

# Discovering Students' Learning Strategies in a Visual Programming MOOC through Process Mining Techniques

Narjes Rohani<sup>1</sup>, Kobi Gal<sup>2</sup>, Michael Gallagher<sup>3</sup>, and Areti Manataki<sup>4</sup>

<sup>1</sup> Usher Institute, University of Edinburgh, Scotland, UK

<sup>2</sup> School of Informatics, University of Edinburgh, Scotland, UK

<sup>3</sup> Moray House School of Education and Sport, University of Edinburgh, Scotland, UK

<sup>4</sup> School of Computer Science, University of St Andrews, Scotland, UK

`Narjes.Rohani@ed.ac.uk`

`Kgal@inf.ed.ac.uk`

`Michael.S.Gallagher@ed.ac.uk`

`A.Manataki@st-andrews.ac.uk`

**Abstract.** Understanding students' learning patterns is key for supporting their learning experience and improving course design. However, this is particularly challenging in courses with large cohorts, which might contain diverse students that exhibit a wide range of behaviours. In this study, we employed a previously developed method, which considers process flow, sequence, and frequency of learning actions, for detecting students' learning tactics and strategies. With the aim of demonstrating its applicability to a new learning context, we applied the method to a large-scale online visual programming course. Four low-level learning tactics were identified, ranging from project- and video-focused to explorative. Our results also indicate that some students employed all four tactics, some used course assessments to strategize about how to study, while others selected only two or three of all learning tactics. This research demonstrates the applicability and usefulness of process mining for discovering meaningful and distinguishable learning strategies in large courses with thousands of learners.

**Keywords:** Process mining · Massive open online courses · Educational data mining · Visual programming · Learning tactic · Learning strategy

## 1 Introduction

The increasing use of digital learning environments enables the collection of large amounts of data, which can be analysed through Educational Process Mining (EPM) to better understand educational processes [1, 2]. A problem that has recently attracted increasing research interest in the EPM community is around detecting students' learning tactics and strategies [19, 20].

Identification of learning tactics and strategies can help customize course design, provide helpful feedback to students and assist them to adopt the best

strategies for learning [20]. A learning tactic is defined as a series of actions that a student carries out to fulfil a specific task, such as passing an exam [13, 7, 17, 21]; whereas, a learning strategy is “a coordinated set of learning tactics that are directed by a learning goal, and aimed at acquiring a new skill or gaining understanding” [17]. Identifying learning tactics and strategies is challenging, as they are invisible and latent [14]. It is even more challenging in courses with large cohorts, which may include more diverse student behaviour. Hence, appropriate analytical methods are needed, such as EPM. Most previous research that applied EPM methods to education are limited to traditional process mining methods, such as Alpha Miner, Heuristic Miner and Evolutionary Tree Miner [2, 3]. On the other hand, Matcha *et al.* [19, 20] proposed a novel EPM-based method for discovering students’ learning tactics and strategies, which combines processes flow, frequency and distribution of learning actions, thus providing a more comprehensive view of student behaviour. However, the generalisability of this method needs to be further investigated, specifically in Massive Open Online Courses (MOOCs), which are less studied. To the best of our knowledge, only one MOOC [20] has been studied with the use of this method, and it involved a student cohort that is relatively small for a MOOC.

To take a step toward addressing this gap, we apply the EPM method by Matcha *et al.* [19] to study students’ learning tactics and strategies in a large-scale visual programming MOOC with thousands of learners. The contributions of this paper are:

- We provide further evidence of the applicability of the method by Matcha *et al.* [19], by replicating their approach on large-scale data from a visual programming MOOC with thousands of students. To the best of our knowledge, this is the first time that this method is applied to such a large student cohort.
- We discover students’ learning tactics and strategies in a visual programming MOOC. This is the first time that such learning patterns are investigated in a visual programming course.

## 2 Related work

A growing number of studies have been conducted recently to analyse the educational behaviours of students, and detect their learning tactics and strategies using process mining and sequence mining [19, 16, 6]. Maldonado-Mahauad *et al.* [16] used the Process Mining  $PM^2$  method [8] on three MOOCs in engineering, education, and management. They identified seven different learning tactics, such as only-video or only-assessment. Then, by applying hierarchical clustering, they discovered three learning strategies (i.e. comprehensive, targeting, and sampling) that involved different levels of self-regulated learning. In another study, Jovanovic *et al.* [14] analysed trace data from an engineering course delivered in a flipped classroom. They discovered four learning tactics using sequence mining techniques and identified five learning strategies by applying hierarchical clustering to the tactics used by students. Fincham *et al.* [10] used the trace data

from the same course and applied a different method. Instead of sequence mining, they used process mining based on Hidden Markov models, which resulted in the identification of eight learning tactics. Then, they clustered the students based on their used tactics and identified four learning strategies in two different periods. Matcha *et al.* [19] also studied the learning tactics and strategies in the same course. They employed a combination of First-order Markov models and the Expectation-Maximization algorithm for discovering learning tactics. This novel method is capable of considering not only the process flow of learning actions, but also their distribution and frequency. By applying hierarchical clustering, they also obtained three learning strategies.

In 2020, Matcha *et al.* applied the same methodology to two additional courses: a blended learning course in biology and a Python programming MOOC [20]. In the latter, they discovered four learning tactics (Diverse-Practice, Lecture-Oriented, Long-Practice, and Short-Practice) and three learning strategies (Inactive, Highly active at the beginning, and Highly active). By using the same methodology as in their previous work, they provided evidence of the generalisability of their method. However, further research is needed in order to draw solid conclusions about its i) generalisability to different learning contexts (e.g. different course designs) and ii) its scalability to large student cohorts and datasets. This is particularly important, as the MOOC analysed in [20] had only 368 students enrolled, which is a much smaller number than the average MOOC size of thousands of learners [4].

### 3 Materials and Methods

In this paper, we applied the EPM-based method in [19] on an introductory visual programming MOOC. We utilise course assessment and clickstream data from the *"Code Yourself! An Introduction to Programming" (CDY)* MOOC, which was delivered on Coursera [5] from January 2016 to December 2017.

CDY teaches the basics of programming using Scratch, which is one of the most popular visual programming languages [23]. It covers five topics (referred to as ‘weeks’ from now on) through 71 videos, 11 reading materials, 5 weekly discussion forums, 5 weekly quizzes/exams, and 2 peer-reviewed projects (on the third and fifth week). Notably, students can submit a quiz or project multiple times, and they receive the highest achieved score among all submissions.

The CDY dataset contains information about 46,018 enrolled students (45% male, 33% female, 22% unknown) and 55,485 learning sessions. A learning session is a series of clickstream actions that a student performs within one login into the platform. In this study, the sessions that have at least one of the following actions were considered for the analysis (i.e. 37,282 sessions in total):

1. **Video-start:** Starting to watch a video for the first time
2. **Video-play:** Playing a video lecture
3. **Video-end:** Watching a video until end
4. **Video-seek:** Skipping forward or backward throughout a video
5. **Video-pause:** Pausing a video lecture

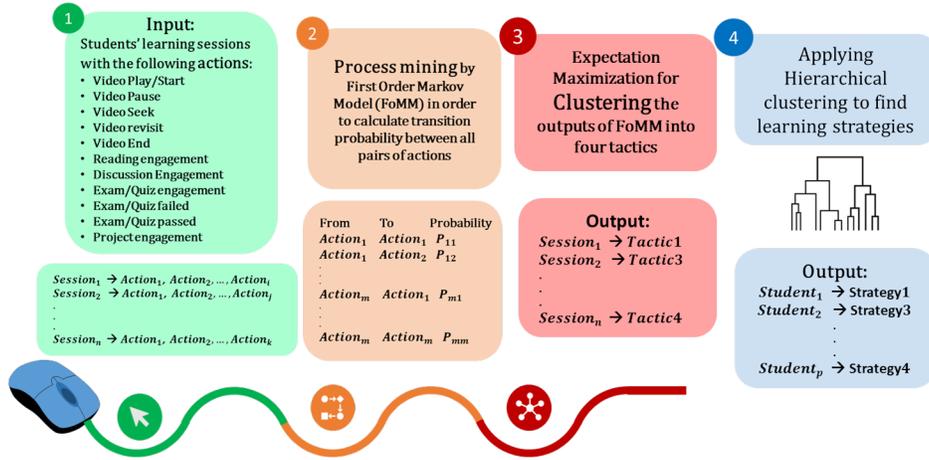
6. **Video-revisit:** Watching a video for the second time or more
7. **Reading engagement:** Any activity related to the reading material such as visiting reading pages
8. **Discussion engagement:** Any activity related to the discussion forums
9. **Exam-visit:** Visiting exam-related pages without submitting answers
10. **Exam-failed:** Failing an exam (score lower than 50% of total score)
11. **Exam-passed:** Passing an exam
12. **Peer-reviewed project engagement:** Any activity related to the peer-reviewed projects, such as submitting or reviewing a submission

### 3.1 Pre-processing

Learning sessions were profiled for each student and analysed to identify their learning tactics. We considered two consecutive sessions with a time gap less than 30 minutes as one session. Due to the high variation between session lengths (i.e. between the number of actions in sessions), very long sessions (higher than the 95th quantile) and sessions with only one action were removed to obtain a more representative dataset. Since the course is a MOOC, there are numerous participants without the intention to take the quizzes and pass the course [15, 25]. Therefore, students without any attempt to submit an assessment were removed. A same approach for pre-processing was used in related work [19, 6]. The pre-processing steps resulted in 3,190 students (sample size 8 times larger than in [19]) and 34,091 sessions. The course completion rate among the 3,190 students considered in this study was 42%.

### 3.2 Detecting Learning Tactics and Strategies through Process Mining and Clustering

To detect the learning tactics and strategies of CDY students, we followed the methodology in [19], the main steps of which are shown in Figure 1. Learning tactics were detected with the use of process mining and clustering methods. In particular, First-order Markov Models, as implemented in the *pMineR* package [12], were employed to calculate the transition probability matrix of actions. The number of possible learning tactics (no. tactics=4) was estimated based on a process flow created by first-order Markov model, Elbow method, Hierarchical clustering dendrogram, and prior contextual knowledge. To identify the learning tactics, the Expectation-Maximisation algorithm [12] was applied to the obtained transition probability matrix. To shed light on the identified learning tactics, the *TraMineR package* [11] was used for analysing the distribution, duration and the order of employed learning actions. A student may apply a range of tactics throughout a course. Therefore, a learning strategy is defined as the goal-driven usage of a collection of learning tactics with the aim of obtaining knowledge or learning a new skill [17]. To extract the various strategies adopted by students, and following methods established in related work [19, 6], we calculated the number of occurrences of each tactic used by each student and we transformed it to the standard normal distribution. Finally, the strategies



**Fig. 1.** Schema of the method: 1) Sessions with at least one coded action were selected. 2) First-order Markov Model was applied to create a process map and a transition matrix for all pairs of actions. 3) The transition matrix was used to cluster the sessions into four tactics using Expectation-Maximization method. 4) Hierarchical clustering was used to cluster students into four groups of strategies based on the frequency of their tactics.

were identified by clustering the students using Agglomerative hierarchical clustering with Ward’s linkage and Euclidean distance of the normalised vectors as the distance of students. The number of clusters (no. strategies=4) was determined based on the dendrogram analysis, Elbow method, and contextual prior knowledge.

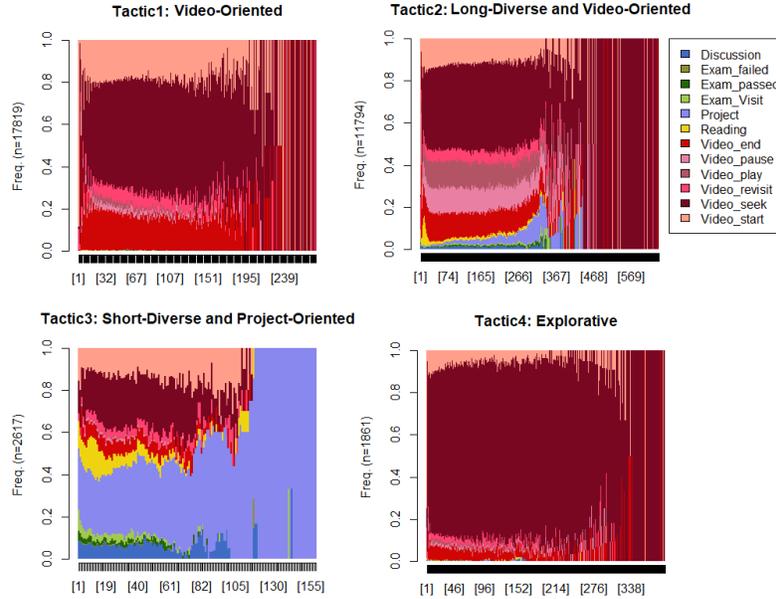
## 4 Results

Four learning tactics were discovered, which are characterised as follows.

**Tactic1: Video-Oriented** (17,819 sessions, 52.3% of all learning sessions) is the most commonly used learning tactic in CDY. It is characterized by relatively short sessions (median = 11 actions per session) that include mostly (over 99%) video-related learning actions. The high proportion of Video-end and Video-revisit actions indicate the high degree of interaction with videos (Figure 4).

**Tactic2: Long-Diverse and Video-Oriented** (11,794 sessions, 35.12% of all learning sessions) are long sessions (median = 74 actions per session) composed of diverse actions, predominantly video-related. The majority of these sessions begin with a high peak in reading- and video-related actions, followed by a peak in project engagement (Figure 4). We can infer that students employed this tactic to first gain knowledge and then do the peer-reviewed projects.

**Tactic3: Short-Diverse and Project-Oriented** (2,617 sessions, 7.7% of all learning sessions) are the shortest (median = 8 actions per session) and most diverse sessions, shaped by a wide range of learning actions and dominated by

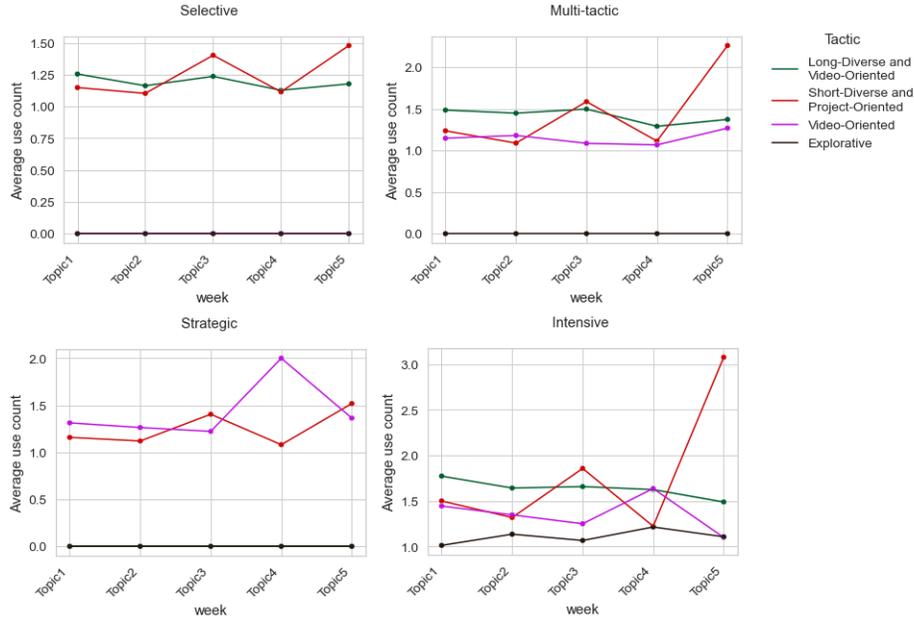


**Fig. 2.** Sequence distribution plot for the learning tactics. The X-axis presents the position of each learning action in the sessions and the Y-axis shows the relative frequency for each action in the corresponding position in the sessions. For example, the top right image for Tactic 2 shows that sessions in this cluster can contain over 500 actions. The relative frequency of reading-related actions decreases throughout these sessions, while the relative frequency of project-related actions increases.

project engagement (Figure 4). The frequency of reading- and exam-related actions is much higher in this tactic than in other tactics. Figure 4 demonstrates that most of these sessions start with understanding theoretical concepts using video and reading actions, and continue with project actions. There is also a noticeable proportion of exam-related actions in these sessions, which indicates that students not only used video and reading materials to understand the concepts, but also they engaged in quizzes for self-assessment.

**Tactic4: Explorative** (1,861 sessions, 5.4% of all learning sessions) is the least frequent learning tactic. It involves relatively long sessions (median = 22 actions per session), largely dominated by video-seeking actions (Figure 4). This indicates the exploratory behaviour of students, i.e. students may use this tactic to explore the videos or look for a specific concept.

After finding the learning tactics, four learning strategies were identified following the methodology described in Section 3. It is worth noting that similarly to other MOOCs, the average number of sessions per student is low (avg: 2) and almost all students across all strategies have a relatively low level of engage-

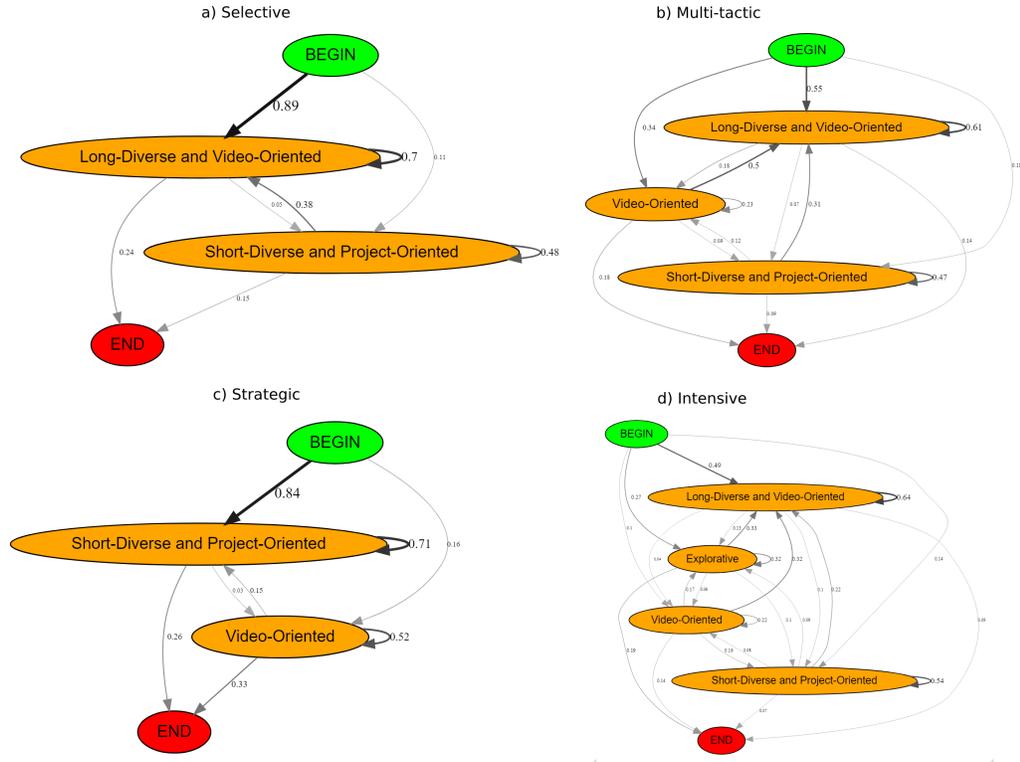


**Fig. 3.** Weekly changes in the applied learning tactics for the discovered learning strategies.

ment, especially with assessments [19]. The characteristics of the four learning strategies are as follows.

**Strategy1, Selective:** This strategy is followed by the majority of students (69.9% of students) and it is characterized mainly by using the *Long-Diverse and Video-Oriented*, and *Short-Diverse and Project-Oriented* learning tactics. In other words, this group of students are highly selective and use only two tactics. Based on the discussion in the learning tactics section, we can infer that students tend to use these two learning tactics to obtain knowledge, with the objective of answering questions in exams or doing peer-reviewed projects. Therefore, this group of students are characterized as *Selective* learners. Figure 4 indicates that the students using this strategy mainly start their learning process by a *Long-Diverse and Video-Oriented* tactic ( $p = 0.89$ ). Afterwards, they tend to keep using this tactic ( $p = 0.7$ ). The highest probable tactic to finish their learning process is also *Long-Diverse and Video-Oriented* ( $p = 0.24$ ), and the most probable transition between the two tactics is the shift to a *Long-Diverse and Video-Oriented* tactic from a *Short-Diverse and Project-Oriented* tactic ( $p = 0.38$ ).

**Strategy2, Multi-tactic:** This strategy contains 13.5% of students who used multiple learning tactics each week (Figure 3). In other words, all learning tactics except *Explorative* are employed in this strategy. Moreover, the frequency of *Long-Diverse and Video-Oriented*, and *Video-Oriented* tactics in this strategy



**Fig. 4.** Process models of the discovered learning strategies, which were created by *pMineR* package.

remains almost the same during the course, while the frequency of *Short-Diverse and Project-Oriented* fluctuates throughout the different weeks. *Multi-tactic* students mainly tend to start their week with a *Long-Diverse and Video-Oriented* ( $p = 0.55$ ) or *Video-Oriented* ( $p = 0.34$ ) tactic; either way, they tend to continue the week with a *Long-Diverse and Video-Oriented*. The most probable shifts between used tactics are the transitions from any tactic to *Long-Diverse and Video-Oriented*, underlining this tactic as the predominantly used tactic by *Multi-tactic* learners (Figure 4).

**Strategy3, Strategic:** This group contains 11.8% of all students, who mostly used the *Short-Diverse and Project-Oriented* and *Video-Oriented* learning tactics. *Short-Diverse and Project-Oriented* was mostly used at the time of submitting peer-reviewed projects, while the rest of the time these students primarily used *Video-Oriented* to learn the course materials (Figure 3). On the other hand, the process flow of these students' sessions (Figure 4) demonstrates that these students tend to start their learning process with a *Short-Diverse and Project-Oriented* tactic ( $p = 0.84$ ) and continue using this tactic until the end of the

session ( $p = 0.71$ ). The second most probable scenario is to start ( $p = 0.16$ ) and continue ( $p = 0.52$ ) to use only the *Video-Oriented* tactic with lower probability tactic. Alternatively, they might start the session with a *Short-Diverse and Project-Oriented* tactic and shift to *Video-Oriented* tactic. This strategy is named *Strategic* due to the high probability of using *Short-Diverse and Project-Oriented* tactic, which is a short tactic including a considerable number of project and exam-related actions along with the rest of the actions. Therefore, it can be inferred that the students strategically started their learning session with this tactic to achieve the required understanding for doing projects and exams.

**Strategy4, Intensive:** This is the smallest group of students (4% of all students) and they are very diligent, with relatively high engagement across all weeks (Figure 3). These students used all learning tactics every week. Although the frequency of employed tactics varies across weeks, the least and most used tactics in this strategy are the *Explorative* and *Short-Diverse and Project-Oriented*, respectively. The *Video-Oriented* tactic was primarily used in the fourth week with two drops in the third and the fifth weeks, which is similar to the frequency trend of this tactic in the *Strategic* group. The average frequency of *Explorative* and *Long-Diverse and Video-Oriented* remains fairly steady throughout the course. Figure 4 shows the process flow of this strategy, which is not as straightforward as the process flow of the other strategies. The learning process in this strategy mainly starts with a *Long-Diverse and Video-Oriented* ( $p = 0.49$ ) or an *Explorative* tactic ( $p = 0.27$ ). Irrespective of the starting tactic, students tend to shift to *Long-Diverse and Video-Oriented* and continue using it with the highest probability. The process flow also highlights the diversity of tactics used in this strategy and the fact that there is no clear structure in terms of learning tactic transitions.

It is worth mentioning that we also investigated the association between learning strategies and academic performance, and found that the learning strategies in CDY do not correlate significantly with students' assessment scores. However, the discovered learning strategies in [20] were significantly associated with student performance. An explanation for this phenomenon can be the fact that the strategies discovered in [20] are indicative of students' engagement level, and students that engage more with a course tend to perform better. The strategies discovered in this study, however, are not indicative of engagement level and are rather characterised by different combinations of tactics.

## 5 Discussion

In this study, we applied an existing EPM-based method [19] to data from a large-scale course in visual programming, and detected novel learning tactics and strategies. Our main contribution is around evidence of the applicability of this method to a different learning context, namely a visual programming MOOC with thousands of learners. Only one other MOOC has been studied with the use of this method. Another important factor is student cohort size –

our study involved 3,190 learners (after pre-processing), while the largest cohort analysed with this method in previous work was 1,135 students [20].

The learning tactics and strategies detected in our study are novel for programming and computing courses. In fact, it is the first time that such learning behaviours are investigated in a visual programming course. Most of the tactics detected include the high employment of video-related actions, which is reasonable given the high volume of video materials in the CDY course. This finding is in line with the fact that learning tactics can represent the different study approaches that are embedded in course designs and supplemented by course materials [10, 19, 18, 9, 20].

The four learning strategies discovered differ in terms of the learning tactics employed, whereas the engagement level does not vary much. However, the strategies found by Matcha *et al.* [20] were primarily focused on student engagement. In particular, most students in [20] used almost all learning tactics; therefore, clustering was based on the number of tactics used, which is an indicator of engagement. On the other hand, in CDY, clustering was based on the different combinations of tactics used. This demonstrates that the EPM method employed can effectively yield conceptually different strategies. Another advantage of this method is that it considers the process flow of learning tactics in order to group students, thus providing further insight into learning processes. Our findings indicate that the process flow of learning tactics in CDY is distinct in each group. For example, the process models of selective and strategic learners are composed of only two learning tactics; while multi-tactic and intensive students used multiple different learning tactics.

Moreover, the learning strategies extracted with the use of process mining are helpful resources for optimizing future course designs and understanding how the course design impacted the students' learning behaviour. The more insights we gain about the learning tactics and their relation to the course design, the better we can design future courses to achieve better student comprehension and fit with their learning preferences. As an example, the high rate of using the *Video-Oriented* tactic may be due to the high number of available videos in CDY. Therefore, the course design can be adjusted by supplementing more diverse resources, such as pre-lab reading, adding some programming lab notes, and making the exams or projects more interactive and attractive, so as to increase student engagement with assessments. Furthermore, informing students about their used learning strategies and other possible strategies that they can apply, can lead to better awareness and improvement of their learning approach. Also, teachers can consider students' learning strategies for providing personalized feedback [22]. For example, identifying a student that is erratic or that is only focusing on projects can help teachers provide personalized suggestions.

### 5.1 Limitations and Future Directions

The learning tactics and strategies that can be detected with the use of EPM methods are limited to the kind of data collected on the learning platform. For a programming course, it would have been interesting to also consider the

programming process, for example when attempting assignments. This was not possible in the case of CDY, but it is worth addressing in future research. Another promising avenue for future research is to combine self-declared information and trace data [24] for analysing students' educational behaviour.

There is also a great opportunity to extend this work to investigate how student's demographic features, such as gender, academic degree, and age, impact the selection of learning tactics and strategies. This is particularly interesting to examine for courses with diverse student populations, such as MOOCs.

Similarly to related work, in this study we assume that learning strategies are static. However, it is plausible that students change their learning strategy throughout a course. Future studies should relax this assumption and consider changes in learning strategies over time.

Finally, we see great value in comparing learning strategies before and after the outbreak of the Covid-19 pandemic. An interesting methodological question is to what extent the method by Matcha *et al.* [19] enables such comparisons.

## 6 Acknowledgements

This work was supported by the Medical Research Council [grant number MR/N013166/1].

## References

1. Baker, R.S.: Educational data mining: An advance for intelligent systems in education. *IEEE Intelligent systems* **29**(3), 78–82 (2014)
2. Bogarín, A., Cerezo, R., Romero, C.: A survey on educational process mining. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* **8**(1), e1230 (2018)
3. Cerezo, R., Bogarín, A., Esteban, M., Romero, C.: Process mining for self-regulated learning assessment in e-learning. *Journal of Computing in Higher Education* **32**(1), 74–88 (2020)
4. Chen, Y.H., Chen, P.J.: Mooc study group: Facilitation strategies, influential factors, and student perceived gains. *Computers & Education* **86**, 55–70 (2015)
5. Coursera: Code yourself! an introduction to programming. <https://www.coursera.org/learn/intro-programming>, accessed: June. 4, 2022
6. Crosslin, M., Breuer, K., Milikić, N., Dellinger, J.T.: Understanding student learning pathways in traditional online history courses: utilizing process mining analysis on clickstream data. *Journal of Research in Innovative Teaching & Learning* (2021)
7. Derby, S.J.: Putting learning strategies to work. *Educational Leadership* **46**(4), 4–10 (1989)
8. Eck, M.L.v., Lu, X., Leemans, S.J., Van Der Aalst, W.M.: Pm2: a process mining project methodology. In: *International conference on advanced information systems engineering*. pp. 297–313. Springer (2015)
9. Fan, Y., Saint, J., Singh, S., Jovanovic, J., Gašević, D.: A learning analytic approach to unveiling self-regulatory processes in learning tactics. In: *LAK21: 11th International learning analytics and knowledge conference*. pp. 184–195 (2021)

10. Fincham, E., Gašević, D., Jovanović, J., Pardo, A.: From study tactics to learning strategies: An analytical method for extracting interpretable representations. *IEEE Transactions on Learning Technologies* **12**(1), 59–72 (2018)
11. Gabadinho, A., Ritschard, G., Mueller, N.S., Studer, M.: Analyzing and visualizing state sequences in r with traminer. *Journal of statistical software* **40**(4), 1–37 (2011)
12. Gatta, R., Lenkowicz, J., Vallati, M., Rojas, E., Damiani, A., Sacchi, L., De Bari, B., Dagliati, A., Fernandez-Llatas, C., Montesi, M., Marchetti, A., Castellano, M., Valentini, V.: pminer: An innovative r library for performing process mining in medicine. In: ten Teije, A., Popow, C., Holmes, J.H., Sacchi, L. (eds.) *Artificial Intelligence in Medicine*. pp. 351–355. Springer International Publishing, Cham (2017)
13. Hadwin, A.F., Nesbit, J.C., Jamieson-Noel, D., Code, J., Winne, P.H.: Examining trace data to explore self-regulated learning. *Metacognition and Learning* **2**(2), 107–124 (2007)
14. Jovanovic, J., Gasevic, D., Dawson, S., Pardo, A., Mirriahi, N., et al.: Learning analytics to unveil learning strategies in a flipped classroom. *The Internet and Higher Education* **33**(4), 74–85 (2017)
15. Koller, D., Ng, A., Do, C., Chen, Z.: Retention and intention in massive open online courses: In depth. *Educause review* **48**(3), 62–63 (2013)
16. Maldonado-Mahauad, J., Pérez-Sanagustín, M., Kizilcec, R.F., Morales, N., Muñoz-Gama, J.: Mining theory-based patterns from big data: Identifying self-regulated learning strategies in massive open online courses. *Computers in Human Behavior* **80**, 179–196 (2018)
17. Malmberg, J., Järvelä, S., Kirschner, P.A.: Elementary school students’ strategic learning: does task-type matter? *Metacognition and Learning* **9**(2), 113–136 (2014)
18. Matcha, W., Gašević, D., Uzir, A., Jovanović, J., Pardo, A., Maldonado-Mahauad, J., Pérez-Sanagustín, M., et al.: Detection of learning strategies: A comparison of process, sequence and network analytic approaches. In: *European conference on technology enhanced learning*. pp. 525–540. Springer (2019)
19. Matcha, W., Gašević, D., Uzir, N.A., Jovanović, J., Pardo, A.: Analytics of learning strategies: Associations with academic performance and feedback. In: *Proceedings of the 9th International Conference on Learning Analytics & Knowledge*. pp. 461–470 (2019)
20. Matcha, W., Gasevic, D., Uzir, N.A., Jovanovic, J., Pardo, A., Lim, L., Maldonado-Mahauad, J., Gentili, S., Perez-Sanagustin, M., Tsai, Y.S.: Analytics of learning strategies: Role of course design and delivery modality. *Journal of Learning Analytics* **7**(2), 45–71 (2020)
21. Rachal, K.C., Daigle, S., Rachal, W.S.: Learning problems reported by college students: are they using learning strategies? *Journal of Instructional Psychology* **34**(4), 191–202 (2007)
22. Reimann, P., Frerejean, J., Thompson, K.: Using process mining to identify models of group decision making in chat data (2009)
23. Resnick, M., Maloney, J., Monroy-Hernández, A., Rusk, N., Eastmond, E., Brennan, K., Millner, A., Rosenbaum, E., Silver, J., Silverman, B., et al.: *Scratch: programming for all*. *Communications of the ACM* **52**(11), 60–67 (2009)
24. Ye, D., Pennisi, S.: Using trace data to enhance students’ self-regulation: A learning analytics perspective. *The Internet and Higher Education* **54**, 100855 (2022)
25. Zheng, S., Rosson, M.B., Shih, P.C., Carroll, J.M.: Understanding student motivation, behaviors and perceptions in moocs. In: *Proceedings of the 18th ACM conference on computer supported cooperative work & social computing*. pp. 1882–1895 (2015)