Toward Contextualized Handwriting Process Analysis: Comparison Between Problem Types in Math

Shunsuke TONOSAKI[a*], Taito KANO[a], Satomi HAMADA[a], Izumi HORIKOSHI[b] & Hiroaki OGATA[b]
[a]*Graduation School of Informatics, Kyoto University, Japan*
[b]*Academic Center for Computing and Media Studies, Kyoto University, Japan*
*tonosaki.shunsuke.75i@st.kyoto-u.ac.jp

Abstract: Handwriting logs in the math answering process have recently been collected, and features related to the answering performance and the process, such as the stroke duration, have been investigated. However, the results reported in previous studies showed inconsistencies, and sufficient consideration had not been given to the differences in the problem types. In this study, we classified some problems into two types and verified whether there is a difference in the effect of the handwriting process on performance in each feature. The result of the analysis showed a significant difference in the effects of problems on the features used in this study, such as answering time and number of strokes. This study contributes to the need to take into consideration the problem type in learning support with handwriting process logs.

Keywords: handwriting process, learning logs, learning analytics, learning support

1. Introduction

In Japan, the GIGA School Program introduced one device per student in K12, accumulating extensive daily learning logs. Handwriting process logs from math problem-answering are particularly valuable for analyzing the link between students' performance and cognitive states, with tools developed to help teachers visualize these processes (Yoshitake, Flanagan, and Ogata 2020). However, understanding the relationship between individual performance and answering processes remains challenging for teachers. Thus, efficiently identifying students requiring attention by finding significant features is crucial. For instance, students with low performance and short answering times may have given up early and need support.

Some studies have explored features related to performance, such as average stroke length and duration, in student Handwriting answers (Zhou et al. 2014; Stahovich and Lin 2016). These analyses suggest significant features for identifying students needing attention. However, the findings are inconsistent: one study linked shorter stroke durations to high performance, while another found the opposite. The discrepancy may occur from differences in problem types, with one study using varied difficulty math problems and the other focusing on static problems like balance of power. Noteworthy features may thus depend on the problem type.

This study's purpose is to quantitatively analyze the relationship between handwriting logs and performance across two math problem types to identify noteworthy features. The research question is as follows. This analysis highlights the necessity of contextualized learning analysis, such as classifying support targets according to problem types, in data-informed learning support and offers significant insights for its implementation.

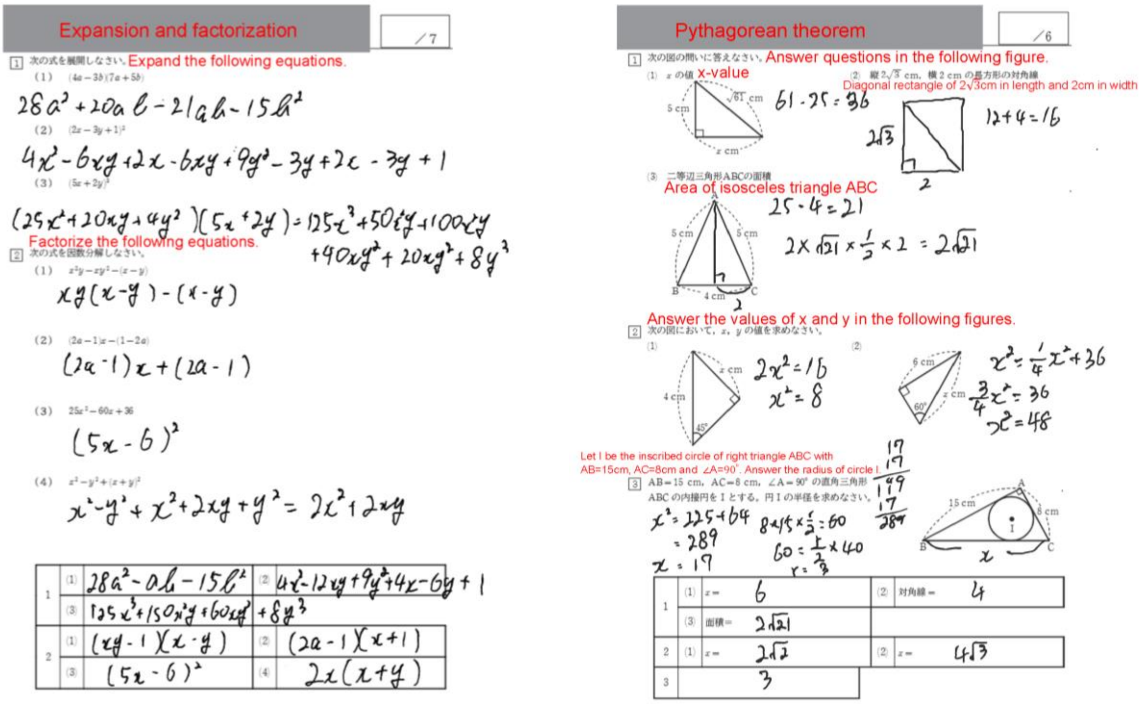
RQ: Is there a difference in each handwriting process feature that affects performance for different problem types?

2. Methods

2.1. Context

In this study, we collected handwriting log data from the Learning and Evidence Analytics Framework (LEAF) system. The LEAF system is an LA platform that collects and analyzes learning log data in the real world to support teaching and learning (Ogata et al. 2018). The LEAF system has a pen stroke system that can collect handwriting log data in daily classrooms and homework situations in junior high and high schools.

The activity analyzed was a series of 4 math quiz-style tests taken by 94 Japanese third-grade in junior high school. For the 4 tests, we chose 2 tests on expansion and factorization as description-type problems (Figure 1 (a)) and 2 tests on the Pythagorean theorem as diagram-type problems (Figure 1 (b)).



(a). Description Problem. *(b).* Diagram Problem.

Figure 1. Sample Problems and Answers of Description-Type & Diagram-Type.

2.2. Data & Analysis

The answer record data used in this study consisted of 96 description-type and 115 diagram-type answers, except the answer that only the last answer and no description of the process at all. The performance indicator, which indicates the level of performance, was the number of problems answered correctly, and scoring was done by graduate students in the laboratory.

In this study, we conducted 6 general linear models, depending on the number of 6 features extracted from the handwriting logs (Table 1), with the problem type and one of the features as the independent variables, and the performance indicator as the dependent variable. The analyses were conducted under these conditions to check whether interaction effects between each feature and problem type appeared significant to confirm that the features have different effects on performance depending on the problem type. In Table 1, *SD* and *LD* are distinguished by the average durations used in this study, with the smaller value representing the short duration and the larger value representing the long duration (Zhou et al. 2014). In addition, the significance level was set at 5%.

Table 1. *Handwriting Features for General Linear Models*

Features	Description
TAT	Total answering time (second) during a problem
TNS	Total number of strokes during a problem
TNE	Total number of erasers during a problem
SD	Total number of short duration (smaller than the average) from the previous stroke
LD	Total number of long duration (larger than the average) from the previous stroke
Speed	The value of *TNS* divided by *TAT* during a problem

3. Result

Table 2 shows the results of the 6 general linear models of the interaction of each feature and problem type. Interactions were significant for all.

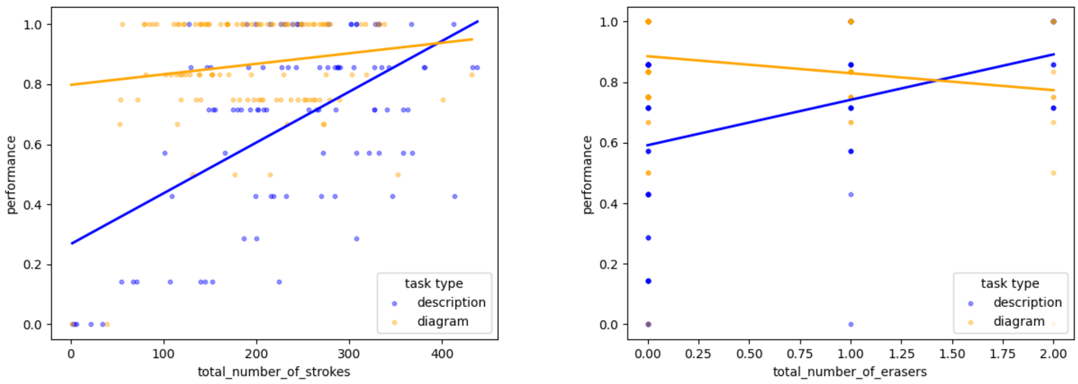
Table 2. *Result of General Linear Models with an Interaction Term between Feature and Problem Type*

Features	Problem type	M	SD	df	sum_sq	F
TAT	Description	524.0	66.5	1.0	0.321	6.08*
	Diagram	505.0	80.0			
TNS	Description	239.0	99.6	1.0	0.738	16.3***
	Diagram	194.0	79.6			
TNE	Description	0.531	0.725	1.0	0.994	18.9***
	Diagram	0.348	0.622			
SD	Description	223.0	95.4	1.0	0.751	16.6***
	Diagram	181.0	77.4			
LD	Description	2.63	8.54	1.0	0.898	18.3***

	Diagram	19.4	6.80			
Speed	Description	0.453	0.183	1.0	0.828	17.2***
	Diagram	0.390	0.166			

*p <.05, **p <.01, ***p <.001

In the features of *TAT*, *TNS*, *SD*, *LD*, and *Speed* where interactions were significant, there was a positive correlation between each feature and performance for both problem types, and the slope of the regression line for the diagram type was slower than for the description type such as *TNS* in Figure 3 (a). Similar results were obtained for other features. On the other hand, in *TNE*, there was a negative correlation between them in diagram type (Figure 3 (b)).



(a). Total Number of Strokes (*TNS*). *(b).* Total Number of Erasers (*TNE*).

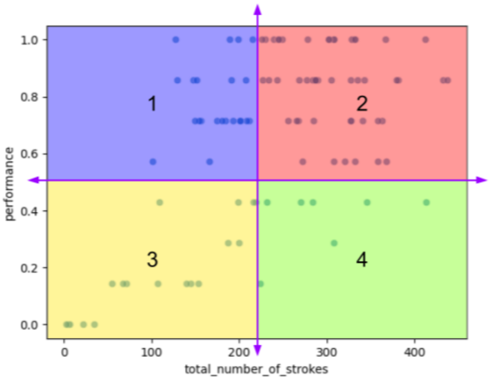
Figure 3. Results of General Linear Models with *TNS* and *TNE* with Interaction.

4. Discussion

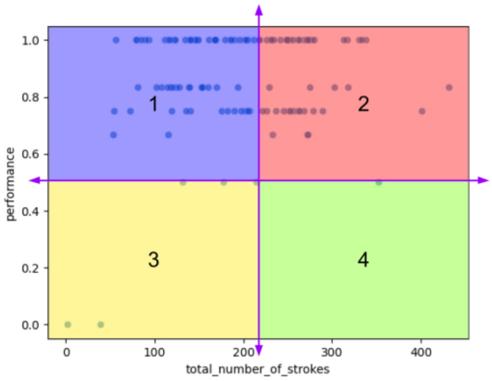
The analysis in this study shows the significant interaction between problem type and features for all features. These findings offer insights into the student problem-answering process, such as revealing that high performers tend to persevere through the handwriting process (high *TAT*, *TNS*, & *LD*) in description-type problems (Stahovich and Lin 2016). However, these features may be less critical to performance in diagram-type problems compared to description-type problems.

The potential applications of the findings of this study are as follows: analyzing the relationship between performance and features according to problem type may help to find the students of need attention and deepen our understanding of each characteristic of high and low performers.

As an example, Figure 4 shows how the results of this study can be used for learning support. Figure 4 plots the *TNS* and performance on two axes, classifies students into 4 quadrants, and is compared two different problem types. In Figure 4 (a), for example, low performers may be facing different issues in the number of strokes, such as a lack of knowledge in the 3 and having mistakes in the 4. On the other hand, in problems such as Figure 4 (b), where high performance can be obtained even with a low *TNS*, it is difficult to characterize performance with the *TNS*. In this way, rather than overgeneralizing support methods, we can hold out hope that contextualized learning support will be achieved by ensuring that appropriate performance-related features are used according to the problem type.



(a). Description Problem.



(b). Diagram Problem.

Figure 4. The Comparison Between Two Problem Types on Classification Example of *TNS*.

5. Conclusion

This study examined whether there is a difference in noteworthy handwriting process features between description and diagram problem types. The analysis revealed significant differences, with features like the number of strokes being noteworthy in descriptions but not in diagrams. These findings suggest the importance of considering problem types to accurately assess student characteristics and provide appropriate support.

Acknowledgements

This study was supported by the CSTI SIP program grant number JPJ012347 and the JSPS KAKENHI (A) JP23H00505.

References

Ogata, H., Rwitajit Majumdar, Gökhan Akçapınar, M. Hasnine, and B. Flanagan. 2018. “Beyond Learning Analytics: Framework for Technology-Enhanced Evidence-Based Education and Learning,” November. https://doi.org/http://hdl.handle.net/2433/237322.

Stahovich, Thomas F., and Hanlung Lin. 2016. “Enabling Data Mining of Handwritten Coursework.” *Computers & Graphics* 57 (June):31–45.

Zhou, Jianlong, Kevin Hang, Sharon Oviatt, Kun Yu, and Fang Chen. 2014. “Combining Empirical and Machine Learning Techniques to Predict Math Expertise Using Pen Signal Features.” In *Proceedings of the 2014 ACM Workshop on Multimodal Learning Analytics Workshop and Grand Challenge*, 29–36. MLA ’14. New York, NY, USA: Association for Computing Machinery.