Majumdar, R. et al. (Eds.) (2025). Proceedings of the 1st International Conference on Learning Evidence and Analytics. Asia-Pacific Society for Computers in Education

Modelling Test and Study Behaviors in Math Exercise Solving

Yiling DAI[a*], Hatsune ICHIDATE[b], Brendan FLANAGAN[c] & Hiroaki OGATA[a]
[a] *Academic Center for Computing and Media Studies, Kyoto University, Japan*
[b] *Graduate School of Informatics, Kyoto University, Japan*
[c] *Center for Innovative Research and Education in Data Science, Institute for Liberal Arts and Sciences, Kyoto University, Japan*
*yilingdaiac@gmail.com

Abstract: Learning happens when learners repeat the test and study behaviors on the materials. However, it is practically not clear how these behaviors relate to the learning outcomes. In this study, we focus on the context of K12 students solving mathematical exercises. We categorized students' test, study, and restudy behaviors based on the attempt transition types. We discovered four clusters with different levels of engagement, understanding towards contents, and reviewing behaviors. We then investigated the relationship between the detected patterns and learning performance, which provides insights on how to improve the effectiveness and efficiency of learning.

Keywords: Repeated practice, math exercise solving, study and test

1. Introduction

Learning can be viewed as a repeated sequence of study and test trials over the materials (Roediger & Butler, 2011). For mathematics, it has been basically taught and learnt through exercises and examples globally (Schmidt et al., 1997). Learners utilize the exercises as materials for studying or testing accordingly. It is considered that learning occurs with the early study and test trials but decreases with additional trials (Roediger & Butler, 2011). However, it is not practically clear how the amount of study and test trials is related to the learning outcomes. To match the gap between theoretical hypothesis and practice, researchers in the field of learning analytics have mined students' engagement patterns from learning logs. Patterns have been found with the focuses on how students interact with the assignments and lecture videos (Boroujeni & Dillenbourg, 2018), and whether the students view course materials before, in, or after class (Akçapinar et al., 2020). Kuromiya et al. (2021) identified four patterns of when students completed the assignments during a vacation. However, they ignored the duplicated answers on the exercises, which is the focus of this study.

In this study, we investigate patterns of study and test trials among K12 students' practice of mathematical exercises. We modelled the students' test and study behaviors based on six attempt transition types--- none to correct, none to incorrect, correct to correct, correct to incorrect, incorrect to correct, and incorrect to incorrect. We then identified four clusters using a standard hierarchical clustering method. We discussed the characteristics of the clusters and investigated the learning performance for each cluster. Promising results help us understand the current study and test behavior of students, which is important for the further investigation of the effectiveness and efficiency of learning, and the potential intervention to optimize the behavior.

2. Learning context

We collected 8-week practice data from an authentic learning environment. During this period, the students were supposed to review the contents they had learnt according to the curriculum, which were then tested in a final test of the semester. A new set of exercises was provided, and the students were able to view the questions and answers through a digital learning platform named LEAF (Ogata et al., 2023). After students attempt the exercises, they were

supposed to check the answers and assess whether their answers are correct or not. The LEAF platform recorded the timestamp and the result of each attempt.

Before we conduct the analysis, we removed the "invalid" records if it is for the same exercise, student, result, and happens in less than five seconds after the previous action. Ultimately, we removed 8.56% of the raw records and collected 21,592 attempt results from 127 eighth-grade students. The average number of exercises attempted was 170.016 (SD=96.996), and the average number of distinct exercises attempted was 132.622 (SD=69.526).

3. Procedure

As shown in Table 1, we model test, study, and restudy behaviors based on the attempt transition types. Intuitively, we assume students learn mostly from study behaviors and then restudy behaviors. To identify the knowledge they know or not, students conduct test behaviors repeatedly. While all three behaviors are necessary parts of efficient learning, we are interested in exploring whether there exist different patterns of these behaviors among students. To identify test and study patterns of students, we first count the numbers of each attempt transition types for each student. Then, we cluster the students using hierarchical clustering method.

Table 1. *Categories of attempt transitions*

Category	Attempt transition type	Implication
Test	None -> correct	The student probably knows the knowledge.
	None -> incorrect	The student probably does not know the knowledge.
Restudy	Correct -> correct	The student probably enhances her/his knowledge.
	Correct -> incorrect	The student probably does not know the knowledge.
Study	Incorrect -> correct	The student probably acquires the knowledge.
	Incorrect -> incorrect	The student probably learns the knowledge.

4. Results and Discussions

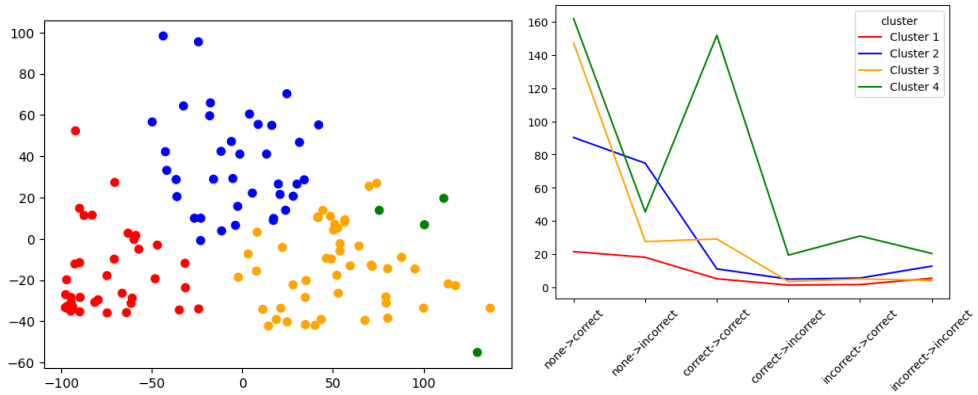

Figure 1. Scatter plot of clusters (left) and line plot of cluster centroids (right).

After conducting a hierarchical clustering analysis, we selected four as the cluster number based on the dendrogram. Figure 1 shows the scatter plot of the students in the clusters after applying principal component analysis, and the plots of attempt transition types of the cluster centroids. We discovered the following characteristics of each cluster: 1) Cluster 1--- Students attempted less exercises and may have struggled in solving the exercises. The students had some restudy and study behaviors but similarly, the correct/incorrect ratio indicates the difficulty of solving the exercises. 2) Cluster 2--- Students attempted an average number of exercises and may have struggled in solving the exercises. The students had little restudy and study behaviors. 3) Cluster 3--- Students attempted a considerable number of exercises and had less difficulties in solving them. The students had little restudy and study behaviors. 4) Cluster 4--- Students attempted an exceptional number of exercises and were able to solve

most of the exercises. The students had many restudy and study behaviors, with a high rate of solving the exercises correctly again.

Table 3. *Statistics of test scores in each cluster (N=127)*

					95% Confidence Interval	
	N	Mean	SD	Skewness	Lower	Upper
Cluster 1	38	58.763	20.494	-0.255	52.027	65.500
Cluster 2	37	74.973	14.943	-1.202	69.991	79.955
Cluster 3	48	85.104	10.368	-0.824	82.094	88.115
Cluster 4	4	86.750	6.898	-0.528	75.774	97.726

We then investigated the learning performance of these clusters using the final test scores. Table 3 shows the statistics of the test scores in each cluster. As the normality test suggested a violation of the assumption of normality, we performed a non-parametric one-way ANOVA (Kruskal-Wallis) test on the test scores of four clusters. The differences between the four clusters were significant, $H(3, n = 127) = 41.890, p < .001$. Post hot comparisons were conducted using Dwass-Steel-Critchlow-Fligner pairwise comparisons. We found that Cluster 1 had significantly lower scores than other clusters, and Cluster 2 had lower scores than Cluster 3. The results approximately echoed with our observation on the test and study behaviors of the four clusters. That is to say, overall engagement and performance of the test behavior are related to the learning performance.

5. Conclusion

In this study, we explored the patterns of test and study behaviors in the context of solving math exercises. Based on the attempt transition types, we discovered four clusters which displayed different levels of engagement, different levels of understanding towards the contents, and different levels of reviewing behavior. We also investigated the learning performance of these clusters, which mostly agreed with our observation of the clusters. In the future, we plan to further analyze the learning performance and elaborate the methods to model the test and study behaviors.

Acknowledgements

This work was partly supported by JSPS Grant-in-Aid for Early-Career Scientists JP24K20902, JSPS Grant-in-Aid for Scientific Research (B) JP20H01722, and JP23H01001, (Exploratory) JP21K19824, NEDO JPNP20006, and JPJ012347.

References

Akçapinar, G., Hasnine, M. N., Majumdar, R., Chen, M.-R. A., Flanagan, B., & Ogata, H. (2020). Exploring Temporal Study Patterns in eBook-based Learning. *Proceedings of the 28th International Conference on Computers in Education*, 342–347.

Boroujeni, M. S., & Dillenbourg, P. (2018). Discovery and temporal analysis of latent study patterns in MOOC interaction sequences. *Proceedings of the 8th International Conference on Learning Analytics and Knowledge*, 206–215. https://doi.org/10.1145/3170358.3170388

Kuromiya, H., Majumdar, R., & Ogata, H. (2021). Mining Students' Engagement Pattern in Summer Vacation Assignment. *Proceedings of the 29th International Conference on Computers in Education*, 559–568. https://library.apsce.net/index.php/ICCE/article/view/4200

Ogata, H., Majumdar, R., & Flanagan, B. (2023). Learning and Evidence Analytics Framework Bridges Research and Practice for Educational Data Science. *Commun. ACM*, *66*(7), 72–74. https://doi.org/10.1145/3589660

Roediger, H. L., & Butler, A. C. (2011). The critical role of retrieval practice in long-term retention. *Trends in Cognitive Sciences*, *15*(1), 20–27. https://doi.org/10.1016/j.tics.2010.09.003

Schmidt, W. H., McKnight, C. C., Valverde, G. A., Houang, R. T., & Wiley, D. E. (1997). *Many visions, many aims A cross-national investigation of curricular intentions in school mathematics*. Springer Dordrecht.