Exploring the Relationship Between Student Pre-knowledge and Engagement in MOOCs Using Polytomous IRT

Jingxuan Liu and Hongli Li
Georgia State University, USA
jliu56@student.gsu.edu, hli24@gsu.edu

ABSTRACT

One of the issues that MOOCs face since its emergence is the low engagement rate and accomplish rate. As an open and free education source, MOOCs are available for people around the world with different motivations and previous knowledge to join. It is a challenge to keep students engaged in a MOOC environment. In the present study, we implement a polytomous item response model (IRT) to explore the relationship between students' self-evaluation of their previous knowledge and students' engagement behaviors in a Geography MOOC. Specifically, we estimate students' latent trait, pre-knowledge, through 15 likert-scale items. Engagement behaviors include assignment, peer review, forum, comment, quiz, and lecture. Each of them is quantified by the aggregated frequency. Then we examine the correlation between pre-knowledge and each type of engagement behavior. We find self-evaluation on previous knowledge cannot predict students' engagement behaviors for any type of engagement. This application indicates that the self-evaluation of pre-knowledge does not predict student engagement in MOOC environment. However, it shows that traditional psychometric models used for standardized tests may be useful and promising in the MOOC context.

Keywords

MOOC, engagement, pre-knowledge, Polytomous IRT

1. INTRODUCTION

A massive open online course (MOOC) is a model for delivering learning content online to anyone who wants to take a course, with no limit on attendance. MOOC engagement is a concept to describe students' involvement of a MOOC. Usually it includes behaviors like posting questions and comments in the MOOC system, submitting assignment and quiz, and other behaviors, which can directly predict students' achievement. Although during the past decade, the number of MOOC students increased tremendously across the world, the low accomplishment and low level of active engagement is always a problem for MOOC development [1]. MOOC engagement is important to predict students' achievement and to show whether students really learned something from the course or not. Students' prior knowledge, which was defined by first two assignments' performance, in computer science and problem solving had impact on their MOOC performance [3]. In the current research, we used pre-course survey data to define pre-knowledge of Geography and to explore if it can predict students' MOOC engagement. Also we use a polytomous IRT model to exam each item and their performance.

2. POLYTOMOUS IRT

Polytomous IRT model is an important model in the IRT family, which is designed for items with more than 2 possible options. Within polytomous IRT models, there are mainly four types: the partial credit model, the rating scale model, the generalized partial credit model, and the graded response model. One example of the application of the graded response model is attitude survey data. Usually the format of item in an attitude survey is likert-scale. For example, for question, "how much do you think you like this opera?", the options can be 5 likert scale from "I like it very much" to "I don't like it at all". The mathematic equation for polytomous IRT model is the following:

$$P^*_{x_{ij}}(\theta_i) = P(X_{ij} \geq x_{ij}|\theta_i) = \frac{e^{Da_j(\theta_i - b_{ij})}}{1 + e^{Da_j(\theta_i - b_{ij})}}$$

In the above equation, D equals to 1.7. For each item j, a_j is a discrimination parameter, and b_{ij} is the difficulty parameter for each option i in each item j ($b_1 < b_2 < ... < b_n$) [2]. Figure 1 indicates a graded response function of a polytomous item. Take the blue line as an example, people with higher theta level seldom choose this option, since the slope is roughly negative.

3. METHOD

3.1 Data

Data comes from a MOOC in Geography. It has enrolled over 100,000 students from 200 countries to date. Data from its 2014 class was used in the present study. In total, after excluding students with little data, there were 3058 students in the current analysis.

3.2 Measure

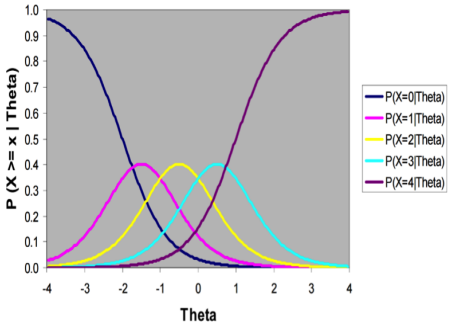

Figure 1: Graded response function

Table 1: Factor loading for each item

Item	1	2	3	4	5
Factor Loading	0.630	0.427	0.608	0.782	0.522
Item	6	7	8	9	10
Factor Loading	0.726	0.705	0.769	0.798	0.697
Item	11	12	13	14	15
Factor Loading	0.668	0.657	0.656	0.698	0.800

There are 15 seven-point likert-scale items, from "strongly agree" to "strongly disagree" designed for students to evaluate their pre-knowledge of Geography. One example is "I enjoy reading maps." In terms of the students' engagement behavior, there are six criteria including assignment, peer review, forum, comment, quiz, and lecture. The method for quantify them is to aggregate the number of times they participate in each type of behavior.

3.3 Procedure

The graded response model was applied using package mirt in R to estimate students' pre-knowledge of Geography. Then the Pearson correlation coefficients between pre-knowledge and each type of engagement behaviors were calculated respectively to examine if students' pre-knowledge influence their engagement behaviors in the MOOC environment.

4. RESULTS

The model fit indices verify a good model fit (RMSEA=0.047, RMSEA_5=0.041, RMSEA_95=0.053, CFI=0.959). The factor loading estimation shows that these 15 items can be used to measure the latent trait, pre-knowledge of Geography (table 1). The parameter estimates are presented in table 2, and the graded response function for each items is shown in the following figure 2. Additionally, table 3 presents the correlation coefficients between pre-knowledge of Geography and each type of engagement behavior.

5. CONCLUSIONS

Table 2: Parameter estimation for each item.

item	a	b1	b2	b3	b4
1	1.38	-1.741	-0.454	1.074	N/A
2	0.804	-2.746	-0.255	2.151	N/A
3	1.302	-6.566	-1.68	0.12	1.842
4	2.133	-5.112	-1.695	-0.553	0.581
5	1.041	-7.801	-2.436	-0.504	0.969
6	1.795	-5.465	-1.72	-0.199	1.096
7	1.693	-5.613	-2.494	-1.345	-0.301
8	2.049	-5.19	-1.768	-0.606	0.639
9	2.257	-5.027	-1.878	-0.757	0.23
10	1.654	-5.692	-1.828	-0.3	0.986
11	1.529	-5.932	-1.553	-0.026	1.268
12	1.482	-6.049	-2.67	-1.278	0.001
13	1.478	-6.048	-2.213	-0.933	0.253
14	1.66	-1.771	-0.307	0.898	N/A
15	2.268	-5.02	-1.621	-0.545	0.52

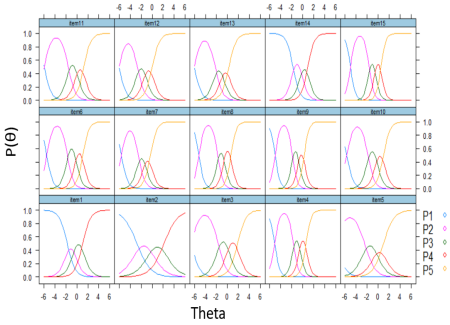

Figure 2: Graded response function for each item.

Table 3: The Pearson correlation coefficient between pre-knowledge and Engagement Behavior Type (EBT)

EBT	Pre-knowledge of Geography
	The Pearson correlation coefficient
assignment	-0.018
peer review	-0.022
forum	-0.016
comment	-0.022
quiz	-0.019
lecture	-0.025

All of the 15 items have relatively good loading on one factor, so it is reasonable to use one-dimensional IRT model. Also, the fit indices show that this graded response model fit well with the data. In terms of the discrimination index, item 8, item 9, item 15 have very good discrimination level. It indicates that these three items can provide more information in terms of students' pre-knowledge of Geography than other items. In terms of the difficulty parameter, b4 cannot be estimated for item 1, item 2, and item 14. This indicates that these items might be problematic.

All of the correlation coefficients are negative and nonsignificant (p-value>.05). This results indicates that although the general trend is students with less pre-knowledge of Geography will have less frequency of engagement behavior, none of them are statistically significant. In other words, whether students report a relative rich or poor pre-knowledge of Geography cannot predict their engagement behaviors. One of the explanation may be the pre-knowledge here is measured by self-evaluation, which relates to the meta-cognitional ability of students. This subjective report is different from objective questions, such as "have you taken any university level courses related to this MOOC course?" In further research, more direct measure of pre-knowledge is needed.

6. REFERENCES

[1] J. Daniel. Making sense of moocs: Musings in a maze of myth, paradox and possibility. *Journal of interactive Media in education*, 2012(3), 2012.

[2] R. J. De Ayala. *The theory and practice of item response theory*. Guilford Publications, 2013.

[3] G. Kennedy, C. Coffrin, P. De Barba, and L. Corrin. Predicting success: how learners' prior knowledge, skills and activities predict mooc performance. In *Proceedings of the Fifth International Conference on Learning Analytics And Knowledge*, pages 136–140. ACM, 2015.