How do business students self-regulate their project management learning? A sequence mining study

Sami Heikkinen[1,2], Sonsoles López-Pernas[1], Jonna Malmberg[3], Matti Tedre[1] and Mohammed Saqr[1]

[1] *Faculty of Science and Forestry, School of Computing, University of Eastern Finland, Joensuu, Finland*
[2] *LAB University of Applied Sciences, Lahti, Finland*
[3] *Department of Educational Sciences and Teacher Education at the University of Oulu, Oulu, Finland*

Abstract
The relation between learning strategies and academic achievement has been proven to be strong in multiple studies. Still, the connection between micro-level SRL processes and the academic achievement of business students in learning project management remains unstudied. The current study aims to find how sequence mining can identify students using different learning tactics and strategies in terms of micro-level SRL processes. Our findings show that there are differences in the use of tactics and strategies between low and high performing students. Understanding the differences in how low and high performing students apply different micro-level SRL processes can help practitioners identify students in need of support for SRL.

Keywords
Sequence mining, micro-level SRL processes, learning tactics, academic achievement, learning analytics, project management

1. Introduction

To succeed in online learning, students need to possess self-regulated learning (SRL) skills. SRL is a dynamic process where the students set goals for their learning, monitor their progress and respond to the challenges during their learning [23]. Different models [1,4,8,23,30,31] describe the processual nature of SRL. Although the different models have distinctive features, the division into three phases (planning, performance, and reflection) is a common characteristic of all models [21]. The three phases can be divided into micro-level SRL processes, which differ between models. The planning phase includes micro-level SRL processes such as task analysis [31] and goal setting [1,30]. The student monitors and controls learning [7,22] in the performance phase with different tactics and strategies [30]. Once the learning task is finished, the SRL cycle ends with reflection, including, e.g., self-judgement, to improve learning in the subsequent SRL cycles [31].

Learning analytics (LA) can be used to track the learning processes in online learning. LA is "the measurement, collection, analysis and reporting of data about students and their contexts for purposes of understanding and optimising learning and the environments in which it occurs" [27]. The clickstream data from learning management systems (LMS) can illustrate the micro-level SRL processes [26]. Learning tactics are sequences of actions students perform during learning [9], whereas strategies are based on patterns of tactics students choose. For example, Siadaty et al. [25] have given an example of how to recode the trace data captured from LMS to show the different micro-level SRL processes students perform during online learning. Uzir et al. [28] have studied how

Proceedings of the Finnish Learning Analytics and Artificial Intelligence in Education Conference (FLAIEC22), Sept 29–30, 2022, Joensuu, Finland
EMAIL: sami.heikkinen@lab.fi (A.1); sonsoles.lopez.pernas@upm.es (A.2); jonna.malmberg@oulu.fi (A.3); matti.tedre@uef.fi (A.4); mohammed.saqr@uef.fi (A.5)
ORCID: 0000-0002-2283-1345 (A.1); 0000-0002-9621-1392 (A.2); 0000-0002-8890-4068 (A.3); 0000-0003-1037-3313 (A.4); 0000-0001-5881-3109 (A.5)

trace data can demonstrate students' time management strategies during the blended learning process and how these strategies associate with academic achievement. In addition, decades of SRL research have shown the relationship between learning strategies and academic achievement [5,17].

López-Pernas and Saqr [14] reviewed different learning tactics with the help of multichannel data to get a holistic view of students' choices. The same data was used by López-Pernas et al. [15] in a study that focused on students struggling with their assignments. This study sought to understand what tactics students use to overcome their challenges. Saqr and López-Pernas [24] have extended the timespan to the entire degree of studies to research the engagement modes of the students. Sequence mining is commonly studied using R with TraMineR and seqHMM [14,24]; pMineR and rENA [28]; BupaR [15,28] packages.

In project management education, a need for SRL is recognised, but the ways to support SRL are lacking [16,20]. LA is used to assess and predict teamwork in the context of software engineering [22], customise scaffolding to automate reflection and feedback loops for virtual business projects [12], and automate the feedback process to help students achieve better grades in engineering education [18,19]. In business management, social network analysis has been used to explore how social factors influence performance and learning [2]. The field of project management, especially in business, requires a better understanding of how SRL could be supported using LA.

1.1. Purpose and aims of the study

Although there are various approaches in the way sequence mining is executed to study learning, the relation between micro-level SRL processes and how to capture it with trace data is not studied. The current study aims to find how sequence mining can identify students using different learning tactics in terms of micro-level SRL processes. This study aims to find an approach that can be used to help low-performing students improve their SRL and thus achieve better learning outcomes.

RQ1: Which micro-level SRL processes do students use for learning project management in LMS?

RQ2: What type of distinct groups of students can be found based on students' use of micro-level SRL processes, and how do they relate to academic achievement?

2. Methods

2.1. Study context

This study was conducted at the LAB University of Applied Sciences. The course participants (n = 96) were first-year undergraduate business students taking part in a project management course arranged entirely online. Only the students who gave their informed consent for research purposes were included in this study. The extent of the course was 5 ECTS (European Credit Transfer System, c. 130 hours of student workload). The course included two topics: creative problem solving and project planning. Both topics are divided into different themes, which follow the chronological order of creative problem solving (i.e., identifying problem, gathering information, creating ideas, and evaluation of ideas) and project planning (i.e., scope and work packages, schedule, budgeting, and compiling project plan document). The course implementation was organised in the autumn semester of 2021. The course started in September and lasted until December for 14 and a half weeks.

The course materials were available for students via the Moodle LMS (learning management system). The learning materials were distributed in video format. There were in total 17 videos which were presenting the topics of the course and the procedures students were expected to follow once working with the course assignments. There were ten assignments, one assignment for each topic of the course. In the first assignment, students chose the problem they worked with during the course. For each assignment, the instructions display the requirements of an accepted deliverable, thus enabling the student to self-assess the output before submitting it.

The course followed the principles of formative assessment, where students should determine their own learning goals. This was done by asking the students to set learning goals in terms of the final grade in the first assignment. In addition, students reflected on their learning goals and learning process in the last assignment. The course's final grade was based on the number of assignments the student completed following the criteria of an acceptable deliverable. All assignments were done individually, and no collaborative activities were included in the course. The minimum requirement for passing the course was submitting three acceptable deliverables.

Along with learning materials and assignments, the course platform included a section for frequently asked questions (FAQ) for troubleshooting; additional material for students who want to dive deeper into the course topics.

2.2. Data sources

There were altogether two data sources from each student: 1) trace data from the Moodle LMS and 2) final grades. LMS data included timestamp of performed actions, user ID, course module ID, and description of learning activity. The final grades achieved were used as an indicator of course performance. The course was graded using a five-tier numeric scale (0 = failed, 5 = excellent).

2.3. Data preparation

LMS data was cleaned and recoded to enable the analysis. First, the actions performed by the teacher were removed. Second, the student IDs were anonymised, and events involving students who did not give their consent to use their data were removed. Third, the actions (e.g., user list viewed) with few instances were removed. Fourth, the event context details were split into two columns, of which 1) recoded to follow the numeric order (1, 2, 3, ..., 10) of the course topics and 2) was the headline of the course topic. Finally, the LMS events were recoded into micro-level SRL processes [25] following the coding plan displayed in Table 1.

Table 1
Recoding LMS events to micro-level SRL processes.

Micro-level SRL processes	LMS event in Moodle
Task analysis	Course module viewed Recent activity viewed Discussion viewed Course viewed The status of the submission has been viewed
Goal setting	Within 1st assignment: - Submission form viewed - Submission created (online text uploaded / a file has been uploaded) - A submission has been submitted - Submission updated
Performing	Course module viewed (pages) and nice to know -materials Within 2^{nd} – 10^{th} assignment - Submission form viewed - Submission created (online text uploaded / a file has been uploaded) - A submission has been submitted - Submission updated Quiz attempt
Reflection	Feedback viewed Course activity completion updated Course user report viewed Grade overview report viewed Grade user report viewed

The recoded trace data were aligned in chronological order. Based on the ordered data, the events were grouped into sessions. The sessions are identified based on the interval between LMS actions. There is no consensus on the optimal interval, and instead, it should be decided considering the course content. In this study, 30 minutes was used as an interval between two actions to consider them as belonging to the same session [13]. This procedure resulted in some sessions which included only one action. These sessions were removed as outliers since they cannot be analysed as manifestations of learning patterns. Also, the learning sessions longer than the 90^{th} percentile of the learning sessions were trimmed

2.4. Data analysis

The data were analysed with sequence mining methods to find how students approach the learning processes. The micro-level learning processes were used to understand the different sequences students take when engaged in the learning process.

2.4.1. Identification of micro-level SRL processes

To answer the first research question, we applied clustering to detect learning sessions with similar patterns. We built a sequence object for each of the sessions identified, containing the chronologically ordered events using the TraMineR R package [6]. The sequences were clustered using Agglomerative Hierarchical Clustering (AHC) and Ward's algorithm. This method has been used previously by [13–15,28].

2.4.2. Relation between course performance and micro-level SRL processes

To answer the second research question, we clustered the students based on the combination of micro-level SRL processes they used during the course using latent profile analysis [10]. Each distinct combination of micro-level SRL processes is referred to as a learning tactics cluster. The relation between clusters and academic achievement was tested using the Games-Howell test accompanied by the Welch's and Holm's tests [7,11,29].

3. Results

There were altogether 26,930 user actions performed during the online course. The distribution of micro-level SRL processes is displayed in Table 2. The micro-level SRL process students use most often is *task analysis*. It is followed by *performing*, while *reflection* and *goal setting* are seldomly used micro-level SRL processes.

Table 2
Distribution of micro-level SRL processes

Micro-level SRL process	n	per cent
Goal setting	1883	6.99%
Performing	4291	15.93%
Reflection	2228	8.27%
Task analysis	18528	68.80%

3.1. Identification of micro-level SRL processes

The overall distribution plot of students' micro-level SRL processes (Figure 1) displays the sequences (n = 2,902) of LMS data. The X-axis describes the order of different micro-level SRL processes in the learning sessions, and on the Y-axis, the proportion of each micro-level SRL process at each step of the learning sessions. The plot shows that students often analyse the tasks, which is the dominating micro-level SRL process throughout the learning sessions. This is most often the state which students start their learning sessions with. In the second step, students either reflect on their learning or focus on goal setting. Starting from the third step of the learning session, students shift to the performing phase. The frequency of *performing* actions increases during the learning sessions; meanwhile, the *reflection* and *task analysis* activities decrease.

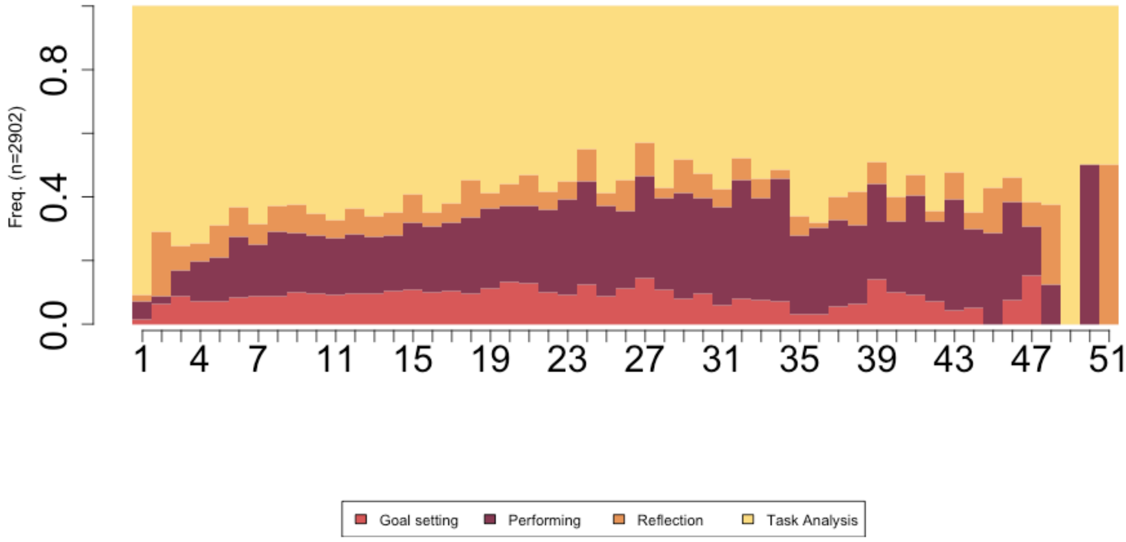

Figure 1: Overall distribution plot of students' micro-level SRL processes

The session length varies a lot. A sequence distribution plot acknowledging the session length (Figure 2) shows that half of the sessions include five or fewer actions taken by students. Here, the *task analysis* dominates, whereas *reflection* has strong distribution in step two. This is followed by an increase in *performing*.

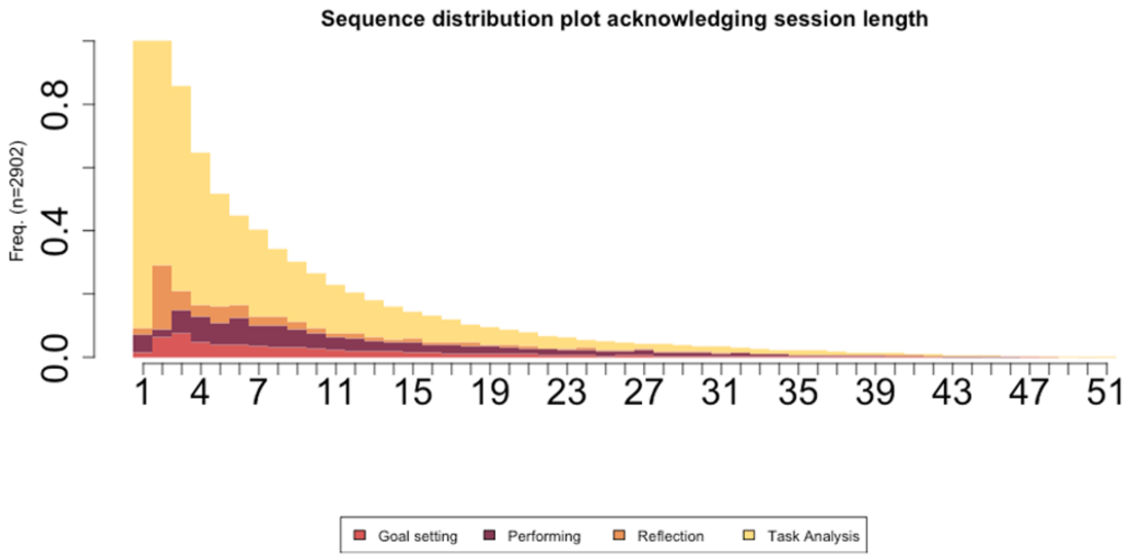

Figure 2: Sequence distribution plot acknowledging session length

The clustered sequence distribution (Figure 3) shows that the learning sessions can be divided into three distinct clusters. The X-axis describes the order of different micro-level SRL processes in the learning sessions, and on the Y-axis, the proportion of each micro-level SRL process at each step of the learning sessions. For identifying learning sessions with similarities, we used AHC. The most distinctive feature is the length of the sequences: 2 for the task analysing tactic, 22 for the short-focused tactic, and 48 for the long-range tactic.

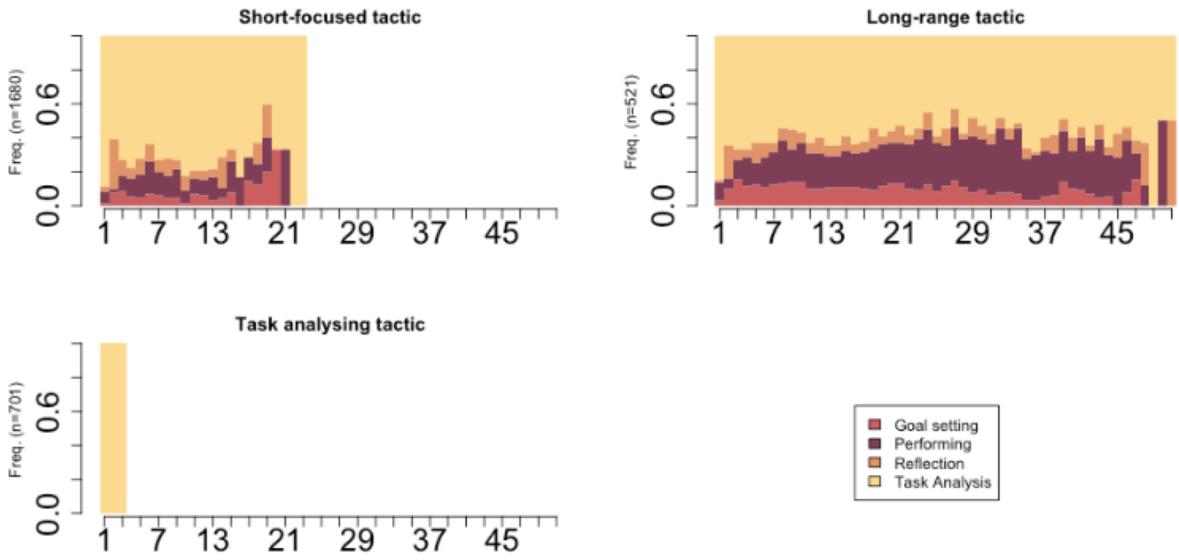

Figure 3: Sequence distribution plot of students' micro-level SRL processes within each cluster

Short-focused tactic (n = 1,680) is the most used tactic. The length of this type of tactic is intermediate. *Task analysis* is the dominating micro-level SRL process. Starting from the second step of the sequence, the proportion of *reflection* activities increases, followed by the rise in the ratio of the *performing* activities.

Long-range tactic (n = 521) is the least often used. It is the tactic with most actions taken resulting in the most extended sequence. The distribution of different micro-level SRL processes is mostly balanced; *task analysis* is most often a micro-level SRL process, but other micro-level SRL processes are present. The performing micro-level processes are strongly present when compared to two other tactics. More effort is also put into *goal setting* and *reflection*.

Task analysing tactic (n = 701) is the shortest tactic (maximum length of two steps), focusing solely on *task analysis*.

3.2. Relation between course performance and micro-level SRL processes

We did clustering using latent profile analysis to classify students according to the number of used tactics. Three learning tactics clusters of students were found.

Engaged (n = 38) students have the highest number of each tactic used. The short-focused tactic is used the most, followed by task-analysing and long-range tactics.

Moderate (n = 46) students have the same kind of distribution between tactics. Here the proportion of the short-focused tactic is the highest, while long-range and task-analysing tactics are relatively less in use.

Disengaged (n = 12) students have a deficient number of tactics in use. Here the short-focused and task-analysing tactics are almost on the same level, whereas the long-range tactic is barely used (Figure 4).

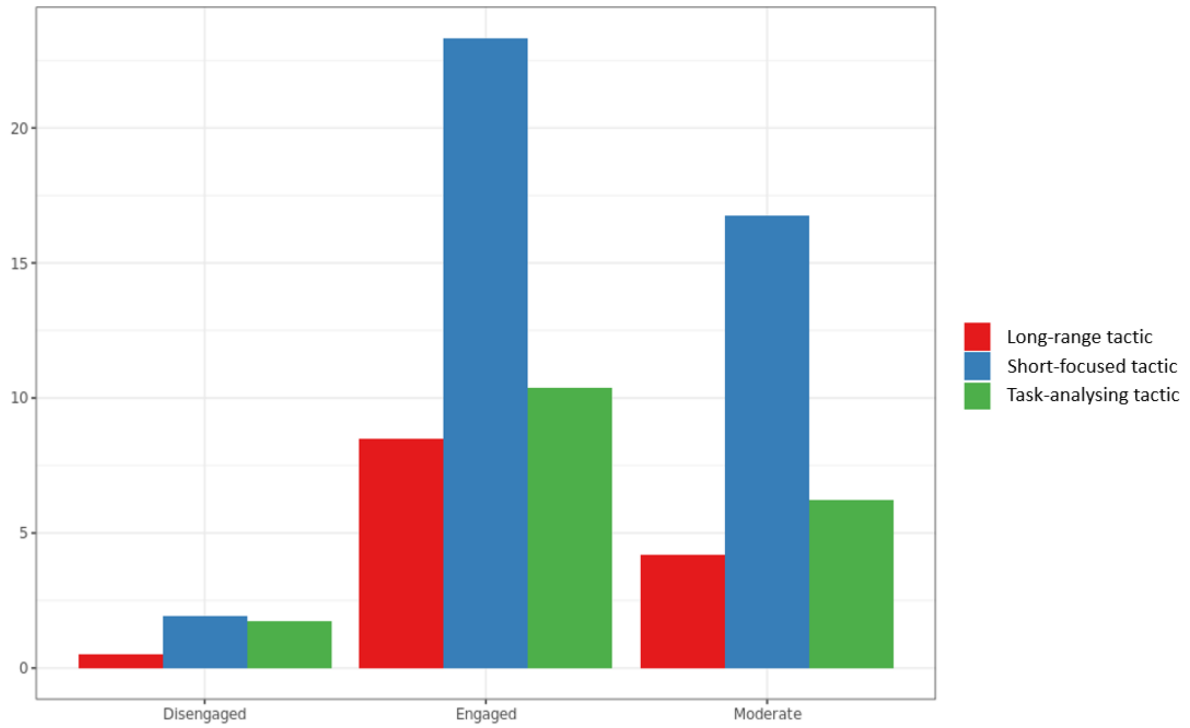

Figure 4: The learning tactics clusters and the number of times each tactic is used per cluster.

As shown in Figure 5, students in the engaged tactics cluster (n = 38) achieved the highest grades (mean 4.00). Students in the moderate tactics cluster (n = 46) achieved mediocre grades (mean 2.43), whereas students in the disengaged tactics cluster (n = 12) were likely to fail (mean 0.17) the course. The differences between every cluster were statistically significant.

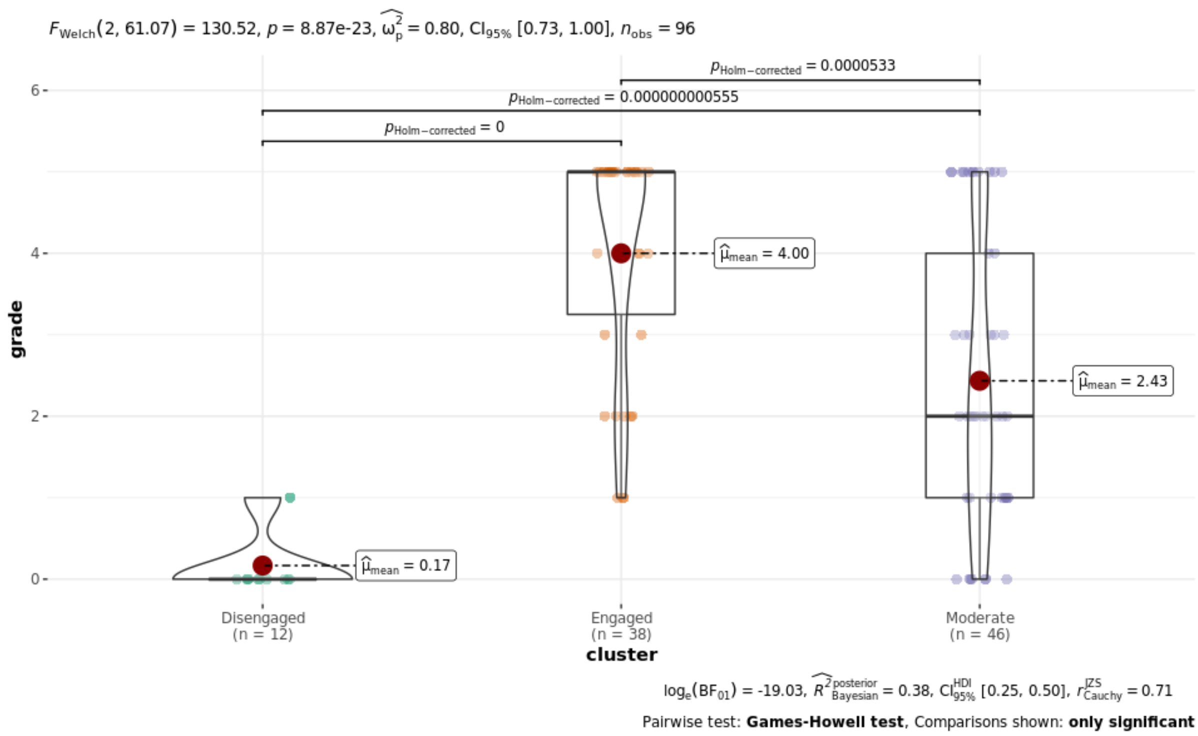

Figure 5: Violin plot compares the grades with tactics clusters using the Games-Howell test and Holm's and Welch's tests.

4. Discussion

According to the current study results, there are differences in the use of SRL tactics between low and high performing students. The engaged students apply SRL tactics to a much greater extent than disengaged students. This finding is in line with the results of previous studies [5,17]. This study sheds light on how business students use tactics when learning project management skills. The findings are similar to the ones previously found in the different disciplines and ascertain the universal nature of SRL between disciplines. Understanding the differences in how low and high performing students apply different micro-level SRL processes can help practitioners identify students in need of support for SRL.

According to the current study results, disengaged students need support in online learning. It might be that they could not figure out how to start working with the assignment, and there was no intervention available at the right time. This finding sets a need for future research. In order to understand the needs thoroughly, the learners' perspectives must be studied with qualitative methods. The ways to support disengaged students should be found. There is a need for interventions that build on the information provided by LA. The first steps of online learning are the crucial part of the learning path that require support. This current study presents the situation of a single implementation of a course. The results of this study must be verified by increasing the number of students involved in the study. With these steps, the learning processes of the business students learning project management can be improved.

Our future research will focus on improving the methods by using a more granular coding of learning activities that describe their project management activities with more details and use a two-step clustering approach to chart the pathway of learning strategies similar to [13]. A possible direction would be to combine analytics methods, e.g., process mining and social network analysis, to obtain a more nuanced and multi-faceted view of the self-regulation process. Another possible direction would be to chart students' pathways throughout the program in order to study the longitudinal pathway of students through the program, such as in the work of [23].

References

[1] M. Boekaerts, L. Corno, Self-regulation in the classroom: A perspective on assessment and intervention, Applied Psychology. 54 (2005) 199–231.

[2] K.S.K. Chung, W.C. Paredes, Towards a social networks model for online learning & performance, Educational Technology and Society. 18 (2015).

[3] S.J. Derry, Putting Learning Strategies to Work., Educational Leadership. 46 (1988).

[4] A. Efklides, Interactions of metacognition with motivation and affect in self-regulated learning: The MASRL model, Educational Psychologist. 46 (2011) 6–25.

[5] E. Fincham, D. Gašević, J. Jovanović, A. Pardo, From Study Tactics to Learning Strategies: An Analytical Method for Extracting Interpretable Representations, IEEE Transactions on Learning Technologies. 12 (2019) 59–72.

[6] A. Gabadinho, G. Ritschard, N.S. Müller, M. Studer, Analyzing and visualizing state sequences in R with TraMineR, Journal of Statistical Software. 40 (2011).

[7] P.A. Games, H.J. Keselman, J.J. Clinch, Tests for homogeneity of variance in factorial designs, Psychological Bulletin. 86 (1979).

[8] A.F. Hadwin, S. Järvelä, M. Miller, Self-regulated, co-regulated, and socially shared regulation of learning, in: B.J. Zimmerman, D.H. Schunk (Eds.), Handbook of Self-Regulation of Learning and Performance, Routledge, New York, NY, 2011: pp. 65–84.

[9] A.F. Hadwin, J.C. Nesbit, D. Jamieson-Noel, J. Code, P.H. Winne, Examining trace data to explore self-regulated learning, Metacognition and Learning. 2 (2007).

[10] M. Hickendorff, P.A. Edelsbrunner, J. McMullen, M. Schneider, K. Trezise, Informative tools for characterizing individual differences in learning: Latent class, latent profile, and latent transition analysis, Learning and Individual Differences. 66 (2018).

[11] S. Holm, A simple sequentially rejective multiple test procedure, Scandinavian Journal of Statistics. 6 (1979).

[12] N. James, A. Humez, P. Laufenberg, Using Technology to Structure and Scaffold Real World Experiential Learning in Distance Education, TechTrends 2020 64:4. 64 (2020) 636–645.

[13] J. Jovanović, D. Gašević, S. Dawson, A. Pardo, N. Mirriahi, Learning analytics to unveil learning strategies in a flipped classroom, Internet and Higher Education. 33 (2017).

[14] S. López-Pernas, M. Saqr, Bringing synchrony and clarity to complex multi-channel data : A learning analytics study in programming education, IEEE Access. 9 (2021) 166531–166541.

[15] S. López-Pernas, M. Saqr, O. Viberg, Putting it all together: Combining learning analytics methods and data sources to understand students' approaches to learning programming, Sustainability (Switzerland). 13 (2021) 4825.

[16] S. Marcelino-Sádaba, A. Perez-Ezcurdia, Competence training for project management: holistic analysis framework, in: Handbook of Research on Project Management Strategies and Tools for Organizational Success, 2020: pp. 196–222.

[17] W. Matcha, D. Gašević, N.A. Uzir, J. Jovanović, A. Pardo, J. Maldonado-Mahauad, M. Pérez-Sanagustín, Detection of Learning Strategies: A Comparison of Process, Sequence and Network Analytic Approaches, in: Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2019.

[18] I. Menchaca, M. Guenaga, J. Solabarrieta, Using learning analytics to assess project management skills on engineering degree courses, in: ACM International Conference Proceeding Series, 2016.

[19] I. Menchaca, M. Guenaga, J. Solabarrieta, Learning analytics for formative assessment in engineering education, International Journal of Engineering Education. 34 (2018) 953–967.

[20] A.G. de Oliveira Fassbinder, M. Fassbinder, E.F. Barbosa, G.D. Magoulas, Massive open online courses in software engineering education, Proceedings - Frontiers in Education Conference, FIE. 2017-October (2017) 1–9.

[21] E. Panadero, A review of self-regulated learning: Six models and four directions for research, Frontiers in Psychology. 8 (2017).

[22] D. Petkovic, Using Learning Analytics to Assess Capstone Project Teams, Computer (Long Beach Calif). (2016) 80–83.

[23] P.R. Pintrich, The role of goal orientation in self-regulated learning, in: M. Boekaerts, P.R. Pintrich, M. Zeidner (Eds.), Handbook of Self-Regulation, Academic Press, San Diego, CA, 2000: pp. 452–502.

[24] M. Saqr, S. López-Pernas, The longitudinal trajectories of online engagement over a full program, Computers and Education. 175 (2021) 104325.

[25] M. Siadaty, D. Gašević, M. Hatala, D. Gasevic, M. Hatala, Associations between technological scaffolding and micro-level processes of self-regulated learning: A workplace study, Computers in Human Behavior. 55 (2016) 1007–1019.

[26] M. Siadaty, D. Gašević, M. Hatala, D. Gasevic, M. Hatala, Trace-Based Microanalytic Measurement of Self-Regulated Learning Processes, Journal of Learning Analytics. 3 (2016) 183–220.

[27] G. Siemens, P. Long, Penetrating the Fog: Analytics in Learning and Education, EDUCAUSE Review. 46 (2011).

[28] N.A. Uzir, D. Gašević, J. Jovanović, W. Matcha, L.A. Lim, A. Fudge, Analytics of time management and learning strategies for effective online learning in blended environments, in: Proceedings of the 10th International Conference on Learning Analytics and Knowledge (LAK '20), March 23-27, Frankfurt, Germany, ACM, New York, NY, 2020.

[29] B.L. Welch, The generalisation of student's problems when several different population variances are involved., Biometrika. 34 (1947).

[30] P.H. Winne, A.F. Hadwin, Studying as self-regulated engagement in learning, in: D. Hacker, J. Dunlosky, A. Graesser (Eds.), Metacognition in Educational Theory and Practice, Erlbaum, Hillsdale, NJ, 1998: pp. 277–304.

[31] B.J. Zimmerman, Attaining self-regulation: a social cognitive perspective, in: M. Boekaerts, P.R. Pintrich, M. Zeidner (Eds.), Handbook of Self-Regulation, Academic Press, San Diego, CA, 2000: pp. 13–40.