}\textbf{CFM-RR:}\ \ln \frac{p_{ij}}{1-p_{ij}} &= \alpha_i + \sum_k \beta_k Q_{kj} \\ &+ \sum_k Q_{kj}(\mu_k S_{ik} +_k F_{ik} + \lambda_k RR_{ik})\end{aligned} \tag{5}$$

where, λ_k is the coefficient which measures the learning rate of a skill from reading opportunities and RR_{ik} is the average reading rate of student i on reading materials associated with skill k.

Reading Opportunities only considers the visit of a student to a text section, which could be a misleading evidence. For example, take 5 different pages with different amounts of information (and thus number of words), A student might spend 5 seconds on all these pages. In this scenario, even though these constitute potentially different opportunities (because it took the student the same time to read a page with differing amounts of information) the model will treat them as similar reading opportunities. Instead, when using *Reading Rate*, if a student remains for only 5 seconds on a long page their reading rate will become small, indicating that this reading opportunity differs from a situation in which the student spent 5 seconds on a short page.

4 EXPERIMENT AND EVALUATION

4.1 Course Description and Data Collection

To test the models we used data from a MOOC, "Introduction to Psychology", offered in 2013. The course was offered through Coursera[1] Platform and included materials from Carnegie Melon University's Open Learning Initiative[2] (OLI) learning environment. In addition to pretest, quiz, lectures, and final exam, the course also included text, examples, video and practice activities, offered through the OLI platform. In OLI, the course is separated into modules. Each module is a set of pages and each page consists of text along with multiple practice activities. Although the text is helpful to learn the *Skills* covered in the module, the two types of practice activities included support students' learning outcomes through practice, hints and feedback ("Learn By Doing" activities) and self-evaluation ("Did I get this?" activities). A snapshot of an OLI page in this course is shown in Figure 1.

[1] http://www.coursera.org
[2] http://oli.cmu.edu

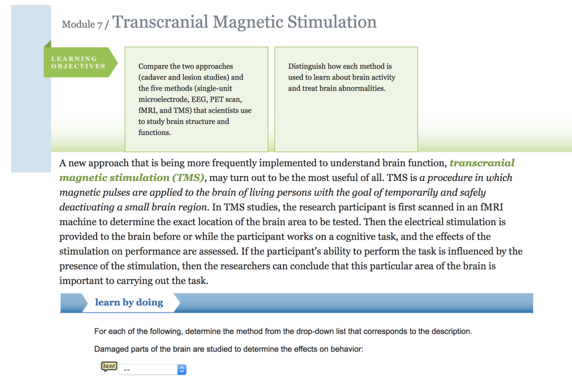

Figure 1: A snapshot of OLI Learning Platform, displaying a typical page which consists of Learning Objectives and text followed by activities

Table 1: Skill Statistics

Total number of skills	226
skills associated with Reading	114
skills associated with practice activities	199
skills associated with both in reading and activity	87

4.2 Skill mapping and statistics

Every module in OLI consists of a set of *learning objectives* as shown in Figure 1. *Learning objectives* are further mapped to *Skills* by experts in the field of study. This allowed us to map the text in the module to expert annotated *Skills* associated with each practice activity. For example, Learning objective - 'Explain emotional intelligence and how it differs from traditional intelligence' is annotated with Skill 'explain emotional traditional intelligence', so the text in pages with the learning objective 'Explain emotional intelligence and how it differs from traditional intelligence' were marked as covering the skill 'explain emotional traditional intelligence'. In addition, practice activities "Learn By Doing" and "Did I get this?" are also annotated with *Skills* by experts. When we use the learning objectives of a given page to identify the skills covered in the text in that page, there is a substantial overlap between *Skills* from reading and *Skills* in practice activities as tagged by experts (see Table 1). Figure 2 depicts the distribution of reading skills tested on practice activities. As evident from the distribution, there are considerable practice questions, which help in understanding students performance after reading. It is important for our model to have common *Skills* between text and activities, to understand the impact of reading on students' performance in the practice activities.

4.3 Reading Interaction Data

Students' interaction logs are available from DataShop[3] repository. The reading behavior was logged using the OLI systems. The OLI system keeps track of how much time the student was with the page (text) before the student started with practice questions (for

[3] ://pslcdatashop.web.cmu.edu/DatasetInfo?datasetId=863

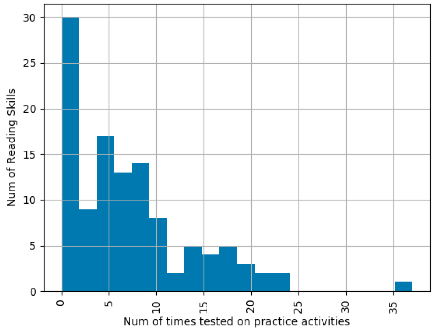

Figure 2: Distribution of *Skills* associated with Reading and tested on practice activities

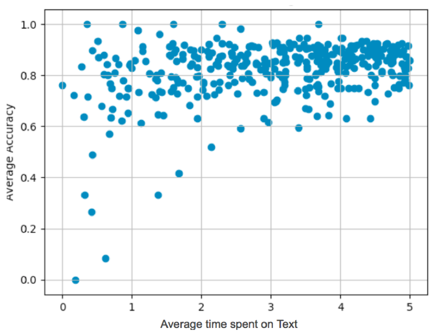

Figure 3: Each dot represents a student. The graph shows that average students' time spent on page is correlated with student performance accuracy (Pearson correlation: 0.322)

details refer to [15]). The reading logs are noisy. A student can open a course content, start reading and then leave for some personal work, as the system will remain open until time out, this will generate a log that suggests the student was reading that content. Similarly, students might open the page and immediately try the activities or open another page. To handle this noise, we took the reading time distribution for each student and removed the records which fell within 5% of left and right ends of the distribution. The graph in Figure 3 is the distribution of students' performance with average time spent across reading material. Each data point in the graph represents a student, vertical axis of the graph is the average accuracy of students' performance across practice activities and the horizontal axis is the average time the student spent across text pages. As it can be seen in the graph, spending more time in the text is correlated with better performance in corresponding activities (Pearson Correlation = 0.322).

4.4 Model details

5, 615 students enrolled in the course, of which only 777 students completed the final exam. Moreover, only a subset of these students finished their practice on OLI modules (OLI activities were optional materials provided to students for practice). To train our models we considered the interactions of 286 students who accessed at least 90% of the pages and completed at least 90% of the practice activities

Table 2: Reading Interaction details

Student-wise avg Read Interactions	301.64
Student-wise avg Activity Interactions	1243.91
Students attempted at least 90% reading and practice activities	286
Total Readings	154
Total Steps	1913

(refer Table 2). To train the models we used all practice activities, irrespective of whether the target *skills* for the activity was also presented in the text or not. For activities without reading *skill* overlap the model will behave like *PFA*, as the reading parameter (RO_{ik} or RR_{ik}) is assigned a value of zero - performance in those steps is purely predicted by activity practice.

5 RESULTS AND DISCUSSION

To evaluate the quality of models we used Akaike Information Criteria (AIC), Bayesian Information Criteria (BIC), and 10-fold cross-validation Root Mean Squared Error (RMSE). For all the metrics lower values are better. As we have a large number of samples AIC is not enough to compare models, so we include BIC which works well with large data samples [18]. In order to compare the performance among four models, Wilcoxon Signed Ranks Tests were conducted on resulted AICs, BICs and RMSEs. The results of different baseline and target models are presented in Table 3. Similar to Pavlik et al. [23], we found that *PFA* performs better than *AFM* both in terms of model fit criteria (lower AIC & BIC) and least prediction error. The main purpose of this work is to test whether considering students' reading behavior (*CFM*) improves model prediction compared to baseline models (*PFA* and *AFM*) that consider only practice activities. This seems to be the case as shown by the good performance of *CFM-RO* compared to *AFM* and *PFA*. Carvalho et al. [3], found similar results, showing the positive impact of optional reading on students' final quiz performance. Moreover, we tried two different variations of incorporating reading behavior: *Reading Opportunity* (*CFM-RO*) and *Reading Rate* (*CFM-RR*). *CFM-RR* is the best model. One possibility, as described above, is that the *Reading Rate* captures students' reading behavior better than *Reading Opportunity*). This finding is not surprising, as *Reading Rate* captures more details of the student's reading behavior than *Reading Opportunity*, as mentioned in section 3.

6 CONCLUSION AND FUTURE WORK

This paper investigated the significance of integrating students' reading patterns in a student modeling framework for online courses. The model *CFM* that includes reading as a predictor of success in a step significantly outperformed the basic models *PFA* and *AFM* in predicting students' performance. The results indicate that students' reading patterns can help infer student's performance. In the future, we would like to investigate our model on other performance activities such as the final exam, quiz and pre-post test. To incorporate student reading pattern we tried two approaches - *Reading Opportunity* and *Reading Rate*, and found *Reading Rate* to be a better indicator of students' reading behavior. In future work, we would like to investigate other ways to represent reading behavior, one example is to group different reading patterns (slow, fast

Table 3: Student Activity Prediction in terms of BIC, AIC and 10 fold cross-validation RMSE. *Denotes a significant performance. Number in bold indicate the best performance.

Model	Reading opportunity	Reading Rate	BIC	AIC	delta-AIC	weight	10 fold cross validation RMSE	Precision	Recall
AFM	-	-	187335.6	181778	6477	<0.001	1.627e-01	0.836	.841
PFA	-	-	187724.5	181181	5882	<0.001	1.613e-01	0.836	.821
CFM-RO	✓	-	186792.1	181148	5849	<0.001	1.600e-01	0.842	.843
CFM-RR*	-	✓	**181718.7***	**175300***	0	1	**1.502e-01***	**0.856***	**.852***

and moderate reading) and use them to represent reading behavior. Currently our study used expert annotated 'Learning Objectives' and *skill* mapping. Although this approach yielded good results, it might not always be possible to obtain learning objectives for each text page. In ongoing work, we are establishing new ways to automatically extracting *skills* through topic analysis [2], which would allow the use of *CFM* to predict student learning from an array of text and activities in online courses, even if the text and activities are not skill-tagged. This work represents a first demonstration of the power of considering students' reading behavior in logistic regression student models in the context of online courses. Although the present work constitutes only an initial demonstration, we believe these types of models could play a bigger role in the future. For example, *CFM* models could be used to integrate students interactions on passive activities like reading, watching videos, discussion forums to predict student learning in an online course. This information could help students in online courses (especially MOOCs) better monitor their learning and their instructors provide appropriate feedback and scaffold.

ACKNOWLEDGMENTS

This paper was partially supported by the National Science Foundation Grants ACI-1443068 and IIS-1525186 and by funding from Google. All data presented here is available from DataShop[4].

REFERENCES

[1] Ryan Shaun Baker, Albert T. Corbett, Kenneth R. Koedinger, and Angela Z. Wagner. 2004. Off-task Behavior in the Cognitive Tutor Classroom: When Students "Game the System". In *Proc. the SIGCHI Conf. on Human Factors in Computing Systems (CHI '04)*. ACM, New York, NY, USA, 383–390.

[2] David M Blei, Andrew Y Ng, and Michael I Jordan. 2003. Latent dirichlet allocation. *Journal of Machine Learning Research* 3 (2003), 993–1022.

[3] Paulo F Carvalho, Min Gao, Benjamin A Motz, and Kenneth R Koedinger. [n. d.]. Analyzing the relative learning benefits of completing required activities and optional readings in online courses. In *The 11th Int. Conf. on Educational Data Mining*, Vol. 34. 68.

[4] Hao Cen, Kenneth Koedinger, and Brian Junker. 2006. Learning Factors Analysis: A General Method for Cognitive Model Evaluation and Improvement. In *Intelligent Tutoring Systems*. Lecture Notes in Computer Science, Vol. 4053. Springer Berlin / Heidelberg, 164–175.

[5] Hao Cen, Kenneth Koedinger, and Brian Junker. 2008. Comparing two IRT models for conjunctive skills. In *Intelligent Tutoring Systems*. Springer, 796–798.

[6] Albert T. Corbett and John R. Anderson. 1995. Knowledge tracing: Modelling the acquisition of procedural knowledge. *User Modeling and User-Adapted Interaction* 4, 4 (1995), 253–278.

[7] Albert T Corbett, Kenneth R Koedinger, and John R Anderson. 1997. Intelligent tutoring systems. In *Handbook of Human-Computer Interaction (Second Edition)*. Elsevier, 849–874.

[8] Mihai Dascalu, Ionut Cristian Paraschiv, Danielle S. McNamara, and Stefan Trausan-Matu. 2018. Towards an Automated Model of Comprehension (AMoC). In *Lifelong Technology-Enhanced Learning*, Viktoria Pammer-Schindler, Mar Pérez-Sanagustín, Henrik Drachsler, Raymond Elferink, and Maren Scheffel (Eds.). Springer International Publishing, Cham, 427–436.

[9] Nicholas Diana, John Stamper, and Ken Koedinger. 2018. An Instructional Factors Analysis of an Online Logical Fallacy Tutoring System. In *Artificial Intelligence in Education*, Carolyn Penstein Rosé, Roberto Martínez-Maldonado, H. Ulrich Hoppe, Rose Luckin, Manolis Mavrikis, Kaska Porayska-Pomsta, Bruce McLaren, and Benedict du Boulay (Eds.). Springer International Publishing, Cham, 86–97.

[10] Michael Eagle, Albert Corbett, John Stamper, Bruce M. McLaren, Angela Wagner, Benjamin MacLaren, and Aaron Mitchell. 2016. Estimating Individual Differences for Student Modeling in Intelligent Tutors from Reading and Pretest Data. In *Intelligent Tutoring Systems*, Alessandro Micarelli, John Stamper, and Kitty Panourgia (Eds.). Springer International Publishing, Cham, 133–143.

[11] Giles Hooker, Matthew Finkelman, and Armin Schwartzman. 2009. Paradoxical results in multidimensional Item Response Theory. *Psychometrika* 74, 3 (2009), 419–442.

[12] Yun Huang, José P González-Brenes, Rohit Kumar, and Peter Brusilovsky. 2015. A framework for multifaceted evaluation of student models. In *Proc. the 8th Int. Conf. on Educational Data Mining*. 203–210.

[13] Brian W Junker and Klaas Sijtsma. 2001. Cognitive assessment models with few assumptions, and connections with nonparametric item response theory. *Applied Psychological Measurement* 25, 3 (2001), 258–272.

[14] Walter Kintsch and David M Welsch. 1991. The construction-integration model: A framework for studying memory for text. (1991).

[15] Kenneth R. Koedinger, R. S. J. Baker, K. Cunningham, A. Skogsholm, B. Leber, and John Stamper. 2010. A data repository for the EDM community: The PSLC DataShop. In *Handbook of Educational Data Mining*. CRC Press, Boca Raton, FL, 43–55.

[16] Kenneth R Koedinger, Albert T Corbett, and Charles Perfetti. 2012. The Knowledge-Learning-Instruction framework: Bridging the science-practice chasm to enhance robust student learning. *Cognitive science* 36, 5 (2012), 757–798.

[17] Kenneth R Koedinger, Jihee Kim, Julianna Zhuxin Jia, Elizabeth A McLaughlin, and Norman L Bier. 2015. Learning is Not a Spectator Sport: Doing is Better than Watching for Learning from a MOOC. In *Proceedings of the second (2015) ACM conference on learning@ scale*. ACM, 111–120.

[18] Jouni Kuha. 2004. AIC and BIC: Comparisons of assumptions and performance. *Sociological methods & research* 33, 2 (2004), 188–229.

[19] Jung In Lee and Emma Brunskill. 2012. The Impact on Individualizing Student Models on Necessary Practice Opportunities. In *Proc. the 5th Int. Conf. on Educational Data Mining*. www.educationaldatamining.org, Chania, Greece, 118–125.

[20] Allen Newell and Paul S Rosenbloom. 1981. Mechanisms of skill acquisition and the law of practice. *Cognitive skills and their acquisition* 1 (1981), 1–55.

[21] Zachary A Pardos, Yoav Bergner, Daniel T Seaton, and David E Pritchard. 2013. Adapting Bayesian Knowledge Tracing to a Massive Open Online Course in edX. *EDM* 13 (2013), 137–144.

[22] Zachary A. Pardos and Neil T. Heffernan. 2010. Modeling individualization in a bayesian networks implementation of knowledge tracing. In *Proc. the 18th Int. Conf. on User Modeling, Adaptation, and Personalization*. Springer-Verlag, Berlin, Heidelberg, 255–266.

[23] P.I. Pavlik, H. Cen, and K.R. Koedinger. 2009. Performance Factors Analysis–A New Alternative to Knowledge Tracing. In *Proc. the 2009 Conf. on Artificial Intelligence in Education: Building Learning Systems that Care: From Knowledge Representation to Affective Modelling*. IOS Press, 531–538.

[24] Radek Pelánek. 2017. Bayesian knowledge tracing, logistic models, and beyond: an overview of learner modeling techniques. *User Modeling and User-Adapted Interaction* 27 (2017), 313–350.

[25] Khushboo Thaker, Yun Huang, Peter Brusilovsky, and He Daqing. 2018. Dynamic Knowledge Modeling with Heterogeneous Activities for Adaptive Textbooks. In *The 11th Int. Conf. on Educational Data Mining*. 592–595.

[26] Amery D Wu and Bruno D Zumbo. 2007. Thinking About Item Response Theory from a Logistic Regression Perspective. *Real data analysis* (2007), 241–269.

[4]https://pslcdatashop.web.cmu.edu