Yin Yang, Yanjie Song

The Education University of Hong Kong, HKSAR, China

yyin@s.eduhk.hk, ysong@eduhk.hk

Examining primary student self-regulated vocabulary learning behavioural patterns and vocabulary learning outcomes leveraged by the mobile app with a self-regulation scheme

Bio data

Yin Yang is a PhD candidate at The Education University of Hong Kong. Her research interests include technology-enhanced language learning and second language acquisition. She received the Best Teacher Forum Paper Award in GCCCE 2019 and Best Oral Presentation Award in PPTELL 2021.

Yanjie Song is an associate professor in the Department of Mathematics and Information Technology at The Education University of Hong Kong. Her research areas focus on mobile learning, seamless learning, technology-enhanced language learning (e.g., augmented reality), vocabulary acquisition, teacher orchestration with technology, learning analytics, innovative pedagogy design, and implementation.

Abstract

This research reports a case study which provides insights into primary students' SRVL processes and behaviours that interact over time in authentic learning environments beyond the classroom, leveraged by a mobile app with a self-regulation scheme. The participants were 44 grade four students in a government-funded primary school in Mainland China. Data collection included log data on the app, and pre- and post-vocabulary tests. Data analysis included clustering, progress-mining techniques, and Kruskal Wallis tests. The findings showed that (1) students' SRVL behaviours leveraged by a mobile app with a self-regulation scheme could be clustered into three groups, (2) the characteristics of SRVL behavioural patterns among three clusters were discussed, and (3) a significant association between the identified three clusters and the students' vocabulary learning outcomes was observed.

Conference paper

Introduction

It is widely acknowledged that self-regulated learning (SRL) is one of the most essential capabilities for lifelong learning to cope with the challenges of the twenty-first century (Lehmann et al., 2014; Zheng et al., 2018). A number of studies have suggested that students who can regulate language learning tend to perform better than those who do not engage in self-regulation (Saks & Leijen, 2019; Zheng et al., 2018). Recently, due to the advancement of mobile technologies, the vocabulary learning experience has become ubiquitous and contextualised by integrating multimedia. But only limited mobile technologies adopted in current studies appear to support the whole process of self-regulated vocabulary learning (SRVL) (Yang et al., in press). Furthermore, little is known about how mobile technologies support primary students' SRVL in forethought, performance and reflection processes (Zimmerman, 2002), and whether student SRVL behaviours are related to their vocabulary learning outcomes or not.

This study adopted a case study to get insights into the students' SRVL processes and behaviours in authentic learning environments beyond the classroom leveraged by a mobile app with a self-regulation scheme. The following research questions were addressed:

(1) What were primary students' SRVL behavioural patterns leveraged by the mobile app with a self-regulation scheme?
(2) If there were different SRVL behavioural patterns, do students with different patterns differ in their vocabulary learning outcomes?

Literature Review

Self-regulation theory

SRL refers to students' self-initiated actions involving setting goals, monitoring their efforts to achieve goals, regulating their cognitive and metacognitive processes and learning behaviours in their learning processes, and reflecting (Pintrich, 2000; Zimmerman, 2002). Zimmerman (2002) stated that SRL involved three cyclic phases, namely, forethought (e.g., goal setting, strategic planning), performance (e.g., self-observation, self-control), and self-reflection (e.g., self-judgement, self-evaluation). In the forethought phase, the students analyse the learning task which involves goal setting and strategic planning. In the performance phase, the students perform the task while monitoring the learning process. At the same time, they use self-control strategies to keep themselves engaged in learning tasks. In self-reflection, students assess their learning performance and satisfaction, evaluate the strategies used, and reflect on what they will do in the next round of learning.

Self-regulated vocabulary learning using learning analytics

To date, a growing number of studies aim at enhancing students' SRVL in real life learning settings supported by technologies. Aligned with the modern SRL research, SRL was considered as a dynamic process (Li, et al., 2020; Panadero et al., 2016). Despite this, few studies have attempted to investigate SRVL processes and behaviours utilising learning analytics. Even fewer studies have been conducted to examine the relationship between students' SRVL behavioural patterns and vocabulary learning outcomes. In most cases, researchers have taken a variable-centered approach, for example, exploring the features of SRL behaviours between learners with high and low academic performance (Yang et al., 2018). Yet only a limited number of studies have sought to examine how specific SRL behaviours cluster among individual learners (Jang et al., 2017; Li et al., 2020). Thus, there is a significant need to cluster students' SRVL behaviours by putting the students into homogenous groups with similar profiles in order to obtain insights into the generalised patterns of students' SRVL behaviours in mobile learning environments.

This study firstly identified student SRVL behavioural patterns via clustering homogenous groups with similar SRVL behaviours; then, features of each distinct SRVL behavioural pattern were explored using learning analytics. Finally, the relationship between the identified groups of SRVL behavioural patterns and vocabulary learning outcomes was investigated.

Methods

Participants

Participants were 44 grade 4 students in a government-funded primary school in Mainland China. The study lasted for four weeks. The learning unit reported in this study was "Dinner's ready."

Data collection and analysis

Data collection included log data on the mobile app, and pre-and post-vocabulary tests. To understand students' SRVL behaviours and processes leveraged by a mobile app with a self-regulation scheme, a theoretical lens of micro-level SRVL based on the SRL model (Saint et al., 2020; Zimmerman, 2002) was adopted. Table 1 represents students' SRVL behaviours and processes leveraged by a mobile app with a self-regulation scheme in terms of SRL cyclic phases, micro-level SRVL processes and specific activities on the mobile app. In general, SRL was composed of three cyclic phases: forethought, performance, and self-reflection. During the forethought phase, the students analysed the learning task, including goal setting and strategic planning. The app adopted in this study enabled students to set goals regarding words, learning time, and expected ranking. In addition, students could plan specific activities and adjust their learning goals. During the performance phase, students could create learning logs by taking pictures, recording and inputting words or sentences, while simultaneously monitoring their progress. Last, the students performed self-evaluation by taking quizzes, assessing their learning performance, evaluating strategies utilised, and reflecting on what they would do in the next learning cycle. To measure students' learning performance, vocabulary learning tests were used. Vocabulary tests consisted of the Vocabulary Levels Test (VLT) and self-constructed curriculum-based vocabulary learning outcomes tests. The tests were designed to examine both breadth and depth of word knowledge (Schmitt, 1999). The internal consistency has been confirmed in a pilot study with a Cronbach's alpha value above 0.80.

Table 1. *Overview of students' SRVL behaviours on the app*

SRL cyclic phases	Micro-level SRVL	SRVL Descriptions
Performance	Goal-setting	Students set learning goals (e.g., number of learning logs, time, ranking)
	Strategic planning	Students plan the learning strategies to reach the goal.
		Students reset learning goals.
Performance	Self-observation	Students check the overview of recorded learning logs, time spent on the app, and ranking.
	Self-control	Students post learning logs.
		Students review/edit their own logs
		Students review peers' logs
Self-reflection	Self-judgement	Students take quizzes.
		Students evaluate the performance and the efforts.
	Self-reaction	Students evaluate self-satisfaction.
		Students evaluate planned strategies.
		Students reset/modify plans.

Data analysis included agglomerative hierarchical clustering, progress-mining techniques using the R package - PMineR, and Kruskal Wallis tests.

Results

Primary students' SRVL behavioural patterns

After applying the agglomerative hierarchical clustering and using the silhouette method (Dinh et al., 2019) to choose the optimal number of clusters, three clusters were identified. Figure 1 shows the cluster plot, which groups similar SRVL behaviours using a Euclidean distance metric.

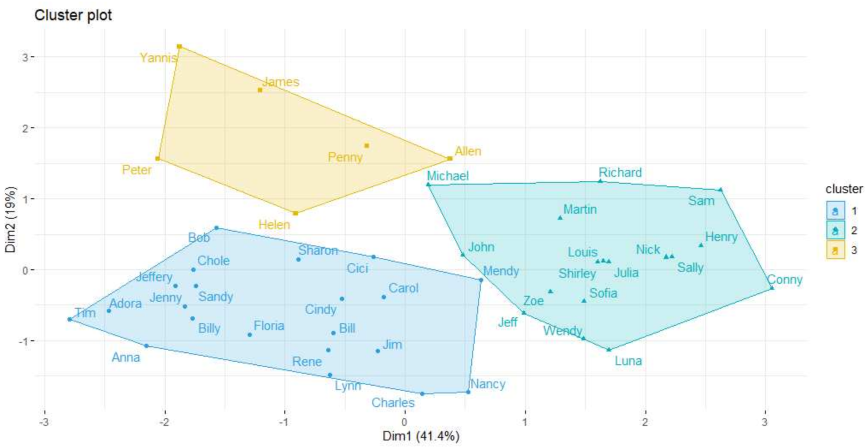

Figure 1. *Cluster plot of the clustering result of primary students' SRVL behaviours*

Then, the process mining technique "First Order Markov Models (FOMMs)" in the pMineR package was employed for primary students' SRVL behaviours. It showed the likelihood of transition among each micro-level SRVL process (Matcha et al., 2019; Peeters et al., 2020). The lines between one node and the next represent the transition probability (TP), which refers to a stochastic measure of the likelihood of transition between one node to another (Saint et al., 2020). The FOMM graphs were generated to compare the differences among the identified three clusters. The similarities and differences of students' SRVL patterns among three clusters could be identified in Figure 2, Figure 3, and Figure 4.

As for the similarities, three aspects were identified. First, three clusters shared similar learning patterns in goal-setting, strategic planning and self-observation. For example, Figure 2 shows that students in Cluster 1 tended to initiate SRVL by setting goals with a transition probability of 0.84, making strategic planning with a transition probability of 0.25, and monitoring the learning process with a transition probability of 0.26. Figure 3 shows that students in Cluster 2 initiated SRVL by setting goals with a transition probability of 0.74, making strategic planning with a transition probability of 0.25, and monitoring the learning process with a transition probability of 0.21. Figure 3 shows that students in Cluster 3 also exhibited a similar pattern, starting with setting goals with a transition probability of 0.67, followed by making strategic planning with a transition probability of 0.24, and monitoring the learning process with a transition probability of 0.27. Second, students in the three clusters showed non-linear learning trajectories between goal-setting and strategic planning, indicating students would revise learning goals. Third, self-reaction was the last activity among students in the three clusters, with a transition probability of 0.41, 0.39, and 0.42, respectively.

As for the differences, two main aspects were identified. First, compared to students in Cluster 1 starting from goal-setting, some students in Cluster 2 moved to self-control with a transition probability of 0.16 to post learning logs, and students in Cluster 3 would engage themselves in self-observation with a transition probability of 0.33. Second, the students in Cluster 2 adopted comprehensive SRL strategies with non-linear learning trajectories. Figure 3 shows that starting from the performance (self-observation and self-control) and the reflection phase (self-judgement and self-reaction), students in cluster 2 adopted comprehensive SRL strategies with various learning trajectories that were mostly non-linear and across different micro-level SRVL processes. While Figure 2 and Figure 4 show that students in Cluster 1 and Cluster 3 followed a linear learning trajectory in the performance phase (self-observation and self-control), and focused much on posting and/or viewing learning logs. Students seldom reviewed the learning process or ranking after posting/viewing learning logs.

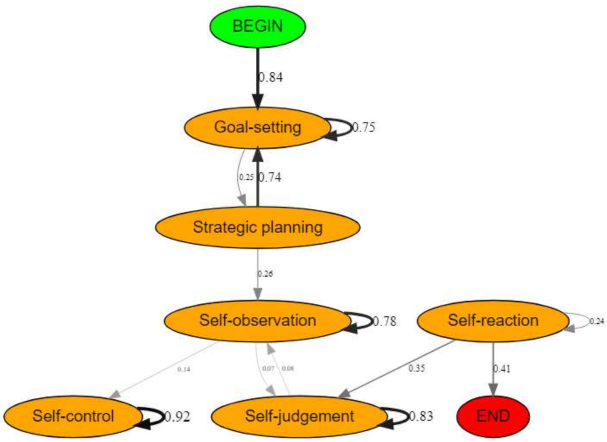

Figure 2. *First Order Markov Models (FOMMs) of Cluster 1*

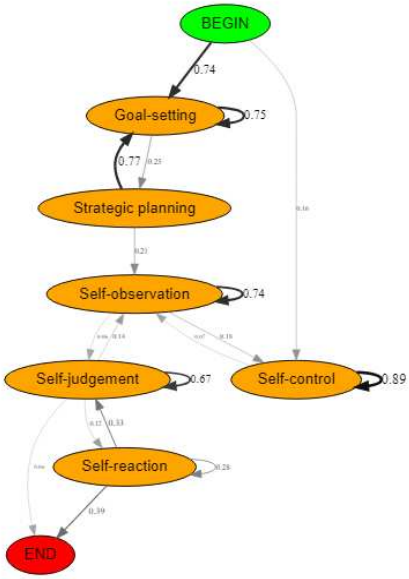

Figure 3. *First Order Markov Models (FOMMs) of Cluster 2*

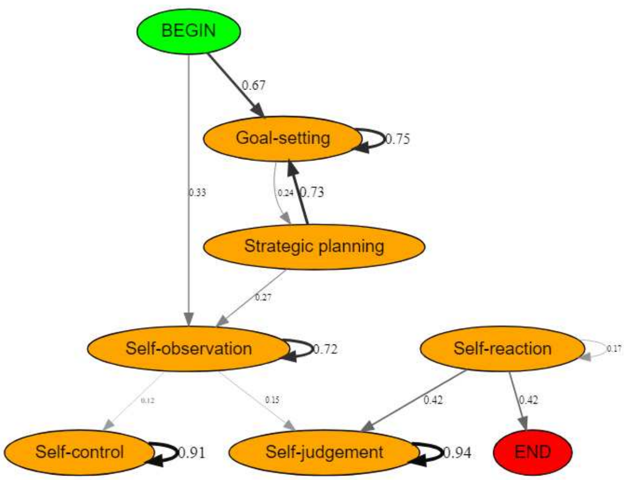

Figure 4. *First Order Markov Models (FOMMs) of Cluster 3*

Relationship between identified clusters and vocabulary learning outcomes

A non-parametric test named Kruskal Wallis Test was adopted to examine the difference of pre-vocabulary tests among students in the three identified clusters, as the pre-test was not normally distributed. The results indicated that there was no significant difference in students' prior English levels among the three identified clusters before the study.

The scores of post-vocabulary learning outcomes for all clusters are presented in Table 2. Kruskal Wallis tests showed a significant association between the identified three clusters and the students' vocabulary learning outcomes, $H\,(2) = 7.775$, $p < 0.05$. Pairwise comparisons were conducted to further examine the relationship between the identified groups and vocabulary learning outcomes. Pairwise comparisons showed significant differences in post-vocabulary test scores between students in Cluster 1 and Cluster 2 ($p < 0.05$). The students in Cluster 2 performed significantly better than students in Cluster 1. However, none of the other comparisons were significant after the Bonferroni correction (all p values >0.05).

Table 2. *Results of Kruskal Wallis tests of post-vocabulary learning outcomes*

Groups	*N*	*Mean*	*SD*	*Mean Rank*	X^2	*p*
Cluster 1	19	68.21	18.59	16.53	7.775	0.021*
Cluster 2	20	86.80	28.43	27.98		
Cluster 3	5	81.00	24.83	23.30		

Conclusions

The findings of this study were summarised into three aspects. First, this case study showed that primary students' SRVL behaviours leveraged by the mobile app with a self-regulation scheme could be clustered into three groups. Second, the similarities and differences of SRVL behavioural patterns using process-mining techniques in the three clusters were discussed. Compared to students in Cluster 1 and Cluster 3, students in Cluster 2 adopted more comprehensive SRVL learning trajectories using the mobile app with a self-regulation scheme. Third, a significant association between the identified three clusters and the students' vocabulary learning outcomes was observed, and students in Cluster 2 outperformed the other students.

This research was not without limitations. First, the study duration was short. Future studies should be conducted to understand students' SRVL behaviours over a longer period of time. Second, this study mainly relied on quantitative data. Future studies should include qualitative data (e.g., interviews) to explore the detailed characteristics of students' SRVL behaviours.

Acknowledgements

The study was funded by the General Research Fund (Ref. 18611019), Research Grants Council, University Grant Committee, Hong Kong.

References

Dinh, D. T., Fujinami, T., & Huynh, V. N. (2019). Estimating the optimal number of clusters in categorical data clustering by silhouette coefficient. In J. Chen, T. Hashimoto, X. Tang & J. Wu (Eds.), *Knowledge and systems sciences* (pp. 1-17). Springer.

Lehmann, T., Hähnlein, I., & Ifenthaler, D. (2014). Cognitive, metacognitive and motivational perspectives on preflection in self-regulated online learning. *Computers in Human Behavior*, *32*, 313-323.

Li, S., Chen, G., Xing, W., Zheng, J., & Xie, C. (2020). Longitudinal clustering of students' self-regulated learning behaviors in engineering design. *Computers & Education*, *153*. https://doi.org/10.1016/j.compedu.2020.103899

Jang, E. E., Lajoie, S. P., Wagner, M., Xu, Z., Poitras, E., & Naismith, L. (2017). Person-oriented approaches to profiling learners in technology-rich learning environments for ecological learner modeling. *Journal of Educational Computing Research*, *55*(4), 552-597. https://doi.org/10.1177/0735633116678995

Matcha, W., Gašević, D., Uzir, N. A. A., Jovanović, J., & Pardo, A. (2019, March). Analytics of learning strategies: Associations with academic performance and feedback. In D. Azcona & R. Chung (Eds.), *Proceedings of the 9th International Conference on Learning Analytics & Knowledge* (pp. 461-470). Association for Computing Machinery.

Panadero, E., Klug, J., & Järvelä, S. (2016). Third wave of measurement in the self-regulated learning field: when measurement and intervention come hand in hand. *Scandinavian Journal of Educational Research*, *60*(6), 723-735. https://doi.org/10.1080/00313831.2015.1066436

Peeters, W., Saqr, M., & Viberg, O. (2020). Applying learning analytics to map students' self-regulated learning tactics in an academic writing course. In H. -J. So, M. M. Rodrigo, J. Mason, A. Mitrovic (Eds.), *Proceedings of the 28th International Conference on Computers in Education* (Vol. 1, pp. 245-254). Asia-Pacific Society for Computers in Education.

Pintrich, P. R. (2000). The role of goal orientation in self-regulated learning. In M. Boekaerts, P. R. Pintrich & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 451-502). Elsevier.

Saint, J., Gašević, D., Matcha, W., Uzir, N. A. A., & Pardo, A. (2020, March). Combining analytic methods to unlock sequential and temporal patterns of self-regulated learning. In A. Ahmad & I. Jivet (Eds.), *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge* (pp. 402-411). Association for Computing Machinery.

Saks, K., & Leijen, Ä. (2019). The efficiency of prompts when supporting learner use of cognitive and metacognitive strategies. *Computer Assisted Language Learning*, *32*(1-2), 1-16. https://doi.org/10.1080/09588221.2018.1459729

Schmitt, N. (1999). The relationship between TOEFL vocabulary items and meaning, association, collocation and word-class knowledge. *Language Testing*, *16*(2), 189-216.

Yang, K. -H., Chu, H. -C., & Chiang, L. -Y. (2018). Effects of a progressive prompting-based educational game on second graders' mathematics learning performance and behavioral patterns. *Journal of Educational Technology & Society*, *21*(2), 322-334.

Yang, Y., Wen, Y., & Song, Y. (in press). A systematic review of technology-enhanced self-regulated language learning. *Educational Technology & Society, 26*(1), 31-44.

Zheng, L., Li, X., & Chen, F. (2018). Effects of a mobile self-regulated learning approach on students' learning achievements and self-regulated learning skills. *Innovations in Education and Teaching International, 55*(6), 616-624. https://doi.org/10.1080/14703297.2016.1259080

Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory into Practice, 41*(2), 64-70. https://doi.org/10.1207/s15430421tip4102_2