

PUBLICATIONS OF
THE UNIVERSITY OF EASTERN FINLAND

**Dissertations in Science,
Forestry and Technology**



UNIVERSITY OF
EASTERN FINLAND

SAMI HEIKKINEN

**Using Learning Analytics
to Capture Self-Regulation
and the Complex Temporal
Dynamics of its Correlates**

**Using Learning Analytics to Capture Self-
Regulation and the Complex Temporal
Dynamics of its Correlates**

Sami Heikkinen

Using Learning Analytics to Capture Self-Regulation and the Complex Temporal Dynamics of its Correlates

Publications of the University of Eastern Finland
Dissertations in Science, Forestry and Technology
No 94

University of Eastern Finland
Joensuu
2025

Academic dissertation

To be presented by permission of the Faculty of Science, Forestry and
Technology for
public examination in the Auditorium AU100 in the Aurora Building at
the University of Eastern Finland, Joensuu, on June 18, 2025,
at 12 o'clock noon

Punamusta

Joensuu, 2025

Editors: Pertti Pasanen, Nina Hakulinen, Raine Kortet,
Matti Tedre, Pasi Vahimaa and Timo Lähivaara

ISBN: 978-952-61-5624-8 (Print)

ISBN: 978-952-61-5625-5 (PDF)

ISSNL: 2954-131X

ISSN: 2954-131X (Print)

ISSN: 2954-1484 (PDF)

Author's address:

Sami Heikkinen
University of Eastern Finland
School of Computing
P.O. Box 111
80101 JOENSUU, FINLAND
email: samp@uef.fi

Supervisors:

Associate Professor Mohammed Saqr, Ph.D.
University of Eastern Finland
School of Computing
P.O. Box 111
80101 JOENSUU, FINLAND
email: mohammed.saqr@uef.fi

Assistant Professor Jonna Malmberg, Ph.D.
University of Oulu
Depart. of Educational Sciences and Teacher
Education
P.O. Box 2000
90014 OULU, FINLAND
email: jonna.malmberg@oulu.fi

Professor Matti Tedre, Ph.D.
University of Eastern Finland
School of Computing
P.O. Box 111
80101 JOENSUU, FINLAND
email: matti.tedre@uef.fi

Reviewers:

Associate Professor Rwitajit Majumdar, Ph.D
Kumamoto University
Research and Educational Institute for
Semiconductors and Informatics
KUMAMOTO, JAPAN
email: majumdar@kumamoto-u.ac.jp

Academy Research Fellow Juho Leinonen, Ph.D
Aalto University
Depart. of Computer Science
P.O. Box 11000
00076 AALTO, FINLAND
email: juho.2.leinonen@aalto.fi

Opponent:

Professor Hanni Muukkonen, Ph.D
University of Oulu
Faculty of Education and Psychology
P.O. Box 2000
90014 OULU, FINLAND
email: hanni.muukkonen@oulu.fi

Heikkinen, Sami

Using Learning Analytics to Capture Self-Regulation and the Complex Temporal Dynamics of its Correlates.

Joensuu: Itä-Suomen yliopisto, 2025

Publications of the University of Eastern Finland

Dissertations in Science, Forestry and Technology; 94

ISBN: 978-952-61-5624-8 (Print)

ISBN: 978-952-61-5625-5 (PDF)

ISSNL: 2954-131X

ISSN: 2954-131X (Print)

ISSN: 2954-1484 (PDF)

ABSTRACT

This thesis investigates the complex interplay between self-regulated learning, student engagement, and learning analytics in online educational environments. Through three related studies, the research examines how learning analytics can support self-regulated learning, how students' self-efficacy beliefs influence their learning behaviors, and how engagement patterns evolve alongside self-regulation throughout online courses.

The research employs a mixed-methods approach that combines systematic literature review, sequence analysis, process mining, and multi-channel clustering techniques. Data were collected from multiple sources including learning management system trace data, self-reported surveys, and longitudinal self-regulated learning measures across various cohorts of business students.

Key findings reveal that while learning analytics interventions show promise in supporting self-regulated learning, only 46% demonstrated positive impacts on learning outcomes. The research identified distinct patterns in how students with different self-efficacy beliefs approach online learning, with high self-efficacy students demonstrating more varied and

adaptive learning tactics. Analysis of engagement patterns revealed three distinct clusters of students: low-regulating improvers (4.85%), disengaged low regulators (44.85%), and high-regulating improvers (50.30%), each showing different trajectories in their self-regulation development.

The thesis makes several theoretical and practical contributions to the field of educational technology. It elucidates how self-regulation and engagement co-evolve in online learning environments, provides evidence-based recommendations for designing learning analytics interventions, and demonstrates the value of combining multiple analytical approaches to understand student learning processes. These findings have important implications for the design of online courses and learning support systems, particularly their adaptation to support students with varying levels of self-regulation skills and engagement patterns.

Keywords: Learning analytics, self-regulated learning, student engagement, online education, sequence analysis, process mining, self-efficacy

Acknowledgements

Learning analytics is an interesting topic. This was the initial idea I had when applying for PhD studies at the University of Eastern Finland. Studying remotely is not easy. When there is no home base, self-regulated learning and self-efficacy beliefs are required to overcome this challenge. Along with these essential skills, one needs support from fellow people.

I want to thank my main supervisor, Associate Prof. Mohammed Saqr, for his vision and support in each step of my journey. I have heard so many horror stories about supervisors who do not respond for weeks or months. This was definitely not the case with Mohammed. I am grateful for everything I have learned from him and every nudge he gave me when I did not realize I would need them.

Furthermore, I want to thank my supervisor, Assistant Prof. Jonna Malmberg, who guided me into the key concepts of self-regulated learning and self-efficacy beliefs. I was lost while writing the second article, but with the help of Jonna, I was once again found. Furthermore, I want to thank my supervisor, Prof. Matti Tedre, whose guidance was invaluable in the first steps of my PhD journey. With his tips, I got the bird's eye perspective on the journey. Online PhD school was one of the most influential links, which I have been sharing with almost everyone considering their first steps on this path that I am about to complete.

I want to express my gratitude to the pre-examiners, Associate Prof. Rwitajit Majumdar and Dr. Juho Leinonen, for their thoughtful comments on my thesis. Their feedback has improved my work. I am also deeply grateful to Prof. Hanni Muukkonen for accepting the opponent position for my dissertation defense.

Other important persons from the School of Computing have helped during this journey. Dr. Jarkko Suhonen has helped me by answering any questions I have. As a remote student, I have been far from my fellow students. Still, Dr. Ramy Elmoazen has answered all the questions I have had

during my journey. I also want to thank Juha Hakkarainen for taking care of the virtual environment, where I could do my coding and analysis.

I have been blessed to have the opportunity to cooperate with and learn from experienced researchers during my studies. Dr. Sonsoles López-Pernas, thank you for helping me get started with R and sequence analysis methods. With my second article, I got to expand the network to the Netherlands. Along with Dr. Tudor Cristea, we worked on two articles – one that is a part of Tudor's dissertation and another that is the second article of my dissertation. I am also grateful to Tudor's supervisors Prof. Chris Snijders, Assistant Prof. Ad Kleingeld, and Associate Prof. Uwe Matzat, for the insightful discussions and their hospitality in Eindhoven.

I must thank all the students who were willing to share their multi-channel learning sequences with me. Furthermore, I must thank the peer reviewers and editors of the journals that have published our articles for your commitment to improving the quality of our output.

I want to thank LAB University of Applied Sciences for the opportunity to conduct this research along with my daily work. The LAB-D programme has given me the opportunity to share experiences with my peer colleagues who have been going through their steps in an academic endeavour. This has been a great support while not being present in the alma mater scientific community. LAB has also allowed the collection of the data needed for the empirical research. My supervisor, Pirjo Haverinen, has been a great support in tackling the daily challenges that have given me more possibilities to carry on researching.

Iso merkitys tämän väitöskirjan valmiiksi saattamiselle on myös vanhemmillani. Ensituntumani tietokoneisiin olen saanut Commodore64®:n välityksellä. Kuulemani mukaan kirjoitin Kuusnepan komentokehotteeseen käskyjä pelien lataamiseksi, vaikka en vielä lukea osannutkaan. Mahdollisuus päästä käsiksi teknologiaan jo varhain on varmasti ollut yksi syy sille, että en ole pelännyt tutkia tietokoneiden tarjoamia mahdollisuuksia. Kiitos siis siitä ympäristöstä, jossa olen saanut mahdollisuuden kasvattaa uteliaisuuttani.

Several colleagues, friends, relatives, and fellow citizens have asked what my PhD is all about. Thank you for your interest and questions. Every time I have tried to answer your questions, I have understood my topic a bit better.

Finally, I want to thank Hanna and Lahja for their support and understanding. I have read many PhD dissertation acknowledgements and found the words about absent minds. I acknowledge that my mind has been wandering in scientific dreams. Sorry for that; thanks for your patience and support. We made it!

Lahti, 13 May 2025

Sami Heikkinen

LIST OF ABBREVIATIONS

AHC	agglomerative hierarchical clustering
AI	artificial intelligence
BIC	bayesian information criterion
DFG	directly-follows graph
ECTS	European credit transfer system
EDI	equity, diversity, and inclusion
HMM	hidden Markov models
LA	learning analytics
LAD	learning analytics dashboard
LMS I	earning management systems
MHMM	mixture hidden Markov models
MOOC	massive open online course
MSLQ	motivated strategies for learning questionnaire
OM	optimal matching
OMslen	spell-length-sensitive optimal matching
SE	self-efficacy
SEB	self-efficacy beliefs
SELF	self-efficacy for learning form
SRL	self-regulated learning
TNA	transition network analysis

LIST OF INCLUDED PUBLICATIONS

This thesis is based on data presented in the following articles, referred to by the Roman Numerals I-III.

- I Heikkinen, S., Saqr, M., Malmberg, J., & Tedre, M. (2023). Supporting self-regulated learning with learning analytics interventions – a systematic literature review. *Education and Information Technologies* **28**, 3059–3088. <https://doi.org/10.1007/s10639-022-11281-4>
- II Heikkinen, S., Cristea, T., Saqr, M., Malmberg, J., Kleingeld, A., Snijders, C., Matzat, U., & Tedre, M. (2024). Sequence analysis and process mining perspectives to goal setting: What distinguishes business students with high and low self-efficacy beliefs? *Smart Learning Environments*. **11**, 40. <https://doi.org/10.1186/s40561-024-00327-4>
- III Heikkinen, S., Saqr, M., Malmberg, J., & Tedre, M. (2025). A Longitudinal Study of Interplay Between Student Engagement and Self-Regulation. *International Journal of Educational Technology in Higher Education*. **22**, 21. <https://doi.org/10.1186/s41239-025-00523-3>

The above publications have been included at the end of this thesis with their copyright holders' permission.

LIST OF PUBLICATIONS NOT INCLUDED

Cristea, T., Heikkinen, S., Snijders, C., Saqr, M., Matzat, U., Conijn, R. & Kleingeld, A. (2025). Dynamics of Self-Regulated Learning: The Effectiveness of Students' Strategies Across Course Periods. Computers & Education. <https://doi.org/10.1016/j.compedu.2025.105233>

López-Pernas, S., Heikkinen, S., Tlili, A., Majumdar, R., Saqr, M. (2025). Unpacking learning in the age of AI: Bridging AI, Complexity, and Precision Education. In M. Saqr & S. López-Pernas (Eds.), *Advanced Learning Analytics Methods: AI, Precision and Complexity* (in press). Springer. <https://lamethods.github.io/book2/chapters/ch01-intro/ch01-intro.html>

Heikkinen, S., Saqr, M., Malmberg, J. & Tedre, M. (2024). The interplay of engagement and learning regulation in online learning. Proceedings of TEEM 2024. The Twelfth International Conference on Technological Ecosystems for Enhancing Multiculturality, 23-25 October 2024, Alicante, Spain. Lecture Notes in Educational Technology. Springer, Singapore.

Heikkinen, S., López-Pernas, S., Malmberg, J., Tedre, M. & Saqr, M. (2022). How do business students self-regulate their project management learning? A sequence mining study. Finnish Learning Analytics and Artificial Intelligence in Education Proceedings, September 29–30, 2022, Joensuu, Finland.

AUTHOR'S CONTRIBUTION

- I) The author set the search string for the review together with the co-authors. The author searched databases using search strings. The author then screened the titles and the abstracts and excluded records that did not meet the inclusion criteria. In the following phase, the author read the full texts and excluded records that lacked the three elements of the research question. The author examined the included records and extracted the information from the studies RQ-wise. The author met a co-author repeatedly to agree on a coding scheme. The author then coded papers independently, followed by a meeting to review the initial results. After consensus, the author completed the coding. The author conducted the analysis and wrote the initial manuscript for the publication.
- II) The author contributed to the study conception and design with the co-authors. The author performed material preparation and data collection. The author performed an analysis with the co-authors. The author wrote the first draft of the manuscript together with a co-author. The author finalized the manuscript.
- III) The author contributed to the study conception and design with the co-authors. The author performed material preparation and data collection. The author performed an analysis together with a co-author. The author wrote the first draft and finalized the manuscript.

Table of contents

ABSTRACT	7
Acknowledgements.....	9
1 INTRODUCTION.....	25
1.1 Motivation, Aim, and Research Questions.....	26
1.1.1 Integration of SRL and SE Research	28
1.1.2 Engagement Patterns and Learning Analytics.....	28
1.1.3 Research Gaps in Current Literature	28
1.2 Research Process	31
1.2.1 Study I: Systematic Literature Review.....	31
1.2.2 Study II: Sequence Analysis and Process Mining.....	32
1.2.3 Study III: Multi-channel Analysis	32
1.2.4 Integration and Progression	32
1.3 Structure of The Dissertation	34
2 LITERATURE REVIEW	37
2.1 Literature Review Introduction	37
2.2 Theoretical Framework for Current Research.....	38
2.3 Student Engagement	40
2.4 Self-Regulated Learning	45
2.5 Self-Efficacy Beliefs.....	53
2.6 Learning Analytics	58
2.6.1 Foundations and Key Concepts	58
2.6.2 Analytical Methods and Techniques	61
2.6.3 Supporting SRL and Engagement.....	63
2.6.4 Sequence Analysis	67
2.6.5 Multi-channel Sequence Analysis	75
2.6.6 Process Mining.....	79
3 RESEARCH DESIGN	87
3.1 Identifying LA Interventions Aimed to Support Learners	87
3.2 Relation Between SEB and Temporality of Learning.....	91
3.2.1 Research Methodology.....	92

3.2.2 Data Collection	93
3.2.3 Data Preparation	94
3.2.4 Data Analysis	94
3.3 Longitudinal Interplay Between Student Engagement and SRL	98
3.3.1 Research Methodology	98
3.3.2 Data Collection	99
3.3.3 Data Preparation	101
3.3.4 Data Analysis	101
4 ETHICAL CONSIDERATIONS	103
4.1 Systematic Literature Review	103
4.2 Empirical Studies	104
5 RESULTS	107
5.1 LA Interventions Supporting SRL	107
5.1.1 Methodological Approaches	108
5.1.2 Intervention Channels.....	108
5.1.3 Support for SRL Phases.....	109
5.1.4 Impact and Effectiveness.....	109
5.2 Sequence Analysis and Process Mining Explaining Differences Between High and Low Self-Efficacy Beliefs	110
5.2.1 Learning Tactics	110
5.2.2 Self-Efficacy Beliefs and Learning Behaviors	111
5.3 Interplay Between Student Engagement and SRL	116
5.3.1 Student engagement and Learning Regulation States	117
5.3.2 Multi-Channel Analysis and Clustering	118
5.3.3 Hidden State Analysis.....	120
6 DISCUSSION	123
6.1 Research Contribution.....	127
6.1.1 Theoretical Contributions.....	127
6.1.2 Theoretical Integration Across Studies	129
6.1.3 Methodological Contributions	131
6.2 Recommendations To Practitioners And Technology Developers	133
6.2.1 Practitioners' Perspective	133
6.2.2 Educational Technology Developers' Perspective	136
6.3 Limitations.....	137

6.4 Future Research	139
Appendices	177

List of Figures

Figure 1. Overview of the research process showing the progression and relationships between the three studies 33

Figure 2. Structure of the dissertation showing the organization and flow of chapters..... 36

Figure 3. Relationships between SRL, SEB, and Student Engagement 40

Figure 4. Student engagement framework in the context of LA 41

Figure 5. Cyclical nature of the SRL process 46

Figure 6. Relations between self-efficacy beliefs, learning behavior, and academic achievement..... 55

Figure 7. LA cycle uses collected learning data to support learning processes through interventions 59

Figure 8. Simplified representation of sequences representing typical sequences and alphabet used in LA studies 69

Figure 9. Simplified version of multichannel sequences used to form a combined sequence 77

Figure 10. Simplified version of the process map. Nodes represent typical learning activities and edges represent transitions between nodes. The width of the edge shows the relative commonness of the particular transition 81

Figure 11. Process mining approaches in educational research..... 83

Figure 12. PRISMA flow diagram of the study selection process for the systematic literature review 90

Figure 13. Summary of the results of the [I] 108

Figure 14. Summary of the results of [II]..... 110

Figure 15. Identification of learning tactics through sequence analysis for high efficacy beliefs students. Three complementary visualizations of the learning tactics are the (top left) distribution plot showing the proportion of different tactics over time, (bottom left) sequence index plot displaying individual student sequences color-coded by tactic, and (top right) representative sequences displaying the most common patterns within each

tactic type. The four identified tactics are project-focused, long-diverse, quiz orientation, and short-focused, each representing distinct patterns of student engagement with the learning environment	112
Figure 16. Identification of learning tactics through process analysis for high efficacy beliefs students. Sequential process map showing transitions between learning activities, with node size indicating activity frequency and edge thickness representing transition probability between activities	113
Figure 17. Identification of learning tactics through sequence analysis for low efficacy beliefs students. See the legend in the caption of Figure 15	114
Figure 18. Identification of learning tactics through process analysis for low efficacy beliefs students. See the legend in the caption of Figure 16	114
Figure 19. Identification of learning tactics through sequence analysis for students who did not set learning goals. See the legend in the caption of Figure 15	115
Figure 20. Identification of learning tactics through process analysis for students who did not set learning goals. See the legend in the caption of Figure 16	115
Figure 21. Summary of the results of [III]	117
Figure 22. Temporal network visualization of engagement-regulation state transitions. Nodes represent different engagement and regulation states (color-coded), with node size proportional to their total connectivity in the network. Edges (arrows) represent temporal transitions between states, with edge thickness and numeric labels indicating transition probabilities.	120

List of Tables

Table 1. Overview of the systematic literature review 89

Table 2. Overview of the empirical studies 97

Table 3. Variables, original statements used to derive the variables, and the
source of original statements 100

1 INTRODUCTION

The landscape of higher education has undergone significant transformation in recent years, particularly in how students engage with and regulate their learning in online environments. This transformation has been accelerated by the COVID-19 pandemic and technological advances, leading to increased adoption of online and blended learning approaches (Ali, 2020; Lapitan et al., 2021). In this evolving educational context, three critical aspects of learning have risen to particular prominence: self-efficacy beliefs (SEB), student engagement, and self-regulated learning (SRL). While these constructs have been studied extensively in traditional classroom settings, their manifestation and interplay in online environments present new challenges and opportunities for research (J. Zheng et al., 2020).

SRL is particularly critical in online settings where students must exercise greater autonomy. Recent research has consistently shown that students in online learning environments demonstrate varying levels of self-regulation capabilities (Broadbent et al., 2023), and the interplay between student engagement and SRL in these contexts is not fully understood (Järvelä et al., 2019). As educational institutions continue to expand their online offerings, the need to better understand these dynamics has become increasingly urgent.

The digitalization of education has not only changed how learning occurs but has also created unprecedented opportunities for understanding learning processes through the data traces that students leave while interacting with online learning platforms (Siemens, 2013). Learning analytics (LA), defined as "the measurement, collection, analysis, and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning" (Siemens & Long, 2011 p. 34), has emerged as a powerful approach to interpreting this educational data. However, the challenge lies not just in collecting and analyzing this data but

in using it to develop meaningful insights that can improve learning outcomes (Lodge et al., 2018).

Despite the growing body of research on both student engagement and SRL in online learning, several critical problems remain unaddressed. First, while previous research has examined student engagement and self-regulation as separate constructs, their dynamic interplay throughout a course is not fully understood. Second, most studies have relied on either trace data or self-reported measures but have rarely combined both to provide a more comprehensive view of the learning process (Saqr & López-Pernas, 2024). This single-channel approach may miss important aspects of how students actually learn and regulate their learning in online environments. Third, while many studies have examined the interplay between SEBs and learning outcomes, few have investigated how these beliefs influence the specific learning tactics students employ in online courses.

Understanding these relationships is crucial for developing effective interventions to support student learning. By leveraging LA to examine these relationships, we can better understand how students navigate and regulate learning in digital spaces, ultimately leading to more effective support strategies and improved learning outcomes.

This research aims to address these gaps by examining the complex interplay between student engagement, SRL, and SEBs in online learning environments. Through a combination of systematic literature review, sequence analysis, and multi-channel analysis, this dissertation provides new insights into how students navigate and regulate their learning in digital spaces. The findings contribute to both theoretical understanding and practical applications in the field of online education, offering guidance for designing more effective learning environments and interventions.

1.1 Motivation, Aim, and Research Questions

This shift to online learning environments has highlighted the critical importance of understanding how students regulate their learning in these environments, where greater autonomy and self-direction are required. At

the intersection of this educational evolution lies the complex interplay between SRL, SEBs, and student engagement – three fundamental constructs that significantly influence learning outcomes in educational settings.

This research is motivated by several converging developments in educational technology and learning science. First, the increasing sophistication of learning management systems (LMS) has created rich repositories of data about student learning behaviors (Romero & Ventura, 2020). Second, advances in LA have provided new tools and methods to analyze this data meaningfully (Viberg et al., 2020). Third, there is growing recognition that successful online learning requires both active student engagement and effective self-regulation skills (Järvelä & Hadwin, 2024).

Research has consistently shown that students often struggle with self-regulation in online learning environments (Broadbent & Poon, 2015). The interplay between student engagement and SRL in online environments is not fully understood (Matcha et al., 2019). Despite the growing body of research on both student engagement and SRL in online learning, several critical problems remain unaddressed. First, while previous research has examined student engagement and self-regulation as separate constructs, their dynamic interplay throughout a course is poorly understood (Lima et al., 2023). Second, most existing studies have relied on either trace data or self-reported measures but have rarely combined both to obtain more comprehensive insights into the learning process (Saqr & López-Pernas, 2024). This narrow approach may miss important aspects of how students actually learn and regulate their learning in online environments. Third, while many studies have examined the interplay between SEB and learning outcomes (Artino & Jones, 2012), few have investigated how these beliefs influence the specific learning tactics students employ in online courses. Understanding these relationships is crucial for developing effective interventions to support student learning.

In summary, this thesis focuses on several issues in the current state of research and addresses gaps in the same.

1.1.1 Integration of SRL and SE Research

While the research has established strong connections between SEBs and academic performance (Saqr & López-Pernas, 2024), the mechanisms through which these beliefs influence specific learning behaviors and regulatory processes remain incompletely understood. Studies have shown that students with higher SE tend to set more challenging goals and persist longer in the face of difficulties (Efklides, 2011), but our understanding of how SEBs influence the selection and adaptation of specific learning tactics remains limited. Similarly, how these beliefs shape students' responses to feedback and LA interventions, as well as the temporal dynamics of how SEBs interact with regulatory processes throughout a course, requires further investigation (Järvelä et al., 2019).

1.1.2 Engagement Patterns and Learning Analytics

The rise of online learning environments has created new opportunities for studying engagement patterns through LA (Saqr, López-Pernas, Helske, et al., 2023). Recent research has demonstrated that different patterns of engagement can be identified through trace data analysis (Saqr & López-Pernas, 2024), and these patterns often correlate with academic performance and learning outcomes. Furthermore, engagement patterns can change over time and may be influenced by various factors (Matcha, Gašević, et al., 2020). However, the causal relationships between engagement patterns and self-regulatory processes remain unclear. Additionally, our understanding of how different types of engagement relate to specific phases of self-regulation, and the role of LA interventions in supporting engagement and self-regulation, requires further development (Sedrakyan et al., 2020).

1.1.3 Research Gaps in Current Literature

Despite extensive research on these constructs being conducted individually, several important gaps remain in our understanding of their dynamic interrelationships. From a methodological perspective, most studies rely on cross-sectional designs or focus on single phases of SRL, with

limited use of multi-channel data collection approaches. Additionally, few studies have combed trace data analysis with self-report measures of SE and SRL. Theoretical integration remains incomplete, particularly regarding how SEBs influence specific learning tactics and regulatory behaviors in online environments, and how engagement patterns evolve throughout a course and interact with self-regulatory processes. Better integration of LA insights with SRL theory is also needed.

In terms of practical applications, research on how LA can effectively support SRL development while accounting for individual differences remains limited. There is a clear need for evidence-based guidelines for designing interventions that consider both SEBs and engagement patterns, and how to bridge the gap between theoretical understanding and practical implementation in online learning environments should be clarified. Furthermore, our understanding of how these relationships manifest in specific disciplinary contexts, how different learning environments affect these relationships, and how cultural and institutional factors influence these processes requires further development.

Based on these identified gaps, this research aims to investigate the dynamic relationships between SEBs, engagement patterns, and SRL in online learning environments. It seeks to examine how LA can be effectively used to support the development of SRL skills, and to develop and validate interventions that consider individual differences in SEBs and engagement patterns. This research contributes to both theoretical understanding and practical applications in the field of educational technology and LA. By addressing these gaps, it aims to advance our knowledge of how students regulate their learning in online environments and how we can better support this process through LA interventions.

This dissertation addresses these gaps through three interrelated studies guided by the following main research question:

How do self-efficacy beliefs, student engagement, and self-regulated learning interact in online learning environments, and how can learning analytics help us understand this interplay?

This main question is explored through three studies, each addressing specific aspects:

[I]: How are LA interventions currently being used to support SRL?

- What LA methods have been applied to identify SRL support needs?
- What methods of communication channels have been applied in LA interventions to foster SRL?
- Which phases of the SRL cycle were targeted by LA interventions?
- How do studies evaluate the SRL support efficiency and impact, and what types of results were achieved?

[II]: How do learning tactics differ between students with high and low SEB?

- What kind of learning tactics do project management students use when studying online, could the different learning tactics be used as an indicator towards achieving their goals, and what are the differences in learning processes between different learning tactics?
- How do learning tactics differ between low- and high-SEB students, and what are the differences in learning processes between low- and high-SEB students?

[III]: How do patterns of engagement and self-regulation evolve and interact an online course?

- How do patterns of engagement and self-regulation evolve throughout an online course?
- What distinct clusters of students emerge when considering both engagement and self-regulation trajectories?

This dissertation builds upon three interconnected studies that progressively examine the role of LA in supporting SRL. Paper [I] established the foundation by systematically reviewing how LA interventions have been used to enhance SRL, revealing critical gaps in current approaches – particularly the lack of comprehensive support across all SRL phases. Building on these insights, Paper [II] investigated specific learning tactics and their dynamic interplay with SEBs in business education, using sequence

analysis and process mining to understand how students with different SE levels approach learning. Paper [III] expanded this understanding by examining the dynamic interplay between student engagement and SRL through longitudinal analysis, providing deeper insights into how these processes co-evolve during learning.

1.2 Research Process

This research adopted a systematic and progressive approach to understanding the complex dynamic interplay between SRL, SEBs, and student engagement in online learning environments. The research process, illustrated in Figure 1, comprises three interconnected studies, each building upon the findings of its predecessor to elucidate how students navigate and regulate their learning in digital spaces.

1.2.1 Study I: Systematic Literature Review

The research began with a comprehensive systematic literature review [I] in accordance with the PRISMA 2020 statement. This foundational study aimed to establish the current state of knowledge regarding LA interventions supporting SRL. The review initially identified 753 papers through database searches, ultimately including 56 studies that met the strict inclusion criteria. This study revealed several critical gaps in current research:

- Only 16% of existing studies addressed all three phases of the SRL cycle
- Less than half (46%) of LA interventions demonstrated positive impacts on learning
- Insights into how LA could support the development of self-regulation skills over time were limited

These findings directly informed the design and focus of the subsequent empirical studies, highlighting the need for more comprehensive approaches to studying SRL in online environments.

1.2.2 Study II: Sequence Analysis and Process Mining

Building on the gaps identified in Study I, the second study [II] employed sequence analysis and process mining techniques to examine the dynamic interplay between SEBs and learning tactics in an online project management course. This study analyzed data from 209 business students, focusing on the following:

- Students' learning tactics through sequence analysis of log data
- The dynamic interplay between SEBs and learning behaviors
- Process mining to understand how different learning tactics evolved over time

The study revealed distinct patterns in how students with different levels of SEBs approached learning, providing valuable insights into the dynamic interplay between students' beliefs about their capabilities and their actual learning behaviors.

1.2.3 Study III: Multi-channel Analysis

The final study [III] expanded the analytical approach by implementing multi-channel sequence analysis to investigate the temporal dynamics of student engagement and self-regulation. This study examined data from 165 students, combining the following:

- Trace data from the learning management system
- Longitudinal self-reported measures of SRL
- Hidden Markov models for analyzing state transitions

This innovative approach allowed for a more nuanced understanding of how student engagement and self-regulation co-evolve during an online course, revealing distinct patterns and trajectories of student development.

1.2.4 Integration and Progression

The research process was designed to ensure that each study built upon and complemented the others:

- Study I provided the theoretical and empirical foundation, identifying key gaps in current understanding

- Study II focused on understanding individual learning behaviors and their dynamic interplay with SEBs
- Study III expanded this understanding by examining the temporal dynamics of learning processes

Throughout the research process, particular attention was paid to maintaining methodological rigor while developing increasingly sophisticated approaches to analyzing learning behaviors. The progression from literature review to sequence analysis to multi-channel analysis reflects the growing complexity of understanding how students regulate their learning in online environments.

Figure 1 illustrates this research process, showing the temporal progression and relationships between the three studies. The figure highlights how the findings from each study informed the design and focus of subsequent investigations, creating a coherent research narrative that addresses the main research question from multiple complementary perspectives.

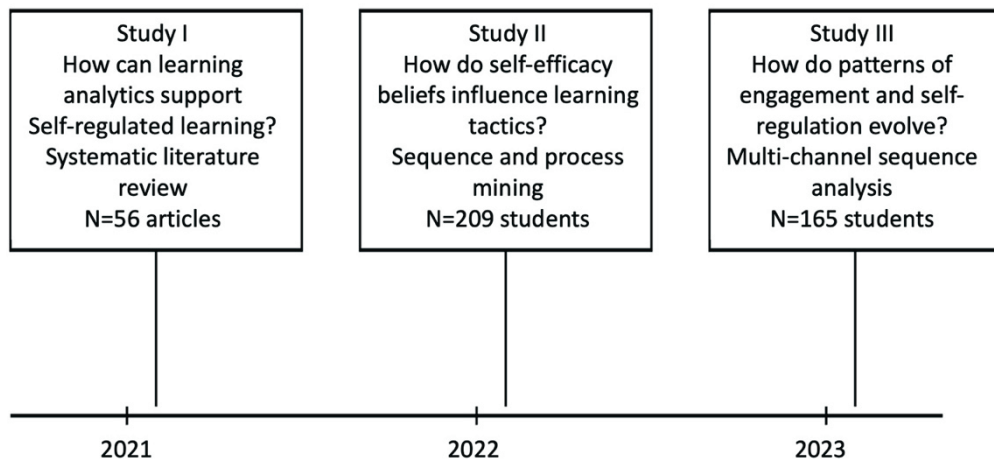


Figure 1. Overview of the research process showing the progression and relationships between the three studies

This systematic approach to the research process enabled a comprehensive examination of SRL and student engagement in online environments, while

ensuring that each study contributed unique insights to the overall understanding of these phenomena.

1.3 Structure of The Dissertation

This dissertation is organized into six main chapters, each building upon the previous to comprehensively examine the dynamic interplay between student engagement and SRL in online environments. Figure 2 provides a visual overview of the dissertation structure.

Chapter 1 (Introduction) establishes the foundation of the research by presenting the motivation, aims, and research questions. It introduces the key concepts of SRL, SEBs, and student engagement in the context of online learning environments. The chapter also outlines the research process and the organization of the dissertation.

Chapter 2 (Literature Review) presents a comprehensive examination of the theoretical foundations and current research in the field. It synthesizes existing knowledge about SRL, SEBs, student engagement, and LA. The chapter critically analyzes previous research to identify gaps in our current understanding and positions this research within the broader academic discourse.

Chapter 3 (Research Design) details the methodological approaches employed in each of the three studies. It explains the rationale behind the chosen methods, describes the data collection procedures, and outlines the analytical techniques used. The chapter emphasizes how each study builds upon the findings of the previous ones to create a coherent research narrative.

Chapter 4 (Ethical Considerations) addresses the important ethical aspects of conducting research in educational settings. It discusses the measures taken to protect participant privacy, ensure informed consent, and handle data responsibly. This chapter also examines the broader ethical implications of using LA in educational contexts.

Chapter 5 (Results) systematically presents the findings from all three studies [I, II, III]. The chapter begins with insights from the systematic literature review [I], followed by the findings on learning tactics and SEBs [II],

and concludes with results on the temporal dynamics of student engagement and self-regulation [III]. Throughout the chapter, the connections between these findings are emphasized to provide comprehensive insights into the research problem.

Chapter 6 (Discussion and Conclusions) synthesizes the findings from all three studies to address the main research question. It discusses the theoretical and practical implications of the research, acknowledges limitations, and suggests directions for future research. The chapter concludes with recommendations for practitioners and researchers in the field of LA and online education.

Each chapter is designed to be self-contained while maintaining clear connections to the overall research objectives. The structure progresses logically from establishing the research context to presenting empirical findings and drawing conclusions that contribute to both theory and practice in the field of LA and online education.

Figure 2 illustrates this structure, showing how each chapter builds upon the previous ones to create a cohesive picture of student engagement and SRL in online environments.

Dissertation Structure

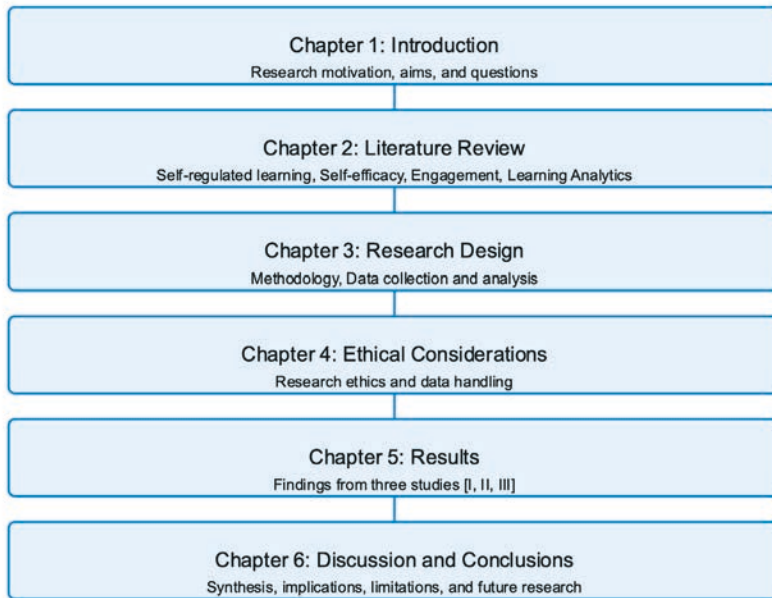


Figure 2. Structure of the dissertation showing the organization and flow of chapters

2 LITERATURE REVIEW

2.1 Literature Review Introduction

This chapter examines the theoretical foundations and empirical research relevant to understanding the dynamic interplay between SRL, SEB, student engagement, and LA in online educational environments. The review begins by exploring SRL as a fundamental framework for understanding how students manage their learning processes. It then examines SEB as a critical component of SRL that influences student behavior and performance. The chapter then examines LA as a methodological approach to understanding and supporting learning processes, including specific techniques, such as sequence analysis and multi-channel analysis. Throughout the review, particular attention is paid to how these concepts interact in online learning environments and their implications for understanding student learning processes.

This research integrates three major theoretical perspectives to understand how students learn in online environments: student engagement, SRL, and SEBs theories. These frameworks, when combined with LA approaches, provide a comprehensive lens for examining the complex dynamics of online learning.

The engagement construct (Martin & Borup, 2022; Redmond et al., 2018) complements these frameworks by providing a multidimensional perspective on how students interact with online learning environments. This theory identifies three crucial dimensions of engagement: a) behavioral engagement (observable participation and effort), b) emotional engagement (affective responses and sense of belonging), and c) cognitive engagement (investment in learning and self-regulation).

Within this broader social cognitive framework, SRL theory (P. R. Pintrich, 2000; B. J. Zimmerman, 2000) offers a more specific model for understanding how students actively manage their learning processes. SRL theory conceptualizes learning as a cyclical process involving three key

phases: forethought (planning), performance (execution), and self-reflection (evaluation). This cyclical nature is particularly important in online environments where students must often independently manage these phases without immediate instructor oversight.

Social cognitive theory (Bandura, 1986) provides the foundational framework for understanding how students' beliefs, behaviors, and environments interact in the learning process. This theory emphasizes the reciprocal dynamic interplay between personal factors (including SEB), behavioral patterns, and environmental influences. In online learning environments, this triadic reciprocal causation becomes particularly relevant as students navigate digital spaces and manage their learning with greater autonomy.

Integrating these theoretical perspectives creates a comprehensive framework for understanding online learning processes. Social cognitive theory explains the broader context of learning and the role of SE. SRL theory provides specific mechanisms for understanding how students regulate their learning. The engagement construct offers concrete dimensions for observing and measuring student participation. LA provides the methodological tools for examining these processes at scale.

2.2 Theoretical Framework for Current Research

The current research adopts Winne and Hadwin's (1998) model of SRL as its primary theoretical foundation, supplemented by insights from engagement research (Henrie et al., 2015) and self-efficacy (SE) theory (Bandura, 1986). This integrated framework views SRL as occurring through four loosely sequenced and recursive phases: 1. Task Definition: where learners develop perceptions about the task and its context, 2. Goal Setting and Planning: where they establish objectives and select strategies, 3 Enactment: where selected strategies are implemented, and 4. Adaptation: where learners evaluate and modify their approaches based on feedback

Within this framework, SEBs are conceptualized as key motivational factors that influence all phases of self-regulation, particularly the interpretation of task requirements, setting of learning goals, selection and

implementation of learning strategies, persistence in the face of challenges, and interpretation of feedback and subsequent adaptations.

Student engagement is viewed as both an outcome of effective self-regulation and a facilitator of regulatory processes. This dual role is particularly evident in online learning environments, where engagement patterns can be tracked through LA and used to understand the effectiveness of self-regulatory processes.

Figure 3 illustrates the theoretical framework guiding this research, depicting the dynamic relationships between SRL, SEBs, and student engagement. The framework is represented as three interconnected constructs, where the overlapping areas signify their mutual influences and interactions. The intersection between SRL and SEBs represents how SE influences goal setting and strategy choice in the learning process, aligning with Winne and Hadwin's (1998) model of SRL. The overlap between SRL and student engagement demonstrates the behavioral and cognitive aspects of engagement that manifest through regulatory processes. The interaction between SEBs and student engagement is characterized by their mutual influence on persistence and motivation, as highlighted in Bandura's (1986) work. The central area where all three constructs converge represents learning outcomes, emphasizing how these interrelated processes collectively contribute to student achievement. This integrated view builds upon existing theoretical models while highlighting the complex interplay between these constructs in online learning environments, providing a foundation for understanding how LA can be used to support student success.

Theoretical Framework

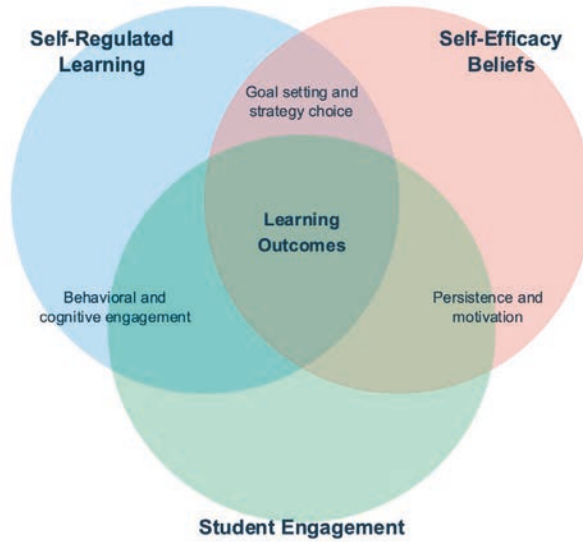


Figure 3. Relationships between SRL, SEB, and Student Engagement

2.3 Student Engagement

Student engagement is a multi-faceted construct that is crucial for understanding learning processes and outcomes in online environments. This section examines the theoretical conceptualization of engagement, including its behavioral, emotional, and cognitive dimensions. It explores how student engagement manifests in online learning environments and its dynamic interplay with SRL and SEB. Particular attention is paid to how engagement can be measured and analyzed in online learning and how this understanding can inform interventions to support student learning. The framework and its relations to the context of LA are illustrated in Figure 4.

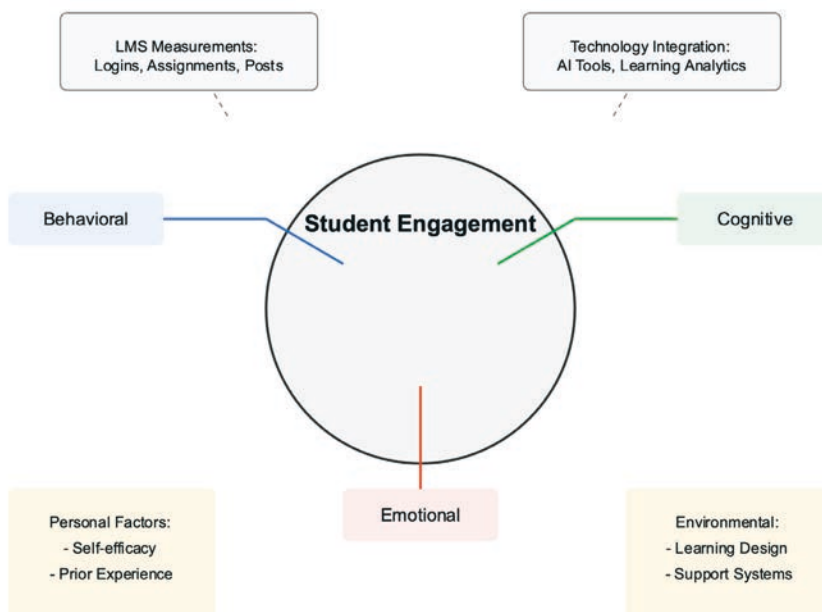


Figure 4. Student engagement framework in the context of LA

Engagement is a three-dimensional construct incorporating behavior, emotion, and cognition (Martin & Borup, 2022; Redmond et al., 2018). Behavioral engagement involves study efforts, such as being present, involvement, and commitment to learning. Emotional engagement explains a student's different emotions toward studying and other learners. Cognitive engagement indicates students' efforts to start and proceed with learning assignments (Borup et al., 2020). In the current study, the behavioral dimension is crucial, as it can be measured using data collected from an LMS. This data includes variables such as completed assignments, frequency of logins, number and frequency of posts, time spent creating posts, time spent online, and the number of accessed learning resources (Fincham, Whitelock-Wainwright, et al., 2019; Henrie et al., 2015). Saqr, López-Pernas, Helske, et al. (2023) have given an example of the methodology that can be used to measure online engagement, by transforming the data to represent states, transitions, and trajectories of different data sources, which can be analyzed by separate or combined states.

Cognitive engagement in online learning has been a focal point of research, particularly in the context of evolving educational technologies and the challenges posed by the COVID-19 pandemic. Cognitive engagement is defined as the psychological commitment students exhibit toward learning tasks, which encompasses deep processing, application of knowledge, and self-regulatory strategies for understanding and problem-solving (Yundayani et al., 2021). The usability of e-learning systems has also been identified as a critical factor influencing cognitive engagement. Research has indicated that students who perceive their online learning platforms as user-friendly exhibit higher levels of cognitive, emotional, and behavioral engagement (Dahleez et al., 2021). This relationship is particularly pronounced in the context of remote learning, which was necessitated by the COVID-19 pandemic, where the quality of the online learning experience directly correlates with student engagement levels.

Previous studies on engagement have shown that prior engagement can predict imminent engagement (Gottfried et al., 2009; Wylie & Hodgen, 2012), but the increase in engagement is not continuous (Wigfield et al., 2007). Numerous variables, including locus of control, self-esteem, parental expectation, communication, and the student's socioeconomic status and grade history (You & Sharkey, 2009) influence the development of student engagement. Furthermore, Wang & Degol (2014) found that emotions, personality characteristics, prior learning experiences, shared values across contexts, and extracurricular activities also affect student engagement. These findings highlight that student engagement is a complex phenomenon that cannot be evaluated in isolation. Therefore, student engagement development should be explored using more advanced approaches. One option is to investigate it using data triangulation along with survey data capturing the essence of SRL in parallel with student engagement measured with LMS-based indicators.

Recent advancements in the field of student engagement have been significantly influenced by the integration of technology, innovative pedagogical strategies, and a deeper understanding of the factors that foster engagement in educational settings. This synthesis aims to explore

these advancements, drawing on a variety of scholarly sources that highlight the multifaceted nature of student engagement and its implications for educational practice.

One of the most notable advancements in student engagement is the increasing recognition of the role that technology plays in enhancing learning experiences. The shift from traditional classroom settings to online and hybrid models has necessitated a re-evaluation of engagement strategies. Research indicates that technology can facilitate deeper levels of engagement by providing diverse content delivery methods that cater to the different needs of various learners. For instance, the use of interactive platforms such as Padlet has been shown to foster greater participation and collaboration among students, thereby enhancing their overall educational experience (Amer et al., 2024). Furthermore, the implementation of LA dashboards has been found to increase student engagement levels by providing real-time feedback and insights into learning behaviors (Ramaswami et al., 2023).

Student engagement is a dynamic, evolving process (Reschly & Christenson, 2022; Salmela-Aro et al., 2016). Longitudinal engagement can be described as the process through which positive perceptions of school environments, a sense of belonging, and academic success motivate students to learn and enhance the behavioral, cognitive, and emotional dimensions of engagement (Tinto, 2022). Previous longitudinal studies have identified distinct patterns of engagement. For instance, Zhen et al. (2020) mapped students' trajectories of academic engagement and identified four types: descending, climbing, persistent, and struggling. Similarly, Li & Lerner (2011) identified four distinct patterns for both behavioral and emotional engagement. In online education, longitudinal studies have been conducted over a full study program to evaluate the development of engagement (Saqr & López-Pernas, 2021), and the interaction between engagement and achievement (Saqr, López-Pernas, Helske, et al., 2023).

The COVID-19 pandemic has also catalyzed advancements in student engagement practices. The abrupt transition to remote learning highlighted the importance of maintaining student engagement in a virtual

environment. Research conducted during this period revealed that active interactions and engagement strategies were crucial for student success in online learning contexts (Diamanti & Nikolaou, 2023; Salas-Pilco et al., 2022). The pandemic underscored the need for institutions to adopt flexible and adaptive approaches to teaching that prioritize student engagement, even in challenging circumstances. This has led to a renewed focus on creating supportive learning environments that encourage student participation and investment in their education.

Another significant advancement in student engagement is the exploration of artificial intelligence (AI) tools, such as ChatGPT, in educational settings. These tools have the potential to revolutionize how students interact with learning materials and engage with their peers and instructors (Glahn, 2024; Halaweh, 2023). Institutions are encouraged to develop guidelines and frameworks that promote the constructive use of AI while addressing potential challenges associated with its integration into educational practices. In LA AI has various potential approaches which can be used as a part of the LA cycle (Saqr & López-Pernas, 2025), including supporting the analysis of multi-modal data while enabling the idiographic analysis of learning (López-Pernas et al., 2025).

The transition into higher education has also been identified as a critical period for fostering student engagement. Research has indicated that successful transition experiences can significantly influence long-term engagement and academic success (Vardal-Ocakli & Ok, 2024). Institutions are increasingly recognizing the importance of providing targeted support and resources during this transitional phase to help students develop the skills and habits necessary for sustained engagement throughout their academic careers.

In conclusion, the landscape of student engagement in higher education is rapidly evolving, driven by technological advancements, innovative pedagogical approaches, and a deeper understanding of the factors that influence engagement. As institutions continue to adapt to the changing educational environment, prioritizing student engagement as a fundamental component of effective teaching and learning becomes

essential. By leveraging technology, fostering collaborative partnerships, and embracing inclusive practices, educators can create dynamic and engaging learning experiences that empower students to take an active role in their education requires SRL skills.

2.4 Self-Regulated Learning

SRL has emerged as a crucial framework for understanding how students actively manage their learning processes, particularly in online environments where learners must exercise greater autonomy (Broadbent & Poon, 2015; Lee et al., 2019; Winne, 2022). This section first examines the theoretical foundations of SRL, including its key components and phases (P. R. Pintrich, 2000; Winne & Hadwin, 1998; B. J. Zimmerman, 2000). It then explores how different SRL models conceptualize the learning process, with focus on their applications in online learning contexts (Hadwin et al., 2011; Lodge et al., 2018). The section concludes by examining current research on SRL interventions and their effectiveness in supporting student learning (Sedrakyan et al., 2020).

SRL is a complex process that encompasses several key components and phases. Broadly defined, SRL involves the setting of goals, monitoring of progress, and regulation of learning (Järvelä et al., 2019; P. R. Pintrich, 2000). Different frameworks are used to explain SRL. A comprehensive review by Panadero (2017) of six prominent SRL models revealed that despite variations in focus and terminology, these models generally agree on a three-phase structure of SRL (Figure 5):

1. **Planning:** This initial phase involves goal setting, task analysis, and strategic planning.
2. **Performance:** During this phase, learners implement their strategies, monitor their progress, and adjust as necessary.
3. **Reflection:** The final phase includes self-evaluation, causal attribution, and adaptation for future learning tasks.

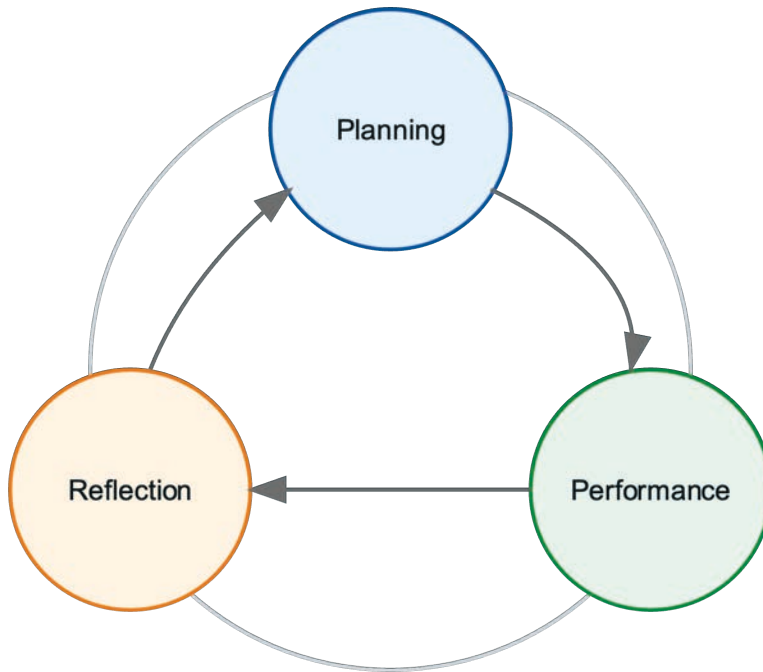


Figure 5. Cyclical nature of the SRL process

While different models emphasize various aspects of SRL, such as motivation (Boekaerts & Corno, 2005), metacognition (Efklides, 2011), or information processing (Winne & Hadwin, 1998), they all acknowledge the cyclical and adaptive nature of SRL. Hadwin et al.'s (2011) model not only considers individual regulation but also how the regulation of learning can be shared between group members. Finally, the models of Pintrich (2000) and Zimmerman (2000) can be described as general models of regulated learning, as they consider motivation and metacognition to be central to SRL.

While SRL models share a common three-phase structure, they differ in their specific subprocesses and areas of focus. In the planning phase, Boekaerts & Corno (2005) included identification, interpretation, primary and secondary appraisal, and goal setting. Efklides (2011) focused on task presentation, whereas Hadwin et al. (2011) emphasized planning. Pintrich

(2000) described this phase as forethought, planning, and activation. Winne & Hadwin (1998) allocated task definition, goal setting, and planning to this phase, and Zimmerman (2000) referred to it as forethought, encompassing task analysis and self-motivation.

The performance phase varies across models as well. Boekaerts & Corno (2005) described it as goal-striving, whereas Efklides (2011) focused on cognitive processing and performance. Hadwin et al. (2011) and Pintrich (2000) highlighted monitoring and control. Winne & Hadwin (1998) emphasized applying tactics and strategies, and Zimmerman (2000) described it as performance with self-control and self-observation.

In the reflection phase, Boekaerts & Corno (2005) focused on performance feedback. Hadwin et al. (2011), Pintrich (2000), and Winne & Hadwin (1998) all included regulating as a key subprocess. Pintrich (2000) also mentioned reaction and reflection, while Winne & Hadwin (1998) emphasized adapting metacognition. Zimmerman (2000) described this phase as self-reflection, including self-judgment and self-reaction. Notably, the model of Efklides (2011) does not explicitly include subprocesses in the reflection phase.

While varying in terminology and focus across models, these subprocesses collectively represent the actions learners take throughout the SRL cycle. They involve estimating how to accomplish a learning task, implementing planned actions, and reflecting on the learning process to inform future learning experiences. Various SRL models share common elements, but they differ in their emphasis and conceptualization of the self-regulation process. Zimmerman's (2000) model, with its cyclical phases of forethought, performance, and self-reflection, provides a clear framework for understanding the temporal aspects of SRL. However, it may oversimplify the complex, often non-linear nature of learning processes. In contrast, Winne and Hadwin's (1998) model, with its emphasis on cognitive and metacognitive processes, offers a more detailed account of the internal mechanisms of SRL but may understate the role of motivation and affect.

Pintrich's (2000) model attempts to bridge this gap by incorporating motivational and contextual factors alongside cognitive processes. While

comprehensive, its complexity can make it challenging to operationalize in empirical studies. Efklides' (2011) model addresses this by focusing on the interactions between metacognition, motivation, and affect, providing a more nuanced understanding of how these factors influence each other during learning. However, it may not fully capture the social aspects of learning that are increasingly important in online environments.

Hadwin et al.'s (2011) model stands out for its consideration of socially shared regulation, which is particularly relevant in collaborative online learning contexts. However, it may overemphasize social aspects at the expense of individual cognitive processes. Each of these models offers valuable insights but also has limitations in fully capturing the complexity of SRL, particularly in diverse online learning environments.

Empirical studies have continuously demonstrated that students do not use efficient self-regulating strategies when learning in virtual environments (Akyildiz & Kaya, 2021; Azevedo et al., 2019; Broadbent & Poon, 2015). Therefore, learners may not achieve the expected learning outcome (Lee et al., 2019; van Alten et al., 2020). This situation has motivated teachers and researchers to design methods to support learners. Recent research on SRL has been focusing on numerous themes. Metacognition, motivation, online learning, higher education, learning strategies, and SE have been the key themes that are present in the majority of the studies during the past few decades (Gambo & Shakir, 2021; Sulistiawati et al., 2023). Analyses of the links between different themes, have revealed that self-regulation is most often connected to LA, SE, motivation, and higher education (Tao et al., 2023).

Many studies have suggested that interventions could enhance SRL (Ceron et al., 2021; Lee et al., 2019; Wong et al., 2019). Interventions can assist at-risk learners in improving their learning performance (Espinoza & Genna, 2021) or can help reduce undesirable learning behavior (Lodge et al., 2018). Furthermore, interventions can be used to increase students' mastery goals (J. Zheng et al., 2020), SE (Samuel & Warner, 2021), and the perceived value of a course (Alamri et al., 2020). This highlights the complex nature of SRL and emphasizes the need for holistic approaches to support online

learners. Although the development of LA interventions has improved learning in the performance phase of the SRL cycle, interventions that would benefit different learning designs still need to be developed (Kleimola et al., 2024). This research contributes to the field of SRL by investigating how SRL is promoted through interventions and exploring how SRL can be measured during online courses. Central to this discussion is the concept of SE.

While various components influence SRL, SEB has emerged as a particularly crucial factor in determining how students approach and persist in their learning tasks. These beliefs about one's capability to organize and execute actions necessary for learning not only influence goal setting and strategy use but also shape how students regulate their learning processes. Understanding the dynamic interplay between SRL and SEB is particularly important in online learning environments, where students must exercise greater autonomy in managing their learning.

The integration of technology has significantly influenced recent advancements in the field of SRL, the understanding of gender differences, and the exploration of various pedagogical strategies. These advancements are crucial in enhancing educational outcomes and fostering independent learning among students across different educational contexts.

One of the notable advancements in SRL is the increasing recognition of the role of technology in facilitating SRL processes. Bembenutty (2023) highlighted that computer-based learning environments have expanded the frontiers of SRL research, allowing for more personalized learning experiences. The integration of LA, as discussed by Ye & Pennisi (2022), provides educators with valuable insights into students' learning behaviors, enabling them to tailor interventions that enhance self-regulation. Furthermore, the use of digital tools has been shown to support students in setting goals, monitoring their progress, and reflecting on their learning experiences, which are essential components of self-regulation (Jaramillo et al., 2022).

Moreover, the impact of perceived teacher support on SRL has emerged as an important area of research. Yin & Luo (2024) demonstrated that students who perceive higher levels of support from their teachers are more

likely to engage in self-regulation strategies, leading to increased learning engagement. This relationship underscores the importance of creating supportive educational environments that empower students to take charge of their learning processes.

The relationship between SRL and academic performance has also been a focal point in recent studies. For instance, Jaramillo et al. (2022) found that SRL skills are predictive of academic success, particularly in virtual learning contexts during the pandemic. This aligns with the findings of Cahyawati et al. (2023), which indicate that self-regulation can mitigate the negative effects of anxiety on academic achievement. Such evidence highlights the critical role of self-regulation in promoting resilience and adaptability among learners, particularly in challenging educational settings.

The incorporation of gamification and interactive learning strategies has also gained traction as a means to enhance SRL. Pausch (2023) discussed how gamification can motivate learners by leveraging their intrinsic motivation, thereby encouraging them to engage more deeply with the learning material. This approach aligns with the findings of Abdul Talib et al. (2023), who identified various SRL strategies, such as goal setting and self-monitoring, that can be effectively integrated into gamified learning environments.

In the context of online learning, the COVID-19 pandemic has accelerated the need for effective SRL strategies. Dai et al. (2022) explored how SRL profiles can influence engagement in online mathematics learning, revealing that students with higher self-regulation are more likely to succeed in virtual environments. This is particularly relevant as educational institutions continue to adapt to hybrid and online learning models.

One of the foundational theories underpinning SRL is the self-determination theory, which posits that fulfilling psychological needs for autonomy and competence is crucial for promoting effective learning behaviors. Zheng (2023) emphasized the importance of SE and informal digital learning as mediators in the relationship between SRL and course satisfaction, suggesting that learners who feel competent and autonomous are more likely to engage in self-regulated behaviors in online environments

(Y. Zheng & Xiao, 2023). This aligns with Schunk and Zimmerman's comprehensive overview of self-regulation, which underscores the significance of motivation and SE in enhancing learners' performance and persistence in online learning settings (Schunk & Zimmerman, 2023).

The readiness of learners to engage in e-learning also plays a critical role in their self-regulation capabilities. Ucar & Ugurhan's (2023) investigation into e-learning readiness indicated that learners who spend more time on digital platforms exhibit higher levels of SRL skills, suggesting that familiarity and comfort with online environments can enhance learners' ability to manage their learning processes effectively (Ucar & Ugurhan, 2023). This is further supported by Wu et al., who found that interaction in online learning environments positively influences satisfaction through the mediating roles of academic emotions and SRL (Wu et al., 2023). The interplay between learner readiness and engagement highlights the importance of creating supportive online environments that foster self-regulation.

The role of artificial intelligence in supporting SRL is another emerging area of interest. Jin et al. discuss how AI applications can assist students in managing their learning by providing personalized feedback and resources, thereby enhancing their self-regulatory skills (S.-H. Jin et al., 2023). This technological integration represents a significant advancement in the field, as it offers scalable solutions to support diverse learners in online environments. Furthermore, Prasetya & Raya's (2023) research highlighted the interplay between SRL behavioral factors and students' performance in online language learning, emphasizing the importance of tailored strategies to enhance self-regulation in specific contexts (Prasetya & Raya, 2023).

The impact of learner support on SRL was further illustrated by Rui & Liu (2023), who found that perceived teacher support positively affects learners' self-regulation, mediated by SEBs. This highlights the importance of instructor involvement and support in online learning environments, as practical guidance can empower learners to take control of their learning processes. Additionally, the findings of Zhou et al. (2022) reinforce this notion by demonstrating that teacher support is crucial for enhancing learners' SRL capabilities in online settings.

The transition from traditional to online learning has necessitated reevaluating students' learning strategies. Park's study comparing learning strategy use between online and offline classes found no significant differences, suggesting that effective learning strategies can be adapted across different modalities (E. H. Park, 2024). This adaptability is essential for promoting SRL, as learners must be equipped with a versatile toolkit of strategies to navigate various learning environments successfully.

Moreover, the COVID-19 pandemic has catalyzed research into SRL, as many educational institutions were forced to shift to online formats. Guo et al. (2022) conducted a multi-institutional survey during this period, revealing that self-regulation is a core concept for understanding the metacognitive, motivational, and emotional aspects of learning in online contexts. This research provides valuable insights into the challenges faced by different learner subgroups, emphasizing the need for targeted interventions to support self-regulation during times of transition.

Recent work by T. S. Cristea et al. (2025) provided important insights into how students' learning strategies evolve throughout a course. Their study revealed that students modify their learning behaviors as courses progress, sometimes adopting less effective strategies over time. Their finding that students increasingly relied on formative assessment strategies, though these did not contribute significantly to academic success, was particularly noteworthy. Their research emphasizes the importance of metacognitive engagement and the ability to balance surface and deep learning strategies, with students who effectively transition between different strategies showing better academic performance.

Recent research has explored how emerging technologies affect SRL and engagement. The dynamic interplay between self-regulated learning and engagement has been influenced by generative AI, leading to improved self-regulated strategies and engagement (M. Pan et al., 2025). However, while AI can help meta-cognitive processes, studies indicate it does not necessarily improve self-efficacy (Deng et al., 2025).

In conclusion, the advancements in SRL within online education are characterized by a growing body of research that highlights the interplay

between engagement, SE, and technological integration. As educational institutions continue to navigate the complexities of online learning, it is imperative to prioritize the development of SRL strategies that empower learners to take control of their educational journeys.

SRL serves as a foundational theoretical framework throughout this research, offering a comprehensive lens for understanding how students regulate their learning processes in online environments. The research specifically employs the three-phase model of SRL (planning, performance, and reflection) to examine both cognitive and behavioral aspects of student learning. This theoretical foundation is strategically integrated with LA to enable deeper insights into student learning behaviors and to provide structured ways of evaluating student engagement and learning strategies. The innovative aspect of this research lies in its systematic integration of SRL theory with LA, enabling more nuanced insights into student learning processes, informing the design of effective interventions, supporting personalized learning approaches, and elucidating the relationship between self-regulation and academic success in online learning environments.

2.5 Self-Efficacy Beliefs

SEB represent a critical psychological construct that influences how students approach, persist with, and ultimately perform their learning tasks. This section examines the theoretical foundations of SEB and their relationship to SRL. It explores how these beliefs develop, their impact on student behavior and performance, and their particular relevance in online learning environments. Special attention is given to the relationship between SEB and online learning, as understanding this connection is crucial for developing effective interventions to support student learning. Recent studies have revealed a relationship between SE and SRL. Students with high self-regulation have better SE than students with low self-regulation (Arianto & Hanif, 2024).

SE is a core concept in social cognitive theory, which refers to individuals' beliefs about their capability to organize and execute actions necessary to achieve specific performance outcomes (Bandura, 1986). SEB are the

specific thoughts and judgments people hold about their ability to perform particular tasks or achieve certain outcomes (Pajares, 1996). The main difference between the two is that SE is the broader construct, whereas SEB are the specific cognitive manifestation of this construct. In academic settings, SEB can be interpreted as students' confidence in their ability to complete academic tasks or reach specific educational objectives successfully (Schunk, 1991).

SEB are crucial in students' motivation, academic performance, and SRL (Pajares, 1996). These beliefs influence various aspects of learning. Students with high SE are more likely to embrace challenges, whereas those with low SE tend to avoid difficult tasks or give up easily (Demirören et al., 2016). Students with higher SE tend to use more effective learning strategies and are better at planning and monitoring their learning (Fadillah et al., 2021). In online learning environments, SE has been identified as an enabling factor, with higher SE being associated with improved problem-solving abilities (Hwang & Oh, 2021). Students with higher SE use more suitable methods at the right time than students with lower SE (Liu et al., 2023). SE is linked to diverse elements of learning, such as interacting with classmates or instructors (Shen et al., 2013). Finally, high efficacy helps students to complete their academic tasks (C. L. Tsai et al., 2020).

Recent studies have focused on exploring the relationship between SE and SRL. Shin (2024) demonstrated that SE can forecast self-regulation and course outcomes, suggesting that students' beliefs in their abilities play a pivotal role in their capacity to manage their learning effectively. Moreover, the practical application of metacognitive strategies has been linked to increased SE, pointing to a potential bidirectional relationship between SRL skills and SEB (Khosravi et al., 2023). Evidence that SE intermediates the dynamic interplay between SRL strategies and student engagement also exists (Goslen et al., 2024; Xu et al., 2024). These results align with the previous findings about SE as a factor in improving student engagement (X. Pan, 2022). Furthermore, students who use more SRL strategies report higher SE (Bai & Guo, 2018), and those who have high SRL skills show the highest SE (J. Chen et al., 2022). Figure 6 simplifies this complex relationship.

Self-Efficacy Beliefs in Learning

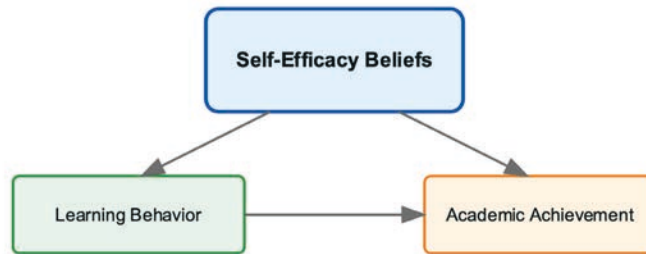


Figure 6. Relations between self-efficacy beliefs, learning behavior, and academic achievement

As education increasingly moves online, researchers have begun to explore how these concepts apply in digital learning environments. Hsu et al. (2023) examined how students allocate their time during exam preparation, revealing that consistency in study time throughout the preparation period led to improved performance. This underscores the importance of time management skills, an element of SRL, within the broader context of SRL. Another perspective is on goal-setting, which is also connected to SE (Su et al., 2019). Implementation of adaptive learning strategies incorporating diverse didactic techniques has been shown to enhance learning outcomes and promote both disciplinary and transversal competencies (Rincon-Flores et al., 2024).

SEB are crucial in online learning environments, where students often need to be more autonomous and self-directed. Artino & Jones (2012) found that SE was positively related to task value and elaboration strategies in online courses, suggesting its importance in motivating students to engage deeply with course material. Moreover, Kuo et al. (2014) demonstrated that internet SE was a significant predictor of student satisfaction in online learning, highlighting the technology-specific aspects of SE in digital environments.

The relationship between SE and actual performance in online settings is complex. While generally positive, some studies have found discrepancies. For example, Broadbent (2016) observed that while high SE was associated with better academic performance in online courses, some students overestimated their abilities, leading to reduced effort and potentially lower achievement. This highlights the importance of accurate self-assessment in online learning contexts.

However, measuring SE in online contexts presents unique challenges. Traditional self-report measures may not capture the nuanced ways in which SE operates in digital environments. Some researchers, like Shen et al. (2013), developed online learning-specific SE scales, which include SE for completing online courses and interacting with instructors and peers online.

With interventions, SE can be improved (Teng & Zhang, 2020). The interventions can have different consequences on SE. Students with higher SE perceive dashboards as helpful for study planning, whereas students with lower SE estimate dashboards as helpful for study monitoring (Silvola et al., 2023). However, interventions that provide feedback in the form of comparative information can be risky for students with low SEB when interpreting this type of feedback (Kaveri et al., 2023) and comparative information can be harmful, especially to SEB.

These findings suggest that understanding SEB and how they relate to other factors is essential in online education. By understanding online learning at a granular level, we can intervene early during online courses and help learners apply more SRL strategies, improve their time management, and set goals, which help them achieve better outcomes.

Goal-setting is closely intertwined with both SRL and SE. An effective learning goal should be specific, measurable, challenging, and attainable (Schunk, 2011). For instance, instead of a vague goal, such as “do well in the course,” a student may set a specific goal, such as “complete all weekly assignments on time and achieve a grade of 80% or higher on the final project.” This goal is specific, measurable, challenging yet attainable, and provides clear criteria for success. Goal-setting is particularly beneficial when implemented at the beginning of a course, as it helps students in the

planning phase of the SRL cycle. Two critical factors that influence students' ability to achieve their goals are SEB and expectancy value.

Complementing the concept of SE, the expectancy-value theory provides another framework for understanding student motivation and achievement. The expectancy-value theory provides a framework for understanding individuals' attitudes and behaviors based on their beliefs and evaluations of potential outcomes. This theory posits that people's attitudes are shaped by two main factors: expectancy and value (Wigfield & Eccles, 2000). Expectancy refers to a student's subjective belief about the probability of a particular outcome occurring as a result of their learning activities. Value pertains to the subjective importance or desirability that a student assigns to a specific outcome or goal. When students believe in their ability to succeed and perceive value in the learning task, they are more likely to persist in their efforts, leading to improved learning outcomes. To illustrate, consider two students in a challenging mathematics course. Student A believes they can master the material with effort (high expectancy) and sees the course as crucial for their future career (high value). Student B doubts their ability to succeed (low expectancy) and doesn't see the relevance of the course to their goals (low value). According to the expectancy-value theory, Student A is more likely to engage deeply with the course material and persist in the face of difficulties. Given the complex interplay of these factors, researchers have sought innovative methods to understand and support student learning. One such promising approach is sequence analysis.

The interplay between SRL, SEB, and goal-setting is complex but crucial for understanding student performance in online learning environments. As online learning continues to grow in importance, further research in these areas will be vital for developing strategies to support all learners, particularly those with low SEB. When learning takes place in online environments, the methods especially designed for these environments, such as LA, can be helpful for understanding learning processes in more depth. This research contributes to this field by applying by investigating the development of SRL during online courses.

While SEB provide insight into students' confidence in their abilities, student engagement represents the actual manifestation of these beliefs in learning behaviors and interactions. The relationship between SEB and student engagement is particularly significant in online learning environments, where maintaining active participation and meaningful interaction can be challenging. Understanding how students' beliefs about their capabilities influence their engagement patterns is crucial for developing effective support strategies.

2.6 Learning Analytics

LA has emerged as a powerful approach for understanding and optimizing learning processes through collecting, analyzing, and interpreting educational data (Siemens, 2013; Viberg et al., 2018). This section begins by examining LA's theoretical foundations and key concepts, including its definition, purposes, and methodological approaches (Romero & Ventura, 2020). It then explores specific analytical techniques relevant to this research, including sequence analysis and multi-channel analysis (López-Pernas & Saqr, 2021; Saqr, López-Pernas, Helske, et al., 2024). The section concludes by examining how LA can be applied to understand and support SRL and student engagement in online environments (Lodge et al., 2018; Matcha, Gašević, et al., 2020).

2.6.1 Foundations and Key Concepts

LA is defined as "the measurement, collection, analysis, and reporting of data about learners and their contexts for the purposes of understanding and optimizing learning and the environments in which it occurs" (Siemens, 2013, p. 1382). LA can be used to analyze and report learner performance and to highlight the parts of learning where SRL processes could be improved (Lodge et al., 2019; Winne, 2022). LA for SRL has two elements, calculation and recommendation (Winne, 2017). These two elements are built within the LA model's four stages: data generation, tracking, analysis, and action (Romero & Ventura, 2020). First, data is generated when learners perform activities in online learning environments. This trace data serves as

input for calculating, analyzing, and tracking learners' processes. Next is the action stage, where instructors execute interventions to support learning activities based on the analytics (Ifenthaler & Yau, 2020). LA interventions aim to support the changes that learners can make to achieve better outcomes. This cycle is presented in Figure 7.

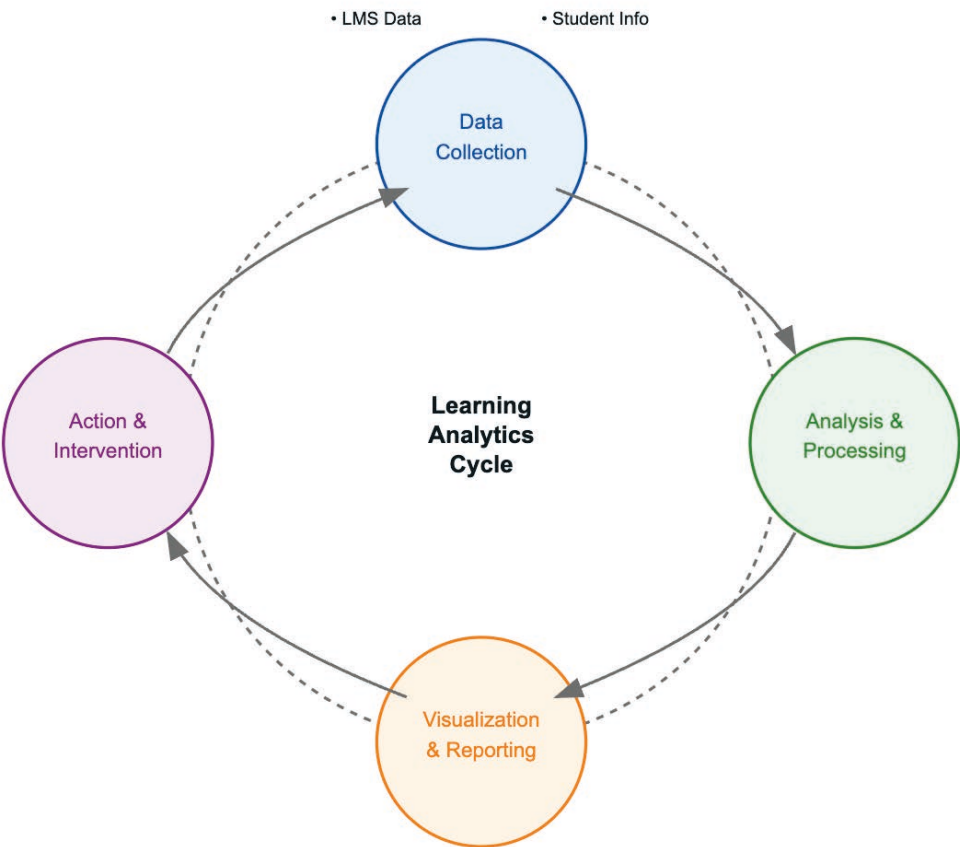


Figure 7. LA cycle uses collected learning data to support learning processes through interventions

While early LA methods relied on fixed factors and had limited predictive capabilities, contemporary dynamic models can be applied to more complex learning contexts. However, implementing effective interventions remains challenging and requires careful design to successfully support learning (Rienties et al., 2017; Sedrakyan et al., 2020; Wise, 2014). A typically used channel for LA interventions is the learning analytics dashboard (LAD; Bodily

et al., 2018). Matcha, Uzir, et al. (2020) focused on LAD interventions to support SRL. Their systematic review highlighted the potential of LADs in supporting various aspects of SRL but also noted challenges in implementation and effectiveness across different contexts. According to Larrabee Sønderlund et al. (2019), most interventions aim to identify at-risk students, namely to students who are highly likely not to pass the course or to pass with poor results (Shafiq et al., 2022).

Viberg et al. (2020) revealed that most studies have concentrated on SRL's forethought and performance phases, with less attention given to the reflection phase. Importantly, they found limited evidence of LA being used to directly support SRL or improve learning outcomes. Instead, LA research focused on measuring SRL, suggesting a need for more interventional studies in this field. Pérez-Álvarez et al. (2018) explored tools supporting SRL in online environments. They found that visualization was the most common means of supporting SRL, and the most used indicators were either action-related or content-related. This is not the only study that has tried to investigate the relationship between LA and SRL. Several studies have investigated SRL with LA, and numerous reviews have summarized the findings of these studies. Still, new measuring methods (Prasse et al., 2024) that could help to elucidate online learning and promote the SEB and SRL of learners are required.

Effective interventions require a broad selection of stakeholders to make the interventions successful. Bayer et al. (2024) emphasized the importance of stakeholder involvement in the design and development of LA tools. By engaging with a wide range of educational professionals, including lecturers, EDI (equity, diversity, and inclusion) leads, and student support teams, they ensured that the dashboard met the specific needs of its intended users. This approach aligns with the growing recognition in the field that the effectiveness of LADs is closely tied to their alignment with user needs and the context in which they are deployed (Ahn et al., 2019; Pozdniakov et al., 2022).

Recent research has demonstrated the potential of LAD to support SRL and feedback, particularly for procedural skills training, though systematic

reviews indicate mixed results in their overall effectiveness (Kaliisa et al., 2023). Villagran et al. (2024) conducted a design-based research study to develop and evaluated an LAD. Their two-phase approach, which included cognitive walkthroughs, surveys, and interviews, revealed that students highly valued features such as performance visualization, peer comparison, and planning resources. Importantly, the study found that higher levels of self-reported SRL strategies predicted higher levels of sensemaking of LAD indicators, suggesting a positive relationship between students' self-regulatory capabilities and their ability to leverage LAD data. The authors also noted that while students found the LAD useful for supporting feedback uptake and strategic planning, there were challenges in motivating the use of self-evaluation features. This aligns with previous findings on the complexities of promoting student engagement with feedback (Carless & Boud, 2018). Villagran et al. (2024) highlighted the potential of LADs to enhance feedback processes in higher education while underscoring the need for careful design considerations to fully support all aspects of SRL. These design considerations are particularly crucial given that recent systematic reviews have found that LADs often have a limited impact on students' motivation and attitudes toward learning (Kaliisa et al., 2023).

Another often neglected viewpoint on interventions is feedback literacy, with a need for the same being identified (F. Jin et al., 2024). S.-H. Jin et al. (2023) found that students perceived feedback tools as potentially helpful in enhancing their feedback literacy skills. Specifically, such tools can support students in processing feedback, eliciting feedback, acting on feedback, appreciating and committing to feedback, and managing emotions related to feedback. These findings suggest that carefully designed feedback analytics tools can scaffold various aspects of student feedback literacy within peer assessment activities.

2.6.2 Analytical Methods and Techniques

LA employs a rich collection of data and methods to analyze online learning processes. Data sources can include clicks, assessments, quizzes, and log traces from learning management systems (LMS). Analysis approaches

include measuring and predicting student performance (Elmoazen et al., 2023; Namoun & Alshantqi, 2020), predicting dropout (de Oliveira et al., 2021), evaluating networking elements (Malmberg et al., 2022; Saqr & Alamro, 2019), and identifying learning patterns (Juhaňák et al., 2019).

While LA has shown promise in providing insights into learning processes, there is a growing recognition of the need to implement these approaches in more ecological settings to increase their generalizability and impact (Cukurova et al., 2020). Modern LA approaches incorporate data from various sources such as speech, text, eye movements, gestures, and facial expressions (Blikstein & Worsley, 2016). Triangulated data from multiple sources offers a more comprehensive understanding of learning processes by integrating and examining data from multiple dimensions. However, implementing multi-source data collection and analysis in real-world settings presents several challenges, including data privacy concerns, and ethical considerations, as well as ensuring validity, setting up complex data collection pipelines, performing data fusion, and conducting analysis (Schneider et al., 2024).

Recent work has focused on addressing these challenges and moving toward more ecological implementations. For example, Schneider et al. (2024) described a project instrumenting a maker space with cameras to capture pose and gaze data, highlighting the importance of working with stakeholders to address privacy concerns and create useful platforms. Martinez-Maldonado et al. (2024) discussed the development of an analytics platform designed to provide actionable feedback on team interactions in healthcare simulations, emphasizing the value of human-centred design approaches. Bernacki et al. (2025) investigated the use of multiple data sources to validate digital traces of SRL and predict student performance in undergraduate courses. This research addresses a key challenge in SRL research: inferring SRL processes from digital events and validating these inferences.

These ecological implementations of multi-source learning analytics are providing new insights into learning processes and opening up possibilities for data-driven interventions. However, they also underscore the need to

carefully consider ethical concerns, stakeholder involvement, and the development of flexible and scalable technical infrastructures. As the field progresses, there is a growing emphasis on creating LA tools that not only inform but also empower users, ensuring that insights can be translated into practical strategies for enhancing teaching effectiveness and student learning outcomes (Worsley et al., 2021).

2.6.3 Supporting SRL and Engagement

Recent advancements have expanded our understanding of student engagement and learning processes in online environments. For example, Nakamura et al. (2024) employed change point detection algorithms to analyze classroom engagement, demonstrating the potential for identifying significant shifts in student behavior and enabling more timely interventions. However, Villalobos, Perez-Sanagustin, et al. (2024) caution that while LA interventions can improve online student engagement, they do not necessarily lead to improved learning outcomes. This highlights the need for a more nuanced approach focused on enhancing learning effectiveness rather than simply increasing student engagement.

Recent research has emphasized that engagement cannot be measured through LMS clicks alone (Fincham, Whitelock-Wainwright, et al., 2019; Johar et al., 2023). There is a growing recognition of the need to develop methods that capture all dimensions of engagement (Johar et al., 2023). Bergdahl et al. (2024) found that while log data is used to capture engagement in 21.4% of studies, it is often accompanied by survey data (33.3% of studies) to elucidate student engagement. This multi-modal approach helps overcome the limitations of uni-dimensional log data analysis (Tempelaar et al., 2020).

Self-reported SRL and student engagement were analyzed together (W. Wang et al., 2024). Self-reporting took place at the end of the research experiment. Our approach improves on this setting by applying a longitudinal approach to assess student engagement with LA data accompanied by self-reported SRL data. The granularity is improved with temporal data that examines the development of student engagement and SRL multiple times during the course.

In previous studies, the longitudinal approach has been used to analyze student engagement in relation to academic achievement (Saqr, López-Pernas, Helske, et al., 2023). However, the longitudinal relationship between student engagement and SRL remains unstudied. Thus, this study examines how the dynamic interplay between student engagement and self-regulation evolves during a single course. We combine the self-reported survey data with the trace data captured from LMS to show a more holistic view of the learning process. In this context, Tepgec et al. (2024) argued for the importance of developing students' feedback literacy skills to translate increased student engagement into improved learning outcomes.

LA uses process mining, sequence analysis, and temporal network analysis to analyze the temporal aspect of learning. These methods are increasingly being utilized to enhance the understanding of learning processes and improve educational outcomes. These methodologies provide a framework for analyzing complex data generated during learning activities, enabling researchers and educators to derive actionable insights from learner interactions and behaviors. Process mining is a technique that focuses on extracting knowledge from event logs generated by information systems. In the context of LA, it allows researchers to analyze the sequences of actions taken by learners, thereby uncovering patterns and inefficiencies in learning processes. For instance, Koschmider et al. (2022) highlighted that process mining can facilitate the automatic analysis of educational processes by discovering process models and conducting conformance checks between expected and actual learning behaviors.

Moreover, the integration of process mining with LA can greatly elucidate how learners engage with educational content. Monroy et al. (2015) demonstrated the application of process mining methods to characterize inquiry-based pedagogical processes, illustrating how these techniques can reveal insights into curriculum use and learner engagement. By systematically analyzing event data, educators can optimize learning environments and tailor instructional strategies to better meet the needs of students (Polat & Eroğlu, 2021).

Sequence analysis complements process mining by focusing on the order and timing of events in learning activities. This method allows researchers to examine how learners progress through educational tasks over time, providing insights into their learning trajectories. Saqr & Nouri (2020) emphasized the importance of understanding the temporal dimensions of collaborative learning through sequence analysis, which can reveal how interactions among learners evolve and impact learning outcomes. This temporal perspective is vital for identifying critical moments in the learning process that may influence student success. E. Park et al. (2023) discussed how sequence analysis can uncover learners' SRL strategies by examining interaction patterns within adaptive learning environments. Similarly, B. Chen & Poquet (2022) noted that sequence analysis provides a robust framework for characterizing learning and learner behaviors, which can be pivotal in tailoring educational interventions.

Temporal network analysis further elucidates learning dynamics by examining the relationships and interactions among learners over time. Saqr & Peeters (2022) illustrated how temporal networks can be employed to analyze collaborative learning environments, revealing the centrality and influence of different learners within a network. This approach not only uncovers the relational aspects of learning but also highlights how the timing of interactions can affect collaborative outcomes. Ifenthaler et al. (2019) emphasized the potential of network graph analysis to inform learning design by analyzing navigation patterns and interactions among learners. Sometimes, the navigation process can be complex.

Complexity in LA refers to the instability or unpredictability of state arrangements in sequential data (Ritschard, 2023). It encompasses multiple aspects, including transitions between states, the number of visited states, and the unpredictability of time spent in each state. This concept has been applied in educational research to understand learning behaviors, with studies showing that the complexity of students' behavior can account for significant variations in academic performance (Tan & Samavedham, 2022). The analysis of complexity in learning requires considering both variable- and person-centered approaches to capture the full complexity of learning

environments (Tempelaar et al., 2023). Person-centered approaches are particularly valuable as they enable the identification of subpopulations within the data, acknowledging the diversity and complexity of learner behaviors and interactions (Howard & Hoffman, 2018). This complementary approach helps researchers better understand the dynamic nature of learning processes and provides a more comprehensive view of how students engage with learning materials over time. Villalobos (2024) applied the person-centered approach to capture the complexity of learning at the subgroup level.

Learning, as a complex dynamic system (Saqr, Dever, et al., 2024), exhibits dynamic behaviors that emerge from the interactions between system components. These dynamics refer to the changes exhibited by the system over time due to the evolution of component interactions and the influence of external conditions in the environment (Favela & Amon, 2023). The dynamics of a system can demonstrate different patterns: cycles of repeated behaviors, chaotic behaviors that appear random, or emergent patterns that suddenly or gradually arise from local interactions. However, these dynamics are highly sensitive to initial conditions, where small variations in the starting point can have multiplicative effects on system behaviors later on – a phenomenon known as the "butterfly effect" (Lorenz, 2000). This sensitivity to initial conditions makes it nearly impossible to employ predictive modeling for complex systems in the long term, though short-term behavior may be relatively predictable. A key concept in understanding system dynamics is that of attractor states, which serve as stable configurations of relationships and behaviors that a complex system tends to evolve toward (Han & Amon, 2021). These attractors can reflect positive stability, where the system maintains a desirable equilibrium or negative stability, where the system becomes "stuck" in potentially maladaptive patterns (López-Pernas & Saqr, 2024).

Saqr et al. (2025) proposed Transition Network Analysis (TNA) as a novel methodological approach that captures both temporal and relational aspects of learning processes by combining process mining and network analysis. This helps to tackle both the complexity and dynamics perspective

of the LA data. TNA models transition probabilities between learning events as a weighted directed network, allowing researchers to analyze the structure, timing, and relationships holistically (Saqr, López-Pernas, Törmänen, et al., 2024). The method extends beyond traditional process mining by incorporating network measures at the node, edge, and graph levels, enabling the identification of important events (through centrality measures), critical transitions (via edge centralities), and process connectivity (through graph density) (J. Helske et al., 2024). TNA also offers unique analytical capabilities including community detection, pattern mining, and statistical validation through bootstrapping and permutation testing. This comprehensive approach allows researchers to identify significant patterns in learning behaviors, validate their findings, and draw meaningful insights about the temporal dynamics of learning processes, making it particularly valuable for understanding complex educational data (Newman, 2018; Saqr, López-Pernas, Törmänen, et al., 2024).

In conclusion, the application of process mining, sequence analysis, and temporal network analysis in LA research offers a robust methodology for analyzing educational data. These techniques enable researchers to uncover patterns in learner behavior, optimize educational processes, and enhance the overall learning experience. In the current thesis, sequence analysis and process mining are used to gain a deeper understanding of students' learning processes. This helps to elucidate, at a granular level, what happens during online learning. Furthermore, understanding the differences in learning processes helps to identify the students who need support during learning to achieve their learning goals.

2.6.4 Sequence Analysis

One promising approach in LA is sequence analysis, which can identify learning tactics students use while studying online (Hadwin et al., 2007). This method constructs sequences of student actions from trace data, and is used, for example, to interpret differences in LMS usage as distinct learning tactics (Jovanović et al., 2020). A well-established connection exists between

the use of learning tactics and strategies and learning outcomes (Jovanović et al., 2017; Siadaty et al., 2016).

Sequence analysis is a methodology used to execute pattern mining (Abbott, 1983). It has been used in various disciplines, but its use in education, has been brief (e.g., Saqr, López-Pernas, Jovanović, et al., 2023). Sequential mining is useful for examining learning from a processual perspective. The process unfolds in time transformative, and it can be described as states (Schmitz, 2006). Sequence mining aims to identify repeating patterns or sequences, which helps us identify learners with similar patterns.

In recent years, LA has begun to use sequence analysis. Typically, researchers use the trace data captured from LMS to perform a sequence mining study (Jovanović et al., 2017; Matcha, Gašević, et al., 2020). A sequence analysis study generally involves six steps: 1) coding the elements of the sequence to form the alphabet, 2) determining the timespan used in the analysis, 3) defining the actor for a sequence object, 4) visualizing the descriptive analysis, 5) grouping the similar sequences into clusters, and 6) analyzing the clusters (Saqr, López-Pernas, Helske, et al., 2024).

The alphabet refers to the states of the sequences (Raab & Struffolino, 2023). Recoding takes place in this step of the process because the raw data (e.g., Moodle log data) is often too complicated to be interpreted meaningfully. Therefore, the raw data must be re-coded carefully to simplify the data structure (Liao et al., 2022).

Trace data forms sessions, which are the basic unit of analysis in sequence analysis. The unit of analysis used in sequence analysis is typically called a session (Jovanović et al., 2017; Matcha, Gašević, et al., 2020) referring to an uninterrupted sequence of activity in online learning. Most often, sessions are determined based on the uninterrupted sequence of the learner's actions (Jovanović et al., 2017; Matcha, Gašević, et al., 2020). While most studies focus on intra-course sequences, some studies expand the timespan over the entire study program (Saqr, López-Pernas, Jovanović, et al., 2023; Saqr & López-Pernas, 2021, 2022).

Figure 8 illustrates how learning activities can be coded into sequence states. The figure shows eight different sequences (vertical axis) over three time periods (horizontal axis), with activities color-coded into five distinct states: Forum (green), Resources (orange), Assignment (purple), Feedback (navy), and Quiz (light blue). Each state represents a different type of learning activity that students engage in during their online learning process. For example, in some sequences, students begin with forum activities before transitioning to assignments, whereas others start with resource consultation. The temporal progression across the three time periods demonstrates how students' engagement with different learning activities evolves during their learning journey. This visualization helps identify patterns in how students transition between different learning states and reveals the varying pathways students take through the course material. Such sequence state representations are fundamental to sequence analysis, enabling researchers to identify common patterns and distinguish different learning strategies among students.

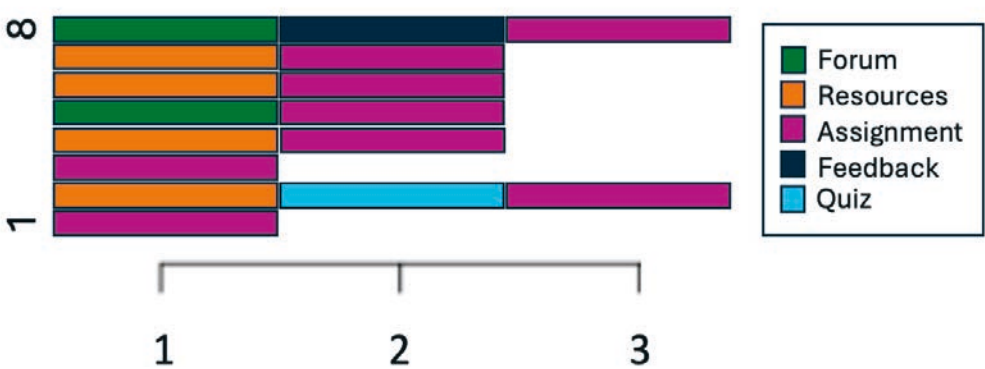


Figure 8. Simplified representation of sequences representing typical sequences and alphabet used in LA studies

The actor in LA studies is typically the learner (López-Pernas & Saqr, 2021). Other approaches have analyzed the group of collaborating students (Saqr & López-Pernas, 2022); in this case, the whole group is the actor. To our

knowledge, no studies have used the teacher as an actor. As LA is focusing on learning, not teaching, this is an understandable approach.

Different libraries are used in the sequence-building step. TraMineR (Gabadinho et al., 2011) is one of the most used libraries, and it can be used to visualize the results of sequence building. Visualization helps to identify the dissimilarities between sequences, and here, different plots are used to get a holistic view of the sequences.

When visualization is performed, and the most suitable distance measurements are selected, the next step is typically to cluster the sequences to identify similar patterns (Studer, 2013). This step aims to find groups of patterns where sequences within a cluster are as similar as possible and as different from other patterns in other clusters as possible. Different clustering techniques have been used, but Agglomerative Hierarchical Clustering (AHC) has been the most popular (Saqr, López-Pernas, Helske, et al., 2024). Clustering uses distance measures to find dissimilarities between clusters. Depending on the data used and the aims of the study researchers have used different measures, such as Euclidean (Matcha, Gašević, et al., 2020; Saqr, López-Pernas, Jovanović, et al., 2023), Longest Common Subsequence (López-Pernas et al., 2023), Longest Common Prefix (Saqr & López-Pernas, 2022), and Optimal Matching (Jovanović et al., 2017). As these approaches have been unable to fulfill the need for analyses, other approaches have been developed. Markovian Models is an example of a method that is gaining more ground (Matcha, Gašević, et al., 2020; Saqr & López-Pernas, 2022). New approaches are used to identify and apply metrics and indicators for understanding learners' sequential patterns in more detail (López-Pernas & Saqr, 2024; Poquet et al., 2023; Saqr, López-Pernas, Jovanović, et al., 2023; Tan & Samavedham, 2022).

States used in sequence mining studies require the data to be presented as categorical. This is typically done by aggregating the raw data into categories. These categories include behavioral states (Raab & Struffolino, 2023), learning activities (Matcha, Gašević, Jovanović, et al., 2020), moves in educational games (López-Pernas et al., 2023), or different roles adopted

during collaborative learning (Saqr & López-Pernas, 2022). A common approach is to define learning actions based on interactions with LMS (López-Pernas et al., 2021; Matcha et al., 2019; Saqr, López-Pernas, Jovanović, et al., 2023; Villalobos et al., 2022). Instead of plain activity descriptions, the interactions can be transformed into meta-cognitive actions (e.g., SRL strategies) to analyze the activity using existing theoretical concepts (Dever et al., 2023; Poquet et al., 2023). In sequence analysis, the concepts of learning tactics and learning strategies are typically used. A learning tactic is defined as a short sequence of a learner's cognitive operations, and a learning strategy is a chronologically ordered learner's sequence of decisions regarding which tactics to use (Winne & Marzouk, 2019). Fincham, Gašević, et al. (2019) created a simplified definition that strategies are sequences of tactics applied by students. It is also possible to continue the analysis by sequencing the sequences and thus analyzing learning, for example at the course level (Fincham, Gasevic, et al., 2019; Saqr, López-Pernas, Jovanović, et al., 2023; Villalobos, 2022).

For example, Matcha, Gašević, Ahmad Uzir, et al. (2020) analyzed how tactics and strategies vary with course design and modality. They identified unique learning actions tailored to each course and found that students' learning tactics reflected the course design. In contrast, learning strategies were more influenced by the course format than the instructional design.

The importance of analyzing strategic transitions in learning has been further highlighted in recent research. T. S. Cristea et al. (2025) employed clustering techniques and Markov modeling to examine how students transition between different learning tactics and strategies throughout their courses. Their methodology for analyzing unobtrusive LMS data complements our approach in [II]. Their work particularly reinforces our findings about the importance of understanding how students adapt their learning strategies throughout a course, though they focus more on the effectiveness of different strategic transitions.

Sequence analysis can be used to track student engagement, including activity level or study regularity. Student engagement indicators can have varying levels of temporal granularity, ranging from daily or weekly data

analysis to actions performed during specific learning sessions at particular times (Jovanović et al., 2021). Other studies have used the COPES model (Winne & Hadwin, 1998) to compute indicators for task definition, goal-setting, enactment, and adaptation (T. Cristea et al., 2024). This approach helps to connect the indicators with a theoretical level to trace SRL-related constructs.

Ritschard (2023) used a different approach to divide metrics into three categories of measuring sequences: a) basic indicators (counts of states), b) diversity (the heterogeneity of visited states), and c) complexity (unpredictability of state arrangement). The basic indicators are similar to the metrics proposed by Jovanović et al. (2021); the indicators are based on the count of interactions. Diversity metrics focus on the state distribution in the sequence, including the course strategies students use during an education program (Saqr, López-Pernas, Jovanović, et al., 2023). Complexity metrics refer to the instability or unpredictability of state arrangement in the sequence, such as students' procrastination behaviors in an online course (Tan & Samavedham, 2022).

Sequence analysis has been used for varied purposes. The sequential relationship was analyzed using students' performance (Kinnebrew & Biswas, 2012; Matcha, Gašević, et al., 2020; Saqr, López-Pernas, Jovanović, et al., 2023), the patterns of course delivery (Matcha, Gašević, Ahmad Uzir, et al., 2020), or dropout (Saqr & López-Pernas, 2021). For example, López-Pernas & Saqr (2024) explored student engagement patterns and observed three patterns of behavior: Active, Average, and Disengaged. They used integrative potential to connect different data sources and theories to gain more comprehensive insights into learning processes and classify students based on this likelihood of achieving a positive state, such as Active student engagement. Tan & Samavedham (2022) used the precarity index (a quantitative measure used to assess instability) to analyze students' procrastination in an online course to cluster students based on the time they spent watching videos. Saqr, López-Pernas, Jovanović, et al. (2023) performed a longitudinal study to analyze students' learning strategies and

behavior variation across a learning program and observed similar context-specific actions across courses.

Temporal sequences can be analyzed in different ways. Molenaar & Wise (2022) proposed a framework for learners' temporal behavior characteristics. They defined the passage of time and order in time as two essential temporal concepts. The passage of time views time as a continuous stream of events, introducing four metrics—position, duration, frequency, and rate—applicable across statistical methodologies. Order in time requires advanced techniques, exploring consistency, recurrent and non-recurrent change, and irregular change within event sequences.

Trajectories can be analyzed with epistemic network analysis, which is used to explore the differences between different efficacy clusters regarding their SRL, and the high SE group demonstrates more SRL behavioral characteristics in the performance and self-reflection phases of the SRL cycle (P. Chen et al., 2024).

Sequence analysis methods offer a promising approach to understanding the needs of students with low SEB – those who believe that higher grades are either too hard to achieve or require too much work. These methods involve analyzing patterns in students' learning behaviors over time, providing insights into their learning strategies and potential areas for intervention.

Sequence analysis in educational contexts involves examining the order and timing of students' learning activities. This method can reveal patterns that are not apparent when looking at aggregate data. For example, Jovanović et al. (2017) applied sequence analysis to data from a flipped classroom environment. They identified several distinct learning strategies, such as “comprehensive” (engaging with all types of learning activities) and “selective” (focusing on specific types of activities). To fully appreciate the potential of this method, it is important to understand how sequence analysis is conducted in educational contexts.

Natthaphatwirata & Matcha (2024) used sequence mining to formulate student engagement sequences. They used hierarchical clustering to recognize similar patterns in formulated sequences and, as a result, found

four groups of learners distinguished by similar student engagement patterns. They found two out of four groups to be successful learners, who are intensively engaged in learning activities in the short and long run, whereas unsuccessful groups tend to be less engaged.

Another study (Lima et al., 2023) explored the interplay between student engagement and SRL using temporal analysis methods to identify distinct profiles of SRL among students based on their student engagement patterns. By analyzing the temporal sequences of student interactions, the study reveals how different levels of student engagement are associated with varying degrees of self-regulation. According to their results, students who exhibit higher student engagement through frequent interactions with learning materials tend to demonstrate more effective SRL strategies, such as goal setting and self-monitoring. This finding underscores the cyclical nature of SRL, where increased student engagement can lead to enhanced self-regulatory behaviors, which in turn can foster further student engagement.

Recent work by Villalobos, Pérez-Sanagustín, et al. (2024) has expanded on traditional sequence analysis methods by introducing a set of indicators to characterize individual students' temporal behavior dynamics. Their approach, which includes basic, diversity, and complexity indicators, offers a more nuanced view of how students engage with online learning environments over time. Unlike traditional clustering methods that group similar behaviors, these indicators can reveal individual differences in learning processes. In their study of a blended learning course, they found that students who engaged with a LAD showed more complex and diverse behavior, although this did not translate to significant differences in final grades. Interestingly, they found few correlations between these sequence indicators and students' self-reported SRL skills or prior achievement, suggesting that pre-course measures may not be good predictors of students' behavioral dynamics during the course. This work highlights the potential of sequence analysis to provide insights into the temporal aspects of learning that may not be captured by other methods, particularly when evaluating the impact of interventions.

Sequence analysis methods can also be used to understand the needs of students with low SEB – those who believe higher grades are either too difficult to achieve or require too much work. While we know students sometimes struggle with assignments (Lopez-Pernas & Saqr, 2021), it is not always clear how best to help them. Students have different goals, and not all aim for the highest grades, sometimes due to low SEB (e.g., Knouse et al., 2014). When designing interventions, it's essential to identify students aiming for higher grades but lacking sufficient skills and distinguish them from those satisfied with passing grades. Targeting interventions at the former group is expected to produce better outcomes (Dahling & Ruppel, 2016).

In the context of SEB, sequence analysis could reveal differences in how students with high and low SE engage with course materials. For instance, students with high SE might show more consistent student engagement patterns and quicker rebounds after setbacks. Meanwhile, those with low SE may display more erratic patterns or prolonged disengagement after difficulties. This research contributes to this discussion by analyzing student engagement along with SRL at a granular level. When learning processes are analyzed multiple times during a course, it can reveal when the distinction between learners with high and low SEB becomes evident. This helps practitioners identify the right time to assist students who need help in achieving their personal goals, thus developing their SRL and helping them to improve their SEB.

Traditionally, sequence analysis studies have focused on a single data channel to analyze the learning progression. While learning is a complex construct, a single data type is not necessarily enough to understand its nuances. Therefore, this research uses parallel datasets to conduct multi-channel sequence analysis to understand the interplay between SRL and student engagement.

2.6.5 Multi-channel Sequence Analysis

Multi-channel sequence analysis has emerged as a sophisticated approach in LA (López-Pernas & Saqr, 2021; Saqr, López-Pernas, Helske et al., 2023),

and it is particularly valuable for understanding the complex, multifaceted nature of learning processes (Jovanović et al., 2021; López-Pernas & Saqr, 2024). This method analyzes parallel sequences of categorical data, providing insights into how different aspects of learning interact and evolve. Unlike single-channel approaches that might focus solely on behavioral traces or self-reported data, multi-channel analysis enables researchers to examine multiple dimensions of learning simultaneously (Saqr & López-Pernas, 2024).

Multi-channel sequence analysis requires at least two temporal datasets offering different types of information, which can be analyzed in parallel. Previous studies have used different types of data to create the parallel channels elemental for multi-channel analysis. For instance, clickstream or log data has been frequently used as one channel to capture behavioral aspects of learning (López-Pernas & Saqr, 2021; Ouyang et al., 2023; Saqr, López-Pernas, Helske, et al., 2023), and complemented with assessment data (Bacci & Bertaccini, 2022; López-Pernas et al., 2021; Saqr, López-Pernas, Helske, et al., 2023), video data (Törmänen et al., 2022, 2023), or even physiological measurements like electrodermal activity (Törmänen et al., 2022, 2023) to analyze learning. A notable example of innovative multi-channel data application was provided by Yan et al. (2024), who conducted a study using audio, positioning, and physiological data to create multi-channel data incorporating 17 multimodal behavioral features to analyze collaborative learning in a classroom environment. This demonstrates the method's flexibility in incorporating diverse data sources to elucidate learning processes.

A crucial step in multi-channel sequence analysis is the aggregation of raw data into meaningful states. This process requires a careful consideration of how to categorize and combine different types of data while maintaining their interpretability. Figure 9 demonstrates how multi-channel sequence analysis combines different data streams to provide a more comprehensive view of learning processes. The visualization shows three parallel channels: engagement (top), combined states (middle), and regulation (bottom). The engagement channel categorizes student behavior into Active (green),

Average (yellow), and Disengaged (red) states. The regulation channel depicts two states of learning regulation: HighReg (dark blue) and LowReg (light blue). The middle channel represents the combined states resulting from the interaction between engagement and regulation levels, showing how these two dimensions interact at each time point. The vertical arrows between channels illustrate how states from the engagement and regulation channels combine to create more nuanced state descriptions in the combined channel. For example, when high engagement (Active) aligns with high regulation (HighReg), it results in an Active/HighReg state (dark green) in the combined channel. This visualization effectively demonstrates how multi-channel analysis can reveal patterns that might not be apparent when examining engagement or regulation in isolation, providing a more complete understanding of students' learning processes.

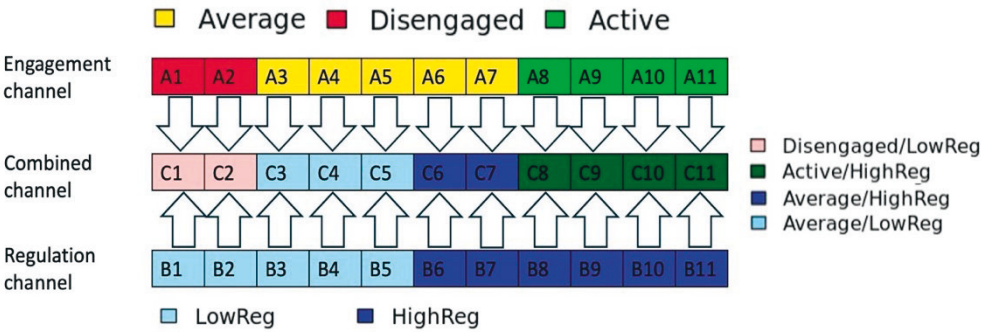


Figure 9. Simplified version of multichannel sequences used to form a combined sequence

The channels used in the current body of knowledge have varied. Previous studies have explored the relationship between emotional valence and arousal (Törmänen et al., 2022), tactics for studying online and automated assessments (Lopez-Pernas & Saqr, 2021), and student engagement and achievement (Saqr, López-Pernas, Helske, et al., 2023). While it is possible to include multiple channels in the analysis, researchers must balance complexity with interpretability. For example, Ouyang et al. (2023) used five channels: social interactions, cognitive dimension, collaboration regulation,

students' online behaviors, and the socio-emotional dimension. However, this high number of channels increases the complexity of interpretation due to the multiplicative effect on possible state combinations.

Hidden Markov Models (HMM) have become a standard tool for analyzing multi-channel sequence data in educational contexts. In HMM analysis, two types of states exist: observed and hidden (also known as latent) (Lopez-Pernas & Saqr, 2021). The observed states are the directly measurable behaviors or responses captured in the data, whereas hidden states represent underlying patterns or processes that cannot be directly observed but can be inferred from the observable data (S. Helske & Helske, 2019).

Recent applications of HMM in educational research have revealed important insights about learning processes:

1. **Strategy Identification:** Studies have used HMM to analyze strategies and transition probabilities of students moving between different learning states (López-Pernas et al., 2021; Saqr, López-Pernas, Jovanović, et al., 2023).
2. **Temporal Analysis:** When hidden states are analyzed over time, distinct trajectories often emerge, revealing patterns in how students progress through their learning journey (Törmänen et al., 2023).
3. **Regulatory Processes:** HMM has been used to assess fluctuations in socio-emotional interaction episodes and regulation of learning (Törmänen et al., 2023).
4. **Group Dynamics:** Studies have applied HMM to understand collaborative group patterns and how they evolve (Ouyang et al., 2023).

Multi-channel sequence analysis has provided several key insights into educational processes. For example, Lopez-Pernas & Saqr (2021) found three hidden states of learning strategies by combining LMS usage data with automated assessment tool data. This revealed distinct patterns in how students transition between different learning approaches and how these transitions relate to academic outcomes.

The multi-channel sequence analysis has been particularly valuable in understanding complex educational phenomena such as the relationship between emotional states and learning regulation (Xu & Zhou, 2024), how different aspects of engagement interact over time (He et al., 2024), the connection between behavioral patterns and learning outcomes (López-Pernas & Saqr, 2021), and the dynamics of collaborative learning processes (Ouyang et al., 2023).

These applications demonstrate the utility of multi-channel sequence analysis in capturing the multifaceted nature of learning processes and how different aspects of learning interact over time. The ability to analyze multiple parallel sequences of data provides a more nuanced and comprehensive understanding of learning processes than single-channel approaches alone could provide.

The integration of multiple data channels has enabled researchers to build more sophisticated models of learning processes, accounting for both observable behaviors and latent constructs. This has proven particularly valuable in online learning environments, where understanding the interaction between different aspects of learning is crucial for developing effective support strategies and interventions.

2.6.6 Process Mining

Process mining has emerged as another valuable perspective for studying patterns in educational data, enabling researchers to analyze how students interact with content and activities in learning management systems (Bogarín et al., 2014). A comprehensive review by Bogarín et al. (2018) mapped the landscape of educational process mining, revealing diverse applications across multiple educational contexts. The technique has proven particularly valuable in investigating student activity sequences in online environments, including MOOCs, blended courses, and computer-supported collaborative learning settings.

Understanding students' learning processes through process mining enables the development of improved teaching processes and learning designs (Alqaheri & Panda, 2022; Stoyanov & Kirschner, 2023). One primary

application has been detecting learning difficulties to provide more targeted student support. For example, López-Pernas et al. (2021) used process mining to explore how students transition between learning management systems and automated assessment tools, thus identifying how struggling students utilize resources to overcome challenges. Similarly, Arpasat et al. (2021) examined differences in activity patterns, bottlenecks, and time performance among students with varying achievement levels in MOOCs.

Process mining studies in educational contexts typically focus on several key methodological approaches. Directly-Follows Graphs (DFGs) are used to understand commonly performed activities and transitions between them. Figure 10 is a simplified example of the process map. The map reveals several key patterns in how students navigate through learning activities. Beginning from the "Start" node, students typically follow three initial pathways: either directly to the forum activities (25% of learners), to the assignment (25%), or to the learning resources (50%). Based on these initial activities, students show various trajectories toward completing assignments. Notably, the map indicates multiple paths leading to assignment completion: students may move directly from resources, transition from forum discussions, or access assignments after attempting quizzes. The process ultimately converges at the "End" node, which can be reached through multiple paths. This convergence suggests that while students may take different routes through the learning activities, they ultimately need to complete certain core elements to finish the course. Such process maps can help educators understand common learning pathways and identify potential points for intervention or support.

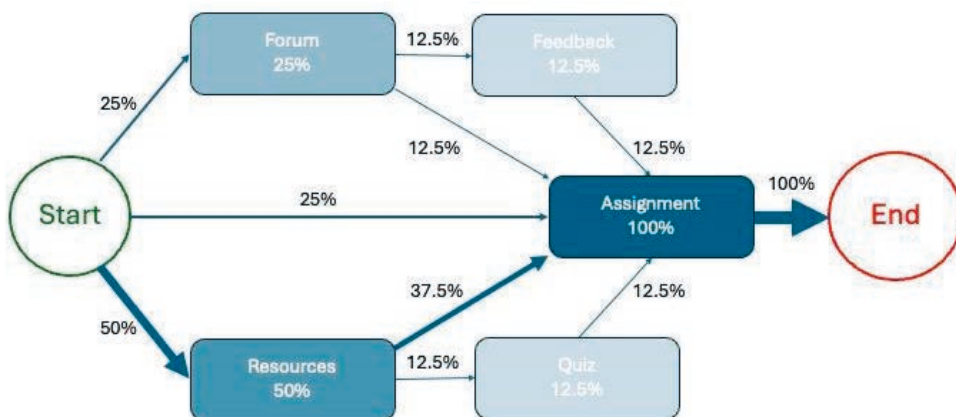


Figure 10. Simplified version of the process map. Nodes represent typical learning activities and edges represent transitions between nodes. The width of the edge shows the relative commonness of the particular transition

For instance, Vartiainen et al. (2022) used DFGs to visualize transitions between in-game activities in educational escape rooms. Researchers have often used DFGs to compare behavioral patterns between different student groups, such as high versus low achievers or between patterns obtained through different methods. Saqr & López-Pernas (2022) demonstrated this by implementing k-means clustering to group students by online activity frequency, and then used DFGs to understand the learning strategies adopted by different groups. Process Discovery is another approach, which has been used by several studies by employing process discovery algorithms to detect underlying patterns in observed data (Bogarín et al., 2018). These include the alpha algorithm (Nafasa et al., 2019), the heuristic algorithm (Bogarín et al., 2014), and the fuzzy miner (Beheshitha et al., 2015). Some researchers perform conformance checks to compare observed processes against designed processes. Pechenizkiy et al. (2009) exemplified this approach by verifying whether students answered exam questions in the prescribed order.

Markovian methods represent another significant approach in process mining, particularly for studying transitions between different learning states. These methods have gained popularity due to their ability to map

students' transitions between learning activities using stochastic processes mining. First-order Markov models are commonly employed to understand these transitions. For example, Matcha et al. (2019) applied first-order Markov models to examine how students transition between different learning tactics. This approach has been extended to various educational contexts, including transitions between academic writing tactics (Peeters et al., 2020), patterns in SRL events (Lim et al., 2023), and dynamics within collaborative learning settings (Saqr & López-Pernas, 2023).

These Markovian approaches complement other process mining methods by providing probabilistic insights into learning transitions and enabling researchers to understand not only what sequences occur but also their likelihood and patterns of change. While earlier work in this area relied heavily on the pMiner R package (Gatta et al., 2017), recent developments have introduced more modern and flexible methods for fitting, plotting, and clustering stochastic process mining models. These newer approaches also allow for the inclusion of covariates to help explain why different transition patterns emerge. Figure 11 shows the different approaches of process mining and the goals for the different approaches.

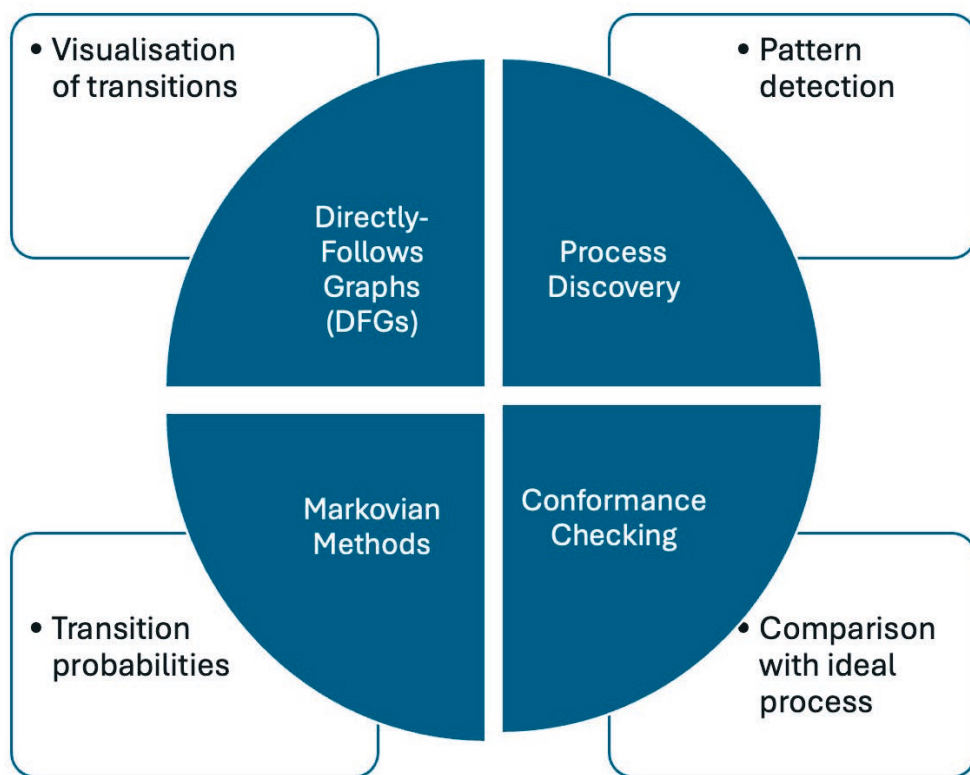


Figure 11. Process mining approaches in educational research

A considerable volume of research has focused on instantaneous events, such as clicks in online learning management systems (Dolak, 2019; Hachicha et al., 2021; Romero et al., 2016). Fewer studies have examined activities with defined start and end times, primarily due to data collection limitations in online platforms. However, researchers have addressed this limitation by grouping clicks into learning sessions, an approach common in studies of learning tactics, strategies, and SRL (Cerezo et al., 2020; López-Pernas et al., 2021; Matcha, Gašević, et al., 2020).

The relationship between instructional design and students' tactics has been a focus of process mining research. Fan et al. (2021) found that students' learning strategies and tactics were significantly influenced by course instructional design and varied based on performance levels. Sequence analysis complements this approach by examining similarities in

student action sequences (Fincham, Gasevic, et al., 2019; Jovanović et al., 2017).

Recent developments have expanded process mining applications in several directions. Process mining research has been integrated with SRL research. He et al. (2024) employed process mining to examine the interplay between SRL and student engagement, viewing SRL as a cyclical process with interrelated phases. Their findings revealed increased task-solving and help-seeking behaviors following formative assessments, indicating positive shifts in SRL patterns. However, their course-level analysis suggests the need for more granular investigations to understand how student engagement and SRL co-evolve during a course.

Recent studies have incorporated multiple data types and analytical approaches. Fan et al. (2024) investigated relationships between performance and motivation using multi-channel process mining analysis. Their study of the impact of AI tools on academic performance revealed that while AI uses improved performance metrics, it does not necessarily translate to enhanced knowledge levels.

Researchers have employed various software tools for process mining analysis. Disco has been widely used for DFG visualization (Juhaňák et al., 2019), whereas ProM dominates process discovery and conformance-checking applications (Ariouat et al., 2016; Cerezo et al., 2020). Many researchers have also utilized the R programming language with the Bupaverse (Janssenswillen et al., 2019) framework for basic metrics and visualization (López-Pernas et al., 2021).

The integration of process mining with other analytical approaches offers insights into learning behaviors in online environments. When combined with sequence analysis and multi-channel analysis, process mining provides valuable insights into how students engage with and regulate their learning. Recent efforts have focused on developing frameworks to support learning design at scale (Wambsganss & Schmitt, 2024) and employing innovative approaches like epistemic network analysis to understand variations in learning processes across educational levels (Cheng et al., 2024). The integration of learning process data with affective or motivational measures

has been garnering increasing interest. For instance, Xu & Zhou (2024) conducted longitudinal research examining SRL alongside social-emotional need satisfaction, demonstrating the value of considering multiple dimensions of the learning process. Similarly, Song et al. (2024) explored the longitudinal relationship between SRL and academic success, finding positive correlations between these factors over time.

The field continues to evolve, incorporating more diverse data sources and analytical methods while maintaining a focus on understanding the complex nature of learning processes. The integration of traditional LA with newer forms of data and analysis techniques promises deeper insights into how different aspects of learning interact and develop over time. This methodological advancement aligns with the broader goals of LA research, namely supporting more effective and personalized learning experiences.

3 RESEARCH DESIGN

This chapter presents the methods used in this research. The first empirical study focused on sequence analysis and the process mining of business students in relation to SEB. The second empirical study continued the analysis by incorporating a multi-channel sequence analysis approach to understand the learning processes.

The methodological progression across the three papers reflects an increasingly sophisticated approach to understanding SRL in online environments. While Paper [I] employed systematic review methodology to map the landscape of LA interventions, Papers [II] and [III] utilized advanced analytical techniques including sequence analysis, process mining, and multi-channel clustering to examine actual student behaviors and their regulation patterns. This methodological evolution allowed for increasingly nuanced insights into how students regulate their learning and how LA can be effectively employed to support this process.

The research context evolved across the three papers to provide complementary perspectives on SRL. Paper [I]'s broad review of LA interventions across various educational settings provided the contextual foundation. Paper [II] focused on business education, examining project management learning in detail. Paper [III] extended this by investigating engagement patterns in online project management education, allowing for a more comprehensive understanding of how students regulate their learning in authentic educational settings. This progression from broad review to specific contextual applications helps illuminate how SRL manifests in different educational environments and how it can be effectively supported.

3.1 Identifying LA Interventions Aimed to Support Learners

The first step of this research initiative was to perform a systematic literature review (SLR). Literature is the first step in research and aims to understand

the state-of-art and identify gaps and challenges in the field (Carrera-Rivera et al., 2022), whereas systematic literature is a methodological approach used to collect, identify, and critically analyze previous studies using a thorough systematic procedure (Pati & Lorusso, 2018). SLR was selected as to obtain a comprehensive understanding of the current state of the field. We wanted to understand the field's current status, considering how LA is used to inform educators in their journey to understand and support online learners.

We followed the PRISMA 2020 (Page et al., 2021) statement in our SLR, as it has been proven to be a robust methodology for conducting a systematic literature review. The systematic nature removes the arbitrary elements, such as snowballing, and will result in replicable findings and increase the transparency of the process. While there were three points of interest, namely a) LA, b) SRL, and c) interventions, this framework set clear boundaries for the body of keywords to be explored. After the initial searches, the concept of intervention was found to be somewhat ambiguous and required a more detailed exploration. Whereas LA and SRL have fewer concepts that could be used to replace them as keywords, intervention might have been replaced with different concepts such as experiment and controlled trial. Therefore, the existence of intervention had to be considered manually to ensure that no critical studies were excluded. While two out of the search term clusters remained fixed, we reduced the reviewer bias. Table 1 represents the overview of our SLR.

Table 1. Overview of the systematic literature review

Research questions	RQ1: What learning analytics methods have been applied to identify the SRL support needs? RQ2: What channels have been applied in LA interventions to foster SRL? RQ3: Which phases of the SRL cycle were targeted by LA interventions? RQ4: How do studies evaluate the SRL support efficiency and impact, and what types of results were achieved?
Databases	Scopus, Web of Science, Google Scholar, Journal of Learning Analytics, the Journal of Educational Data Mining, and the International Conference on Educational Data Mining
Search string	("self-regulat*" OR "self regulat*") AND ("educational data mining" OR "learning analytics")
# of papers identified	753 papers
Inclusion criteria	1. The manuscript reported a learning analytics analysis or an intervention hypothesized to support or enhance learners' SRL. 2. The full text was available. 3. The record was written in English. 4. The record was a peer-reviewed empirical paper. 5. The record was published prior to the end of 2022.
# of papers included	56
Coding process	RQ1: (Romero & Ventura, 2020) RQ2: iterative and inductive RQ3: (Panadero, 2017) RQ4: iterative and inductive
Coding validity	Social validity
Methods	Descriptive and comparison of effect sizes

The systematic review process followed several distinct phases, as illustrated in Figure 12. In the identification phase, database searches initially yielded 753 records. After removing 137 duplicate records, 616 unique records remained for screening. During the screening phase, these records were assessed based on titles and abstracts, resulting in 421 records being excluded as they did not meet the initial criteria. The remaining 195

reports were sought for full-text retrieval, all of which were successfully accessed. These reports were then assessed for eligibility based on detailed inclusion criteria, leading to the exclusion of 139 reports that were not directly relevant to the topic of LA interventions supporting SRL. Finally, 56 studies met all inclusion criteria and were included in the systematic review. This rigorous selection process ensured that only the relevant studies were included in the final analysis.

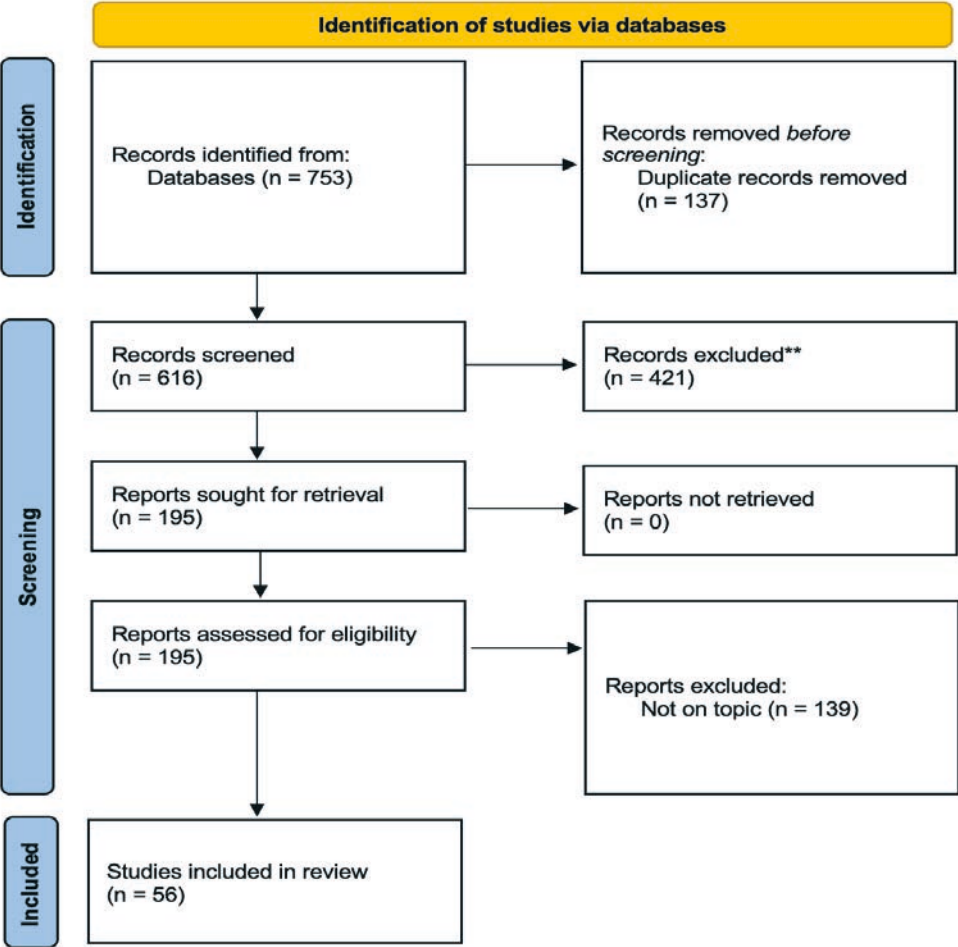


Figure 12. PRISMA flow diagram of the study selection process for the systematic literature review

SLR aimed to identify the research gaps and ground our empirical studies [II and III] on this ground. In particular, this SLR reviewed what LA methods previous studies have used to support SRL (RQ1). Second, we wanted to investigate what channels are used to mediate the LA information to learners (RQ2). Third, we explored which phases of the SRL cycle (Panadero, 2017) have been supported with interventions (RQ3). Lastly, to examine the impact of the proposed interventions to find out if some of the applied methods or channels are the most impactful in promoting SRL (RQ4).

While administering the SLR process, we used Covidence, an online tool, to manage the records database, review the abstracts and full texts, codify the findings and extract the results of the review. Covidence made the reporting and administration of the process more straightforward and concise and mitigated the possible errors caused by manual work during the review process. The coding of RQ1 used the framework proposed by Romero & Ventura (2020), and RQ3 used the division introduced by Panadero (2017). For RQ2 and RQ4, we categorized the findings in themes using an iterative and inductive coding process (Braun & Clarke, 2006) of the data-driven categorization (Thomas, 2006). To improve the coding validity, a study supervisor resolved the disagreements and uncertainties once they were faced during the coding process. We used descriptive analysis to explain the results for RQ1-RQ3, whereas the results of RQ4 were analyzed by performing a comparative analysis of effect sizes to assess the impact of the reviewed studies.

3.2 Relation Between SEB and Temporality of Learning

Once the SLR results were finalized, we were able to determine the goals of the empirical studies. We decided to conduct a sequence analysis of the learning processes of business students in the context of a project management online course. We wanted to explore how the learning processes of project management students vary during online learning. We established a connection with the previous studies by focusing on the learning tactics and learning strategies. To investigate these differences, we decided to use sequence analysis and process mining methodologies.

Furthermore, as we could investigate the differences between students' individual learning goals, we could incorporate SEB theory into our research to explain the differences in the learning processes. Based on these preliminary ideas, we decided to formulate the research questions for the first empirical study.

The research questions for [II] were as follows:

RQ1 What type of learning tactics do project management students use when studying online (RQ1.I), could the different learning tactics be used as an indicator towards achieving their goals and (RQ1.II)?

RQ2 How do learning tactics differ between low- and high-SEB students (RQ2.I) and what are the differences in learning processes between low- and high-SEB students (RQ2.II)?

3.2.1 Research Methodology

In [II], we performed a sequence analysis and process mining study. The goal of this study was to investigate the relationship between students' SEB, goal-setting, and learning tactics in an online business course.

The recent research examined in the literature review section revealed a need to research LA-supported SRL in the context of business education. It aims to address this gap by examining how LA can help understand different learning tactics between students with low and high SEB in a project management course. We used a data-driven approach that has been adopted in previous studies (Elmoazen et al., 2022; López-Pernas & Saqr, 2021; Zhang & Paquette, 2023).

We used TraMiner (Gabadinho et al., 2011), and BupaR (Janssenswillen et al., 2019) libraries to analyze the temporality of the learning processes because these methods were executed using the statistical software R and have been proven to give a solid methodological basis for the research. These methods were used to analyze the data, which was collected from LMS.

3.2.2 Data Collection

We used three data sources: a) the student's numeric goal set (the overall grade for the course) in the deliverable of the first assignment, b) LMS log data, and c) the final grade. The numeric goals were used to assess the students' SEBs. Log data is a standard data type used in the LA study to explore the temporality of learning processes (B. Chen et al., 2018; Siemens & Long, 2011). The final grades were used to assess if the students had achieved their intended goals and to evaluate their SEB.

Several validated instruments have been developed specifically for measuring SEB in learning contexts. P. Pintrich et al. (1991) established a foundation using the Motivated Strategies for Learning Questionnaire (MSLQ), which includes a self-efficacy subscale measuring students' judgments about their ability to accomplish academic tasks through self-reported items on a Likert scale. Building on this framework, B. Zimmerman & Kitsantas (2007) developed the Self-Efficacy for Learning Form (SELF), a scale assessing students' beliefs about their capability to implement various self-regulatory strategies. In online learning environments specifically, researchers have developed context-appropriate measures (Shen et al., 2013; W. A. Zimmerman & Kulikowich, 2016). Our approach of using expected course grades as a proxy for SEB is supported by prior research establishing strong connections between these constructs. B. J. Zimmerman et al. (1992) demonstrated that students' self-efficacy beliefs for self-regulated learning affected their perceived self-efficacy for academic achievement, which in turn influenced the academic goals they set for themselves. More recently, Talsma et al. (2019) observed strong correlations between academic self-efficacy and performance expectations, with self-efficacy being a significant predictor of expected performance.

The course was organized in LAB University of Applied Sciences. The students participating in the course were first-year bachelor students. The course is 5 European Credit Transfer System (ECTS) credits. It was self-paced, and the students had 14 weeks to complete it. At the beginning of the course, students set a numeric goal, which determined the number of

assignments a student had to complete to achieve the grade they had set. All learning materials were already available for the students at the beginning of the course until the end of the course. For each assignment, there were predetermined rubrics available at the LMS showing the minimum criteria for the submissions. A student could evaluate their deliverables before submission by reviewing their materials against the rubric to make sure the minimum criteria were met before submission. The LMS used was Moodle.

3.2.3 Data Preparation

First, we cleaned the data. We removed the non-learning activities (e.g., teacher's actions and non-active learning activities: profile views, logging into the course, etc.) to reduce the noise in the data set. We then anonymized the student ID to follow ethical protocols and protect the privacy of the participants. To suppress the complexity of the trace we used re-coding to categorize the activities into fewer categories (Assignment, Course viewed, Deliverable, Feedback, Forum, Interaction, Learning resource, Monitoring, Quiz) as has been done in previous studies (Molenaar & Chiu, 2014; Saqr, López-Pernas, Jovanović, et al., 2023). Finally, we coded the students into three SEB cohorts, to analyze the differences between these cohorts. After preparing the data, we had 56,047 rows of data ready for analysis.

3.2.4 Data Analysis

We used sequence analysis to explore students' learning tactics and strategies when studying online. A similar approach has been used in previous studies (Lopez-Pernas & Saqr, 2021; Saqr, López-Pernas, Jovanović, et al., 2023).

To identify the distinct learning tactics between low- and high-SEB students, we used TraMineR package (Gabadinho et al., 2011). We used re-coded trace data as input for the analysis, and, as a result, we visualized the learning sequences (Bogarín et al., 2018; Saqr, López-Pernas, Jovanović, et al., 2023) to help us interpret the key differences between different sequences. A time-wise organized dataset was used to create sessions. A

session consisted of activities that were within a 30 min timeframe of the previous and next activities. This time window was chosen to exclude overly long session outliers. After removing the excessively short (< 3 actions) and long (> 50) sessions, we obtained 6,688 sessions.

Next, we used AHC to identify patterns of similar sessions. We used different distance measurements and fit indices (Gabadinho et al., 2011) to define the optimal number of clusters (López-Pernas & Saqr, 2021; Uzir et al., 2020). We tried different options, and the Spell-length-sensitive optimal matching (OMslen) method resulted in four interpretable patterns. OMslen is a modified version of optimal matching (OM) weighing the OM elementary operations inversely with episode length (Halpin, 2009), which was found to be the most suitable solution for the current study. When interpreting the results, we used sequence distribution, index, and implication plots to evaluate the results (Saqr, López-Pernas, Jovanović, et al., 2023), and understand the key differences between different sequences.

To test the differences between different tactics and goal achievement, we used Spearman's correlation. Here, we used different tactics, goal-setting, and final grades as indicators to determine how they are related to achieving learning goals.

We applied process mining methodology to visualize the order of consecutive actions students take to obtain the differences in the learning processes and learning tactics. We used the BupaR (Janssenswillen et al., 2019) library to create process maps representing learning tactics. Visualized process maps helped us to understand how the different tactics evolve over time. The visualizations of the different tactics show the probabilities of different consecutive actions students take (Chiu & Reimann, 2021) while elucidating what are the actions students are most likely to perform during their online learning sessions. We used the processmapR and DiagrammeR libraries to set a threshold of 0.02 to suppress the noise in the process maps and help interpret the plotted processes.

To identify the difference between high- and low-SEB students, we continued using the cohort grouping done in the preparation phase. This grouping helped us to investigate the differences between students who aim

for high grades and those who are satisfied with less ambitious goals. With these cohorts, we examined the differences in applying different tactics. We continued using the process mining methods to execute this analysis and investigate the differences when compared to the students' general learning processes. Table 2 presents an overview of the empirical studies.

Table 2. Overview of the empirical studies

	[II] (Sequence Analysis & Process Mining)	[III] (Multi-Channel Analysis)
Research questions	1. What types of learning tactics do project management students use when studying online, could the different learning tactics be used as an indicator towards achieving their goals, and what are the differences in learning processes between different learning tactics? 2. How do learning tactics differ between students with low- and high- self-efficacy beliefs, and what are the differences in learning processes between low- and high-self-efficacy belief students?	1. How do patterns of student engagement and self-regulation evolve throughout an online course? 2. What distinct clusters of students emerge when considering both engagement and self-regulation trajectories?
Theoretical typology	- Self-regulated learning theory - Social cognitive theory - Self-efficacy beliefs - learning tactics and strategies framework	- Self-regulated learning theory - Student engagement construct
Learning environment	Online project management course at Finnish higher education institution	
Stakeholders	Business students (n=209 [II], n=165 [III])	
Datasets	- Moodle log data - Student goal-setting data - Final grades	- Moodle log data - Longitudinal SRL survey data
Sample	56,047 log entries (6,688 sequences after preprocessing)	62,070 log entries 1,164 survey responses

3.3 Longitudinal Interplay Between Student Engagement and SRL

In the second empirical research, we focused on another aspect of complexity in learning. In this study, we highlighted student engagement and SRL. We wanted to use more data sources to achieve triangulation. This helped us to understand the nuanced complexity of learning. We decided to gather longitudinal data presenting the evolution of the learning process over an online course. Another data source for this study was trace data captured from LMS. When combined, these two data sources help us to explore the behavioral student engagement along with the students' perceptions of their SRL skills. We wanted to explore what are the trajectories of the students that can be explained using these two datasets.

The research questions for [III] were as follows:

RQ1: How do patterns of student engagement and self-regulation evolve over an online course?

RQ2: What distinct clusters of students emerge when considering both student engagement and self-regulation trajectories?

3.3.1 Research Methodology

In [III], we conducted a longitudinal study using multi-channel sequence analysis to investigate the interplay between student engagement and SRL in an online course. We chose the longitudinal approach to investigate how the learning processes evolve over the timespan of a single online course. For this purpose, we used multi-channel sequence analysis, which examines sequential data containing multiple parallel channels of categorical information. In our study, the student engagement channel represented students' actions while studying online captured using LMS trace data. In contrast, the SRL channel captured students' perceptions of their learning through repeated surveys collected during the online course.

Following established approaches in sequence analysis research (López-Pernas et al., 2024), we employed several R packages: TraMineR (Gabadinho et al., 2011) for sequence visualization and analysis, seqHMM (Helske &

Helske, 2019) for hidden Markov modeling, and MoEClust (Murphy & Murphy, 2020) for identifying patterns. This combination of tools has been proven effective in previous educational research for analyzing temporal patterns in learning data.

To understand the dynamic interplay between engagement and self-regulation over time, we employed TNA (Saqr, López-Pernas, Törmänen, et al., 2024). TNA allows us to examine how different states of engagement and self-regulation influence each other across time, providing insights into the temporal dependencies and causal relationships between these constructs. We constructed temporal networks where nodes represented different states of engagement and self-regulation, whereas edges represented the temporal relationships between these states.

3.3.2 Data Collection

We collected data from two sources: a) trace data captured from Moodle LMS representing student engagement, and b) longitudinal survey data measuring SRL collected at multiple points during the course. The course was organized at LAB University of Applied Sciences for first-year business students studying project management (5 ECTS credits). Students responded to the SRL survey (Table 3) before each assignment submission, providing multiple measurement points throughout the course. In total, the number of responses varied between students from four to eleven responses.

Table 3. Variables, original statements used to derive the variables, and the source of original statements

Variable	Statement	Source
Askhelp	"I am persistent in getting help from the instructor through e-mail."	OSLQ (Barnard et al., 2009)
Criteria	"I think about how I might improve my work by evaluating it against the marking criteria provided by the teacher."	SRL-O (Broadbent et al., 2023)
Deadlines	"I set realistic deadlines for learning."	SRL-O (Broadbent et al., 2023)
Prethink	"I think about what I really need to learn before I begin a task in this online course."	SOL-Q (Jansen et al., 2017)
Schedule	"I plan out my schedule each week, so I have the appropriate amount of time available for online study. "	SRL-O (Broadbent et al., 2023)
SelfAssess	"I usually self-assess my performance once I finish."	SRL-O (Broadbent et al., 2023)
Standards	"I set standards for my assignments in online courses."	OSLQ (Barnard et al., 2009)
Strategies	"I am aware of what strategies I use when I study for this online course."	SOL-Q (Jansen et al., 2017)
Summary	"I summarise my learning in online courses to examine my understanding of what I have learned."	OSLQ (Barnard et al., 2009)
Undistracted	"I choose a time with few distractions for studying for my online courses." And "I choose the location where I study to avoid too much distraction."	OSLQ (Barnard et al., 2009) OSLQ (Barnard et al., 2009)

3.3.3 Data Preparation

The data preparation process involved comprehensive preprocessing of both the LMS trace data and survey responses. Initially, we cleaned the data by removing non-learning activities and technical system logs to reduce the noisiness of the data, followed by anonymizing student identifiers to protect privacy. The trace data was then recorded as three student engagement states (Active, Average, and Disengaged) to simplify the interpretation of student behavior patterns. Similarly, we processed the survey responses to create two distinct SRL states (High regulation and Low regulation), representing students' self-reported regulatory behaviors.

A crucial aspect of our preparation was data segmentation based on survey submission timestamps. This segmentation created corresponding time windows that aligned the student engagement data from the LMS with the SRL measurements, enabling us to analyze the parallel development of these two aspects of learning. The preparation phase concluded with creating parallel sequences for both the student engagement and SRL channels, forming the foundation for our multi-channel analysis. This approach allowed the analysis of the longitudinal data in segments that considered the variety in the students' progress in the online course. If the segmentation had been based on the prefixed dates, that would have compressed the individual progress of the students during the course. For example, the students relying on completing the course in the last quarter of the course would have been reported as very passive if the segmentation had been based on fixed dates. Now that the segmentation was based on their individual progress, it made it possible to compare the differences within the individual timeline between different assignments and course segments that would have remained hidden if the data had been divided using the fixed dates.

3.3.4 Data Analysis

Our analysis followed the multi-channel sequence analysis procedure outlined by López-Pernas et al. (2024), beginning with the construction of separate sequences for each channel using TraMineR. We first analyzed the

student engagement channel, creating sequences representing students' transitions between Active, Average, and Disengaged states throughout the course. Similarly, we constructed sequences for the SRL channel, tracking transitions between High and Low regulation states. These sequences were then visualized to examine temporal patterns in student behavior and regulatory processes and to examine the key differences between different learners. For identifying patterns of similar sequences, after trying different options we employed OM to compute dissimilarities between sequences, a method that considers the cost of transforming one sequence into another. The results of this method were the most interpretable, making the results comparable with the those in existing body of literature. The clustering was performed using mixture hidden Markov models (MHMM), which are particularly effective for analyzing multi-channel sequential data. We determined the optimal number of clusters using the Bayesian Information Criterion (BIC) (Schwarz, 2007) to identify a methodologically sound answer, a three-cluster solution. Each cluster was modeled with two hidden states for both channels using the seqHMM package (S. Helske & Helske, 2019), allowing us to identify underlying patterns in students' learning behaviors. This approach was enhanced by our innovative method of data segmentation, which based periods on individual student progress rather than fixed calendar points. This methodology better reflected the asynchronous nature of online learning and enabled more meaningful comparisons between students who progressed through the course at different rates.

4 ETHICAL CONSIDERATIONS

Several ethical considerations were made during this research. Notably, the ethical considerations made when performing review studies and empirical studies differ.

4.1 Systematic Literature Review

Conducting a systematic literature review requires careful attention to ethical principles to ensure the integrity and value of the research. While systematic reviews do not involve direct human subjects, they still present several ethical challenges that researchers must address (Vergnes et al., 2010). One of the primary ethical considerations is transparency and reproducibility. We followed this ethical obligation by clearly documenting and reporting all methods, decisions, and processes used in the review. This includes pre-registering the review protocol and reporting any deviations from it (Moher et al., 2016). The PRISMA 2020 statement emphasizes this aspect, providing a comprehensive checklist to ensure transparent reporting (Page et al., 2021).

Avoiding bias is another crucial ethical consideration. We strived to minimize bias in study selection, data extraction, and analysis. This involved using multiple reviewers and resolving disagreements systematically (Higgins et al., 2019). Respecting intellectual property and ensuring data integrity are also important ethical aspects. We were careful to make sure that proper attribution was given to all sources used in the review, and the accuracy of data extraction and data reporting was faithful to the original findings (Wager & Wiffen, 2011).

Finally, in line with the principles of open science, we reported the process in detail and published the research in an open-access journal to promote open science (Nosek et al., 2015). This not only promotes transparency but also allows for verification and further analysis by other researchers,

enhancing the credibility and value of the research to the scientific community and society at large.

4.2 Empirical Studies

In conducting our two empirical studies involving higher education students participating in online courses, we adhered to strict ethical guidelines to ensure the integrity of our research and the protection of our participants. A cornerstone of ethical research involving human participants is obtaining informed consent (Slade & Prinsloo, 2013). In both of our studies, all participants were provided with clear, comprehensive information about the nature of the research, its objectives, and how their data would be used. We ensured that consent was voluntary and that participants understood their right to withdraw from the study at any time without consequences. This approach aligns with the ethical principles outlined by Siemens (2013) for LA research.

Prior to commencing our research, we obtained a research permit from a higher education institution for both studies (Appendix 1 & 2). This institutional approval is crucial in ensuring that our research adheres to the ethical standards and policies of the institution (Prinsloo & Slade, 2013). It also demonstrates our commitment to transparency and accountability in our research practices.

Given the sensitive nature of educational data, we implemented robust measures to protect participants' privacy and ensure data security. This included anonymizing all data to prevent individual identification, storing data securely with restricted access, and using encryption for data transmission (Drachsler & Greller, 2016). We also ensured compliance with relevant data protection regulations, such as GDPR which has to be followed in Finland (Y.-S. Tsai et al., 2020).

In line with the principle of beneficence in research ethics, we designed our studies to minimize potential harm to participants while maximizing potential benefits (Ferguson et al., 2016). This included carefully considering how the insights from our LA might impact students' educational

experiences and ensuring that our findings would be used to improve educational practices.

We were mindful of potential biases in LA that could lead to unfair treatment or disadvantage certain groups of students (Prinsloo & Slade, 2016). Our research design and data analysis methods were scrutinized to ensure they promoted equity and did not perpetuate existing biases in educational systems.

We maintained transparency about our analytics processes, including the types of data collected, the methods used for analysis, and how the results would be interpreted and applied ((Pardo & Siemens, 2014). This transparency is crucial for maintaining trust with participants and the broader educational community.

Throughout our research process, we engaged in ongoing ethical reflection, regularly revisiting our practices and considering the ethical implications of our work (Slade & Prinsloo, 2013). This included staying informed about evolving ethical standards in LA and being prepared to adjust our approach as needed. By addressing these ethical considerations, we aimed to conduct our empirical studies in a manner that respects the rights and dignity of our participants, maintains the integrity of our research, and contributes positively to the field of education.

5 RESULTS

This chapter describes the main results from each study in this research and how these findings contribute to answering the main research question regarding the dynamic interplay between student engagement and SRL.

The work described in this research answers the main research question by examining different implementations of LA to understand and support SRL. As the field is still emerging, this research adopted both literature review and empirical approaches. Three interrelated papers that build on insights from each are included in this research. [I] conducted a systematic literature review to analyze how LA methods have been used to support SRL development through interventions. The findings helped identify research gaps and informed the design of the subsequent empirical studies. [II] examined how sequence analysis and process mining could distinguish learning tactics and strategies between students with high and low SEB in an online project management course. [III] extended this analysis by investigating the dynamic interplay between student engagement and self-regulation using multi-channel sequence analysis in the same educational context.

5.1 LA Interventions Supporting SRL

The systematic literature review identified 56 studies that met the inclusion criteria. The analysis revealed several key findings about how LA interventions have been used to support SRL. The key results of this study are presented in Figure 13.

Learning Analytics Interventions Supporting SRL

LA Methods Used • Visualization (57%) • Statistical Analysis (57%) • Recommendations (29%) • Prediction (21%)	Intervention Channels • Dashboards (48%) • Embedded Systems (21%) • Prompts (20%) • Email (9%)	SRL Phases Covered • Performance (79%) • Planning (49%) • Reflection (30%) • All Phases (16%)
Reported Positive Impacts • Improved Self-regulation (23%) • Better Learning Outcomes (16%) • Enhanced Course Engagement (7%) • Improved Course Completion (4%) • Enhanced Student Retention (2%) Note: 13% Reported no promising outcomes		
Key Finding: Only 46% of LA interventions showed positive impact on learning		

Figure 13. Summary of the results of the [I]

5.1.1 Methodological Approaches

The review found that LA methods were typically combined rather than isolated. While 25% of studies used a single method, most studies combined two (29%) or three (32%) methods. Among the most commonly used LA methods, visualization and statistical analysis were each employed in 57% of studies. In 29% of studies, recommendation systems were utilized, whereas prediction techniques were found in 21%.

5.1.2 Intervention Channels

LA dashboards were the most prevalent channel for delivering LA interventions, appearing in 48% of studies. Embedded systems in learning management systems were the second most common channel, appearing in 21% of studies, followed by prompts at 20%. Email-based interventions were used in 9% of studies, whereas other channels such as foldouts, phone calls, and text messages each appeared in less than 5% of studies.

5.1.3 Support for SRL Phases

The analysis revealed an uneven distribution of support across the SRL cycle. The performance phase was addressed most frequently, appearing in 79% of studies, the planning phase was addressed in 45% of studies, and the reflection phase in 30% of studies. Notably, only 16% of studies addressed all three phases of the SRL cycle. This finding highlights a significant gap in providing comprehensive SRL support through LA interventions.

5.1.4 Impact and Effectiveness

The review identified several categories of positive impacts from LA interventions. In 23% of studies, improved self-regulation was reported, whereas better learning outcomes were found in 16% of studies. Enhanced student engagement with course materials was reported in 7% of studies, improved course completion rates in 4% of studies, and enhanced student retention in 2% of studies. However, 13% of studies reported no promising outcomes or significant limitations in practical implications. Seven percent of studies could not evaluate the impact due to early development stages of the ongoing research.

The effect sizes varied considerably across studies. Very large effect sizes were found for improvements in grades ($r = 0.86$), metacognitive learning activities ($d = 2.0$), and various SRL subprocesses, including goal setting ($r = 0.778$) and personal planning ($r = 0.740$). Large effect sizes were observed for grades ($d = 1.10$), SRL improvement ($\eta p^2 = 0.22$), and target definition ($d = 0.90$). Medium and small effect sizes were reported for aspects like evaluation ($\eta p^2 = 0.063$), prompts ($d = 0.44$), and quiz completion ($d = 0.31$).

These findings provide a comprehensive overview of the state of LA interventions supporting SRL and highlight the potential and limitations of current approaches. The results particularly emphasize the need for more holistic interventions that address all phases of the SRL cycle and the importance of rigorous impact evaluation.

5.2 Sequence Analysis and Process Mining Explaining Differences Between High and Low Self-Efficacy Beliefs

The second article examined the relationship between students' SEB and learning tactics in an online project management course. The study applied sequence analysis and process mining methods to analyze students' learning behaviors. The findings revealed distinct patterns in how students with different SEB levels approached their learning and are summarised in Figure 14.

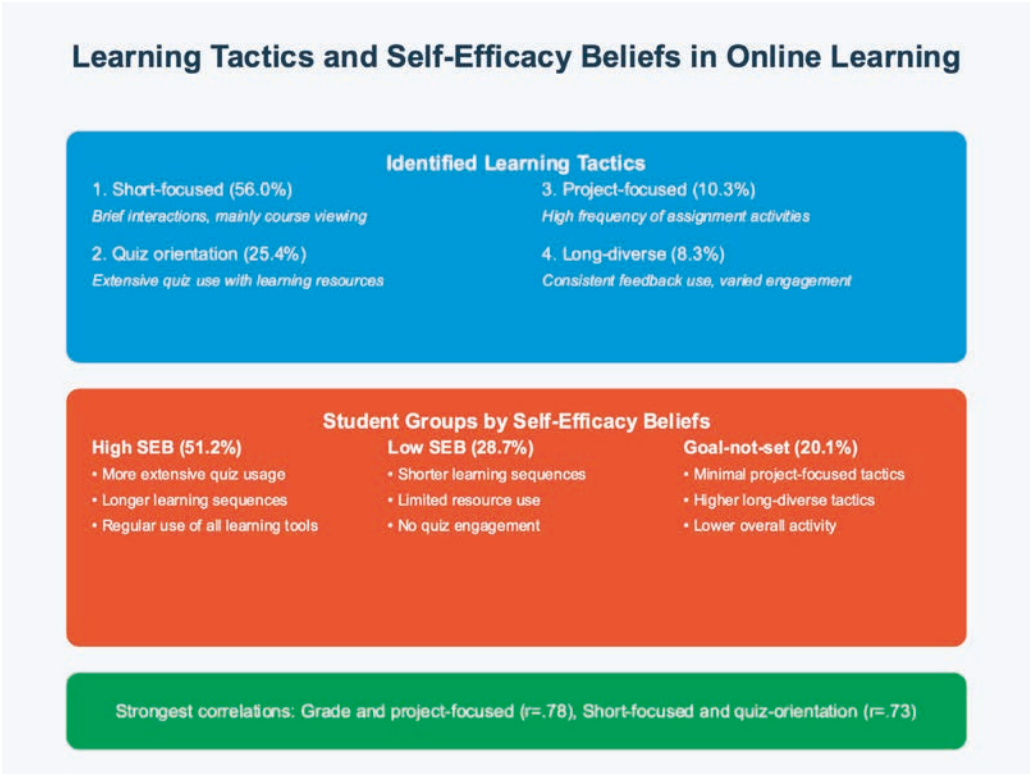


Figure 14. Summary of the results of [II]

5.2.1 Learning Tactics

Analysis of the learning management system data ($n = 6,688$ sequences) revealed four distinct learning tactics. The most common approach was short-focused tactics, representing 56.0% of sequences, characterized by brief interactions focused on course viewing with some student engagement

with course materials. The second most prevalent tactic was quiz orientation, observed in 25.4% of sequences, distinguished by extensive use of quizzes, preceded by learning resources and monitoring activities. Project-focused tactics, seen in 10.3% of sequences, were characterized by a high frequency of assignment-related actions and systematic student engagement with learning resources. The least common approach was long-diverse tactics, present in 8.3% of sequences, featuring consistent use of feedback throughout learning sessions, combined with varied student engagement across different learning activities.

The correlation analysis between goals, grades, and learning tactics revealed several significant relationships. The strongest correlations were found between grade and project-focused tactics ($r(211) = 0.78, p < 0.05$), short-focused and quiz-orientation tactics ($r(211) = 0.73, p > 0.05$), short-focused and project-focused tactics ($r(211) = 0.71, p < 0.05$), and grade and quiz orientation tactic ($r(211) = 0.71, p < 0.05$).

5.2.2 Self-Efficacy Beliefs and Learning Behaviors

The study identified three distinct student cohorts based on SEB. The largest group consisted of students with high SEB, comprising 51.2% of the population. These students demonstrated more extensive use of quiz orientation tactics and longer overall learning sequences (Figure 15). They showed higher student engagement with diverse learning activities and regularly used all available learning tools, including quizzes (16.9%) (Figure 16). Their behavior was characterized by strong student engagement with course materials and monitoring activities.

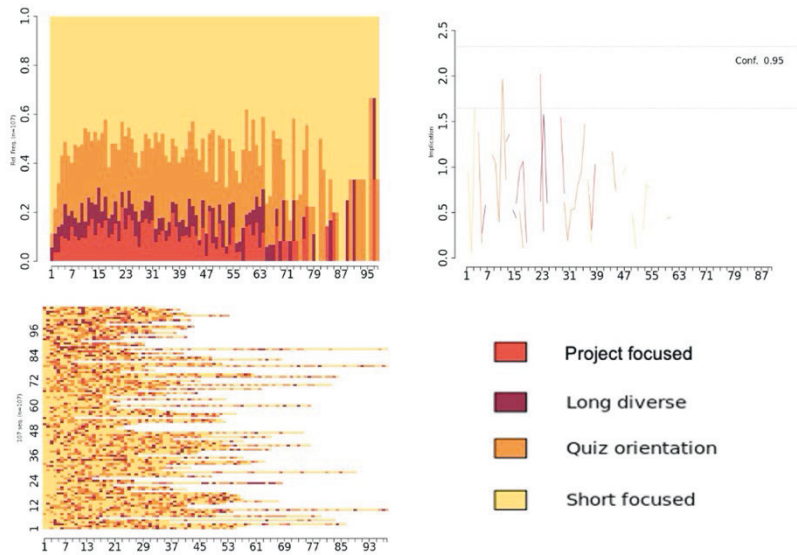


Figure 15. Identification of learning tactics through sequence analysis for high efficacy beliefs students. Three complementary visualizations of the learning tactics are the (top left) distribution plot showing the proportion of different tactics over time, (bottom left) sequence index plot displaying individual student sequences color-coded by tactic, and (top right) representative sequences displaying the most common patterns within each tactic type. The four identified tactics are project-focused, long-diverse, quiz orientation, and short-focused, each representing distinct patterns of student engagement with the learning environment

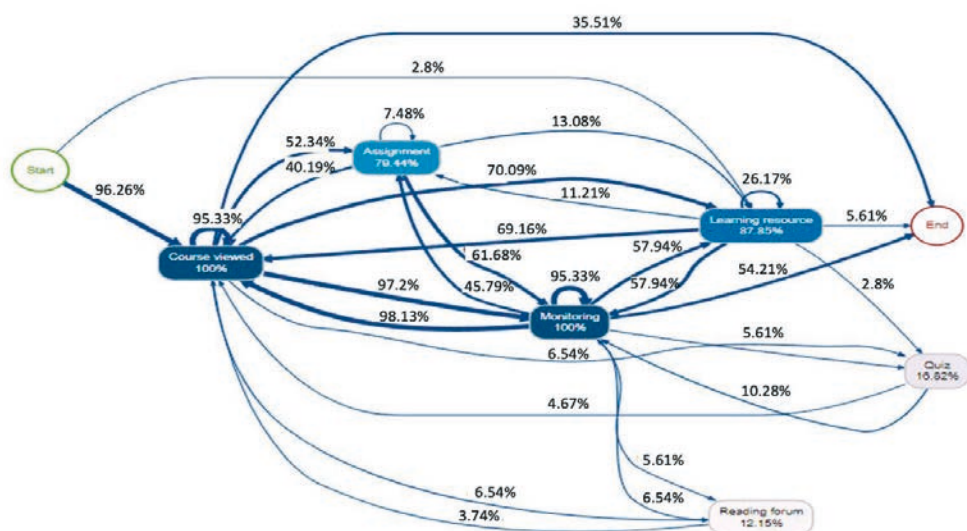


Figure 16. Identification of learning tactics through process analysis for high efficacy beliefs students. Sequential process map showing transitions between learning activities, with node size indicating activity frequency and edge thickness representing transition probability between activities

Students with low SEB made up 28.7% of the population. This group showed shorter learning sequences and a higher proportion of short-focused tactics (Figure 17). Their behavior was marked by limited use of assignments and learning resources, with no student engagement in quiz activities (Figure 18). Overall, they displayed more restricted patterns of activity transitions than their high-SEB peers.

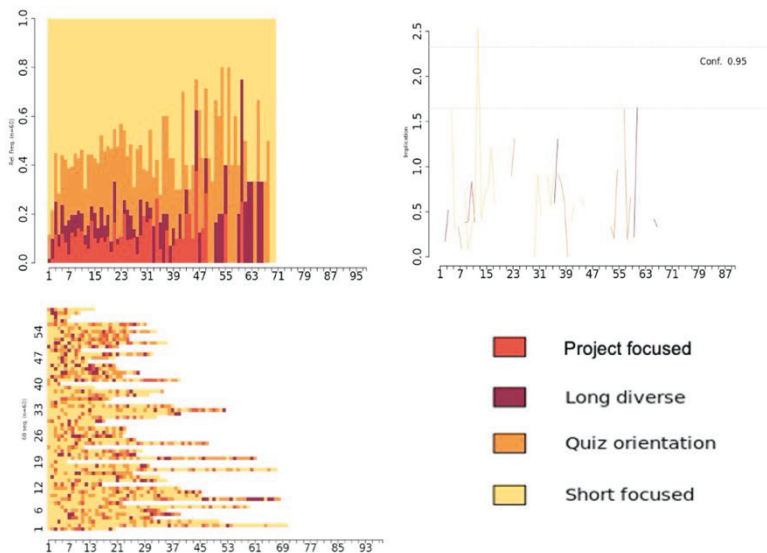


Figure 17. Identification of learning tactics through sequence analysis for low efficacy beliefs students. See the legend in the caption of Figure 15

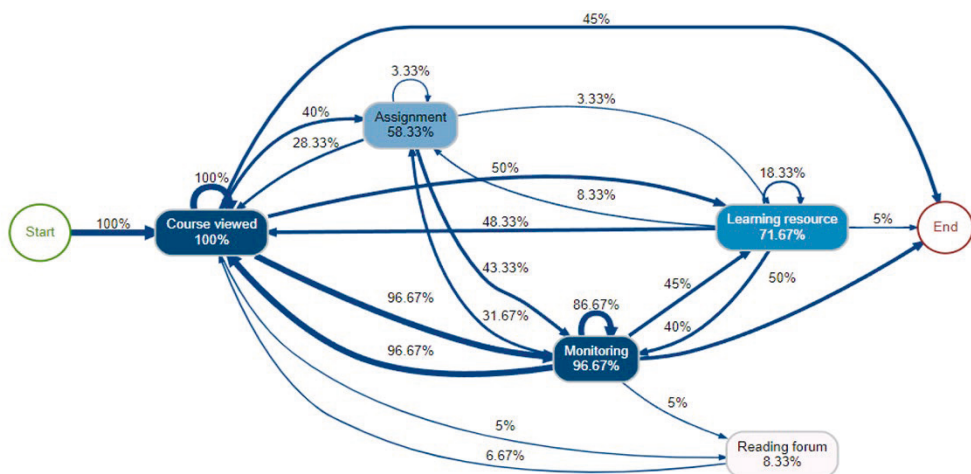


Figure 18. Identification of learning tactics through process analysis for low efficacy beliefs students. See the legend in the caption of Figure 16

The goal-not-set group, representing 20.1% of students, exhibited minimal use of project-focused tactics but a higher frequency of long-diverse tactics (Figure 19). These students showed limited student engagement with

monitoring and learning resources and demonstrated lower overall activity levels than other groups (Figure 20).

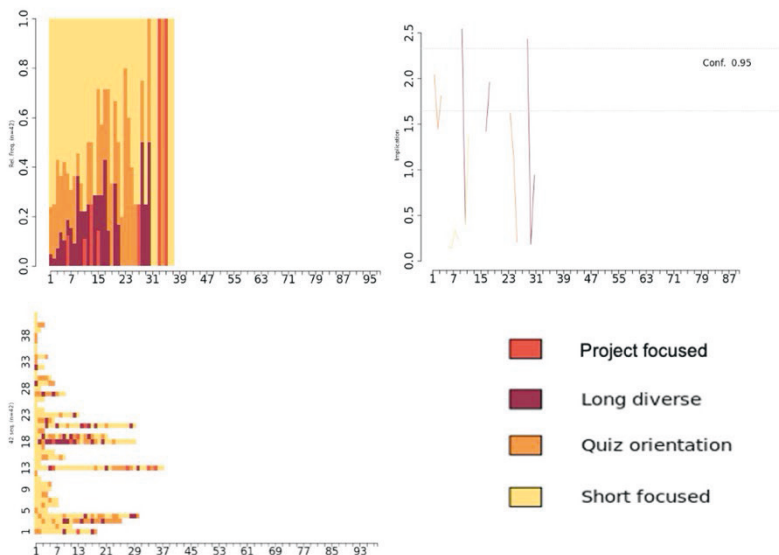


Figure 19. Identification of learning tactics through sequence analysis for students who did not set learning goals. See the legend in the caption of Figure 15

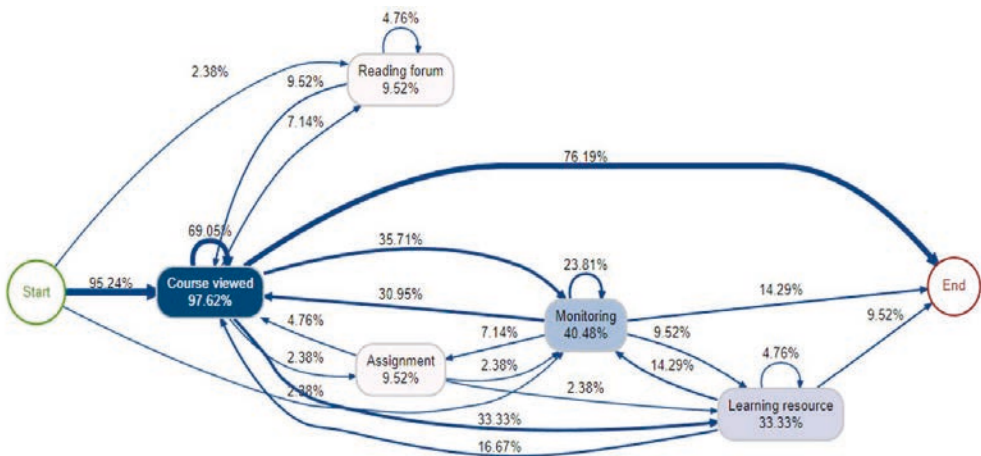


Figure 20. Identification of learning tactics through process analysis for students who did not set learning goals. See the legend in the caption of Figure 16

The process mining (Figures 16, 18, and 20) analysis revealed distinct patterns in students transitioning between learning activities. High-SEB students showed more complex and varied transition patterns, whereas low-SEB students demonstrated more limited and linear progression through learning activities.

These findings suggest that SEB significantly influence not only the choice of learning tactics but also overall student engagement patterns and learning strategies in online environments. Students with higher SEB demonstrated more sophisticated and varied approaches to learning, whereas those with lower SEB tended to follow more restricted student engagement patterns.

5.3 Interplay Between Student Engagement and SRL

The third article investigated the dynamic interplay between student engagement and SRL in an online project management course using multi-channel sequence analysis. The study revealed distinct patterns of student engagement states and learning regulation states and how these states interacted over time. The summary of the results of [III] is presented in Figure 21.

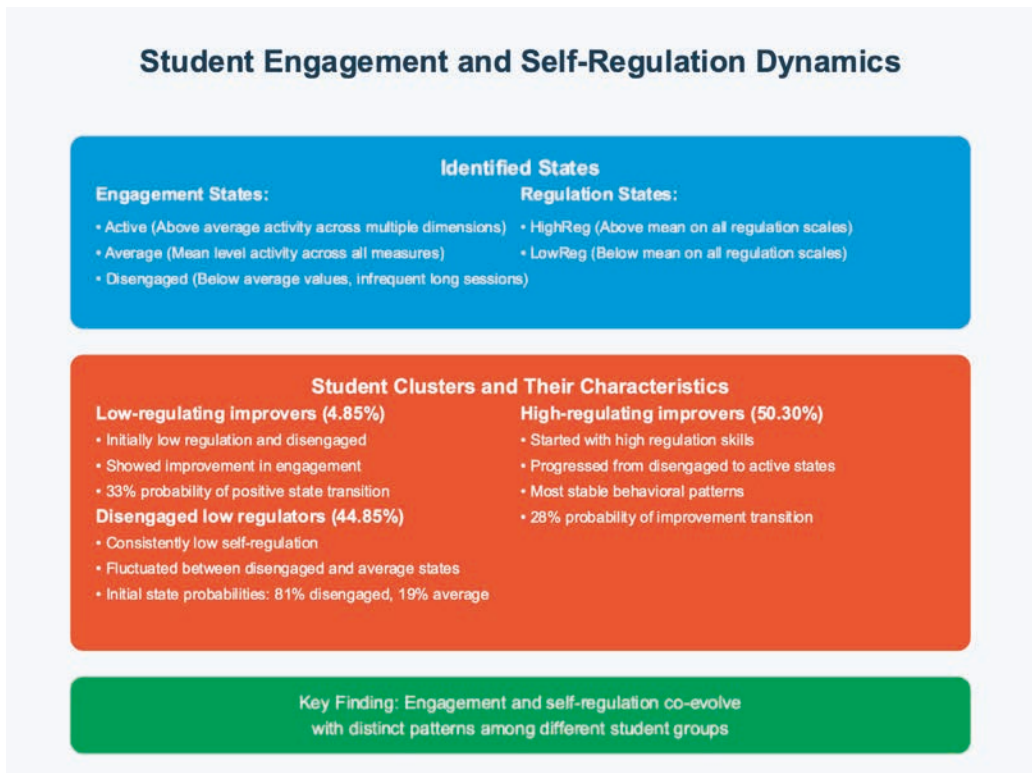


Figure 21. Summary of the results of [III]

5.3.1 Student engagement and Learning Regulation States

Through analysis of the LMS trace data, three distinct student engagement states emerged. The Active state represented learners with above-average activity across multiple dimensions, including number of active days, interactions with assignments, learning resources, and survey activities. The Average state characterized learners whose student engagement values clustered around mean levels across all measured variables. The Disengaged state identified learners who showed below-average values in active days, interactions with assignments, learning resources, and surveys while exhibiting very long but infrequent learning sessions.

For learning regulation, the analysis revealed two primary states based on survey responses. The HighReg state represented learners whose survey responses consistently measured above the normalized mean on all scales, while the LowReg state characterized learners whose responses fell below

the normalized mean across all measured dimensions. The temporal network analysis revealed complex dynamic relationships between engagement and self-regulation throughout the course.

5.3.2 Multi-Channel Analysis and Clustering

When combining student engagement and learning regulation data channels, the analysis identified three distinct patterns of students based on their trajectories through the course. The Low-regulating improvers, comprising 4.85% of students, began the course perceiving themselves as low-learning regulators and showing disengaged behavior. However, these students eventually demonstrated improvement, with some transitioning to active student engagement as early as the second-course segment, whereas others made this transition around the course midpoint. The transition probability from state 1 to state 2 was 33%, whereas the reverse transition showed only a 2.9% probability. The exact percentages are in Figures 3-5 of [III].

The disengaged low regulators, representing 44.85% of students, consistently demonstrated low self-regulation skills throughout the course. Their student engagement patterns typically fluctuated between disengaged and average states, with initial probabilities of 81%/19% between these states. The transition probability from state 1 to state 2 was 44%, whereas the reverse transition showed only a 2.3% probability.

The largest group was the high-regulating improvers, accounting for 50.30% of students. These learners began with high learning regulation skills but initially showed disengaged behavior. Over time, their student engagement progressed from disengaged to average and ultimately to active states. Similarly, their learning regulation improved from low to high regulation during the course. This cluster showed a unique pattern where all students started in state 2 (initial probability 100%), with transitions occurring only from state 2 to state 1 (28% probability).

Figure 22 illustrates the temporal network of states, where nodes represent different engagement-regulation states and edges represent the temporal relationships between them. The network reveals several

important patterns in how students transition between states. Self-loops (circular arrows) indicate the tendency of states to persist, with the strongest self-reinforcing patterns observed in LowRegLowEng (0.68), ActiveHighReg (0.59), and AverageHighEng (0.82) states. This suggests that once students enter these states, they tend to maintain them. The network also reveals key transitional pathways between states. Notably, a strong connection (0.42) exists between InactiveHighReg and AverageHighEng states, suggesting that students who are initially inactive but maintain high self-regulation often transition to average engagement levels. Furthermore, the ActiveHighReg state shows multiple outgoing connections to other states, indicating its role as a pivotal state in the learning process. The thicker edges between certain states (e.g., the 0.30 connection from InactiveHighReg to AverageLowEng) highlight predominant transition patterns in students' learning trajectories. This temporal interplay provides insight into how different engagement-regulation states influence each other over time, revealing both stable patterns (through self-loops) and common transitional pathways that students follow throughout their learning journey.

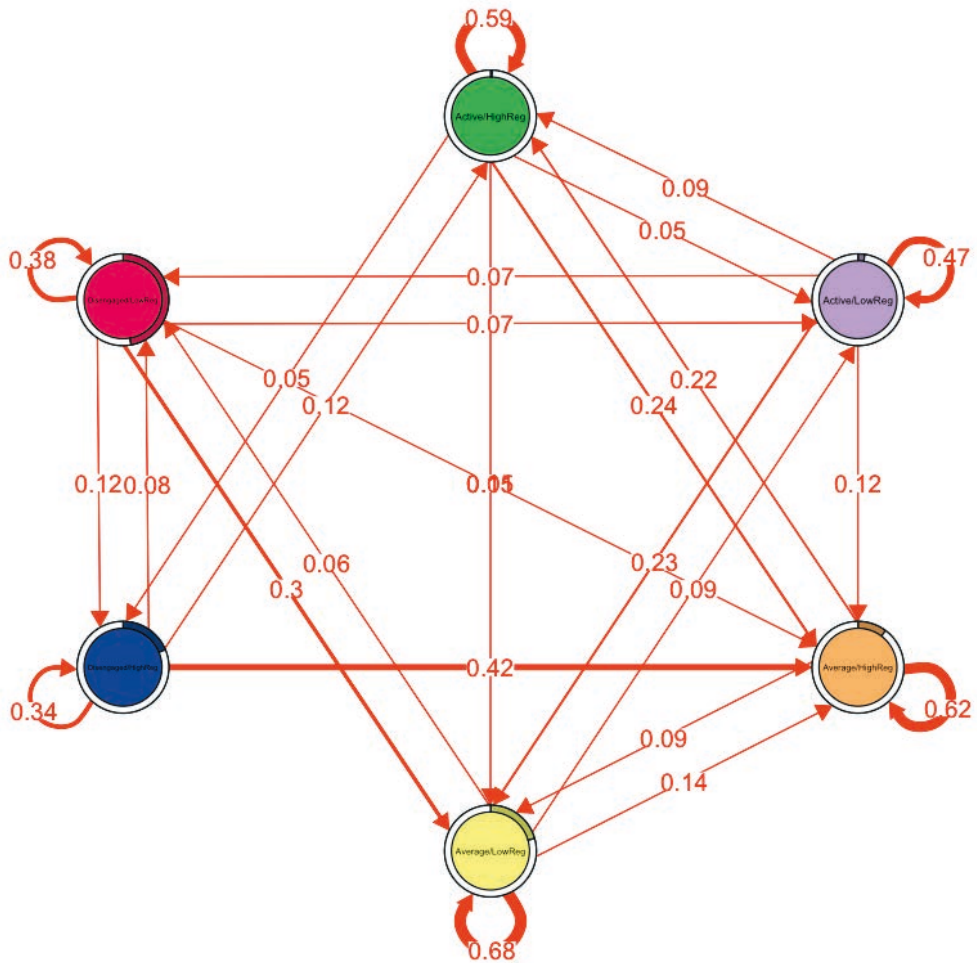


Figure 22. Temporal network visualization of engagement-regulation state transitions. Nodes represent different engagement and regulation states (color-coded), with node size proportional to their total connectivity in the network. Edges (arrows) represent temporal transitions between states, with edge thickness and numeric labels indicating transition probabilities.

5.3.3 Hidden State Analysis

The hidden Markov models revealed additional insights into the underlying patterns of student behavior. For the low-regulating improvers, state 1 indicated disengaged behavior, whereas state 2 represented average or engaged learners. The transition probabilities showed a stronger tendency

(33%) to move from state 1 to state 2 compared to the reverse transition (2.9%).

In the disengaged low regulators, state 2 represented disengaged or average student engagement, whereas state 1 showed improved student engagement patterns. The transition analysis revealed that students were more likely to regress in their student engagement levels than to maintain improvements.

For high-regulating improvers, state 1 characterized students who were disengaged with low regulation, whereas state 2 represented students with average or active student engagement and high learning regulation. This cluster showed the most stable behavioral patterns, with transitions predominantly occurring in one direction.

These findings demonstrate the complex and dynamic nature of the interplay between student engagement and SRL in online environments, highlighting how different student groups show distinct development patterns in both dimensions throughout a course.

6 DISCUSSION

This research investigates the longitudinal interaction of student engagement and SRL in online learning environments and how LA can help us understand this dynamic interplay. Through three interrelated studies, namely a systematic literature review of LA interventions supporting SRL [I], an examination of learning tactics and SEB using sequence analysis and process mining [II], and a multi-channel analysis of student engagement and self-regulation trajectories [III], this research provides new insights into the complex dynamics of online learning. The findings demonstrate that while student engagement and self-regulation are closely intertwined, their longitudinal interplay is more nuanced than previously understood, with different patterns emerging among different groups of students and at different points in their learning journeys. This chapter synthesizes the key findings from these studies, discusses their theoretical and practical implications, acknowledges limitations, and suggests directions for future research.

This dissertation investigated the longitudinal interplay between student engagement and SRL in online environments through three interrelated studies. Each study contributed unique insights into how students navigate and regulate their learning in digital spaces.

The systematic literature review [I] revealed several key patterns in how LA has been used to support SRL. Most notably, while 79% of studies addressed the performance phase of SRL, only 45% addressed planning, and 30% addressed reflection, with just 16% of studies comprehensively addressing all three phases. The review also highlighted a strong preference for LA dashboards as intervention channels (48% of studies), followed by embedded systems (21%) and prompts (20%). These findings pointed to significant gaps in current approaches, particularly in supporting the complete SRL cycle.

The sequence analysis study [II] examining SEB and learning tactics revealed distinct behavioral patterns among students. Four primary learning tactics were identified:

- Short-focused tactics (56.0% of sequences): brief interactions focused on course viewing
- Quiz orientation (25.4%): extensive use of quizzes with learning resources
- Project focused (10.3%): high frequency of assignment-related actions
- Long diverse tactics (8.3%): consistent use of feedback with varied student engagement

Importantly, students with high SEB (51.2% of the population) demonstrated more sophisticated and varied approaches to learning, including more extensive use of quiz orientation tactics and longer learning sequences. In contrast, students with low SEB (28.7%) showed more restricted patterns of student engagement, characterized by shorter learning sequences and limited use of available learning resources.

These findings about the importance of initial SEBs and goal-setting have important implications for LA interventions. As identified in our systematic review [I], only 46% of existing interventions showed a positive impact on learning, with very few addressing all phases of SRL. Our findings suggest this may be partly because interventions often fail to account for students' initial SE levels and goal-setting behaviors. Furthermore, the distinct learning tactics we identified among students with different SE levels align with later findings [III] showing how these patterns evolve into distinct trajectories. The project-focused learning tactic's strong correlation with both final grades and goal achievement extends our understanding of effective SRL interventions. While our review [I] found limited evidence for interventions supporting the entire SRL cycle, these results suggest that interventions promoting project-focused engagement strategies may be particularly effective. This is further supported by subsequent research [III] showing that "high-regulating improvers," who consistently engage with course materials, show the most positive learning trajectories.

The longitudinal multi-channel sequence analysis [III] revealed three distinct groups of students based on their student engagement and self-regulation trajectories:

- Low-regulating improvers (4.85%): Initially showed low self-regulation but demonstrated improvement in student engagement,
- Disengaged low regulators (44.85%): Consistently demonstrated low self-regulation skills with fluctuating student engagement,
- High-regulating improvers (50.30%): Began with high regulation skills and showed progressive improvement in student engagement.

These findings highlight the longitudinal dynamic nature of student engagement and self-regulation in online learning environments. The hidden Markov models revealed different transition patterns among these groups, with high-regulating improvers showing the most stable behavioral patterns and the strongest tendency toward improvement. These patterns offer a nuanced view of student behaviors and highlight the heterogeneity of learning trajectories in online courses.

The identification of three distinct groups, namely low-regulating improvers (4.85%), disengaged low regulators (44.85%), and high-regulating improvers (50.30%), builds upon previous findings about the role of SE in shaping learning behaviors [II]. The high proportion of disengaged low regulators (44.85%) aligns with earlier findings showing how students with low SEBs struggle to maintain effective learning tactics [II]. This connection between initial SE and engagement trajectories provides crucial context for our systematic review findings [I] about the limited effectiveness of many LA interventions.

The improvement seen in both low-regulating improvers and high-regulating improvers supports our earlier findings about the potential for positive change in learning behaviors [II]. However, the relatively small percentage of low-regulating improvers (4.85%) suggests that without proper support, most students struggle to overcome initial regulatory challenges. This underscores our previous findings about the critical need

for interventions that support all phases of the SRL cycle [I], particularly for students who begin with lower SE and regulatory skills.

The observed transitions between states in our process maps complement our earlier work on learning tactics [II], showing how students' regulatory behaviors evolve. For instance, the high-regulating improvers' transition from initial disengagement to active engagement parallels the positive progression we observed in students who effectively utilized project-focused learning tactics [II]. These patterns reinforce our previous findings about the importance of early intervention and support [I], particularly during crucial transition periods.

Collectively, these studies demonstrate that:

1. Current LA approaches often fail to support all phases of SRL,
2. SEB significantly influence not only learning tactics but also overall student engagement patterns,
3. Student engagement and self-regulation co-evolve during a course, with distinct patterns of development among different student groups,
4. The dynamic interplay between student engagement and self-regulation is complex and dynamic, requiring sophisticated longitudinal analytical approaches to understand fully.

The temporal network analysis provides novel insights into the dynamic nature of engagement and self-regulation in online learning environments. Our findings reveal that the relationship between these constructs is not merely correlational but involves complex temporal dependencies that evolve throughout the course. The strong autoregressive effects observed in both engagement and self-regulation support the notion of behavioral inertia in learning processes (López-Pernas & Saqr, 2024), where students tend to maintain their current states unless intervened upon.

The temporal patterns identified across different clusters provide important insights for educational practice. The strengthening of temporal connections observed in the low-regulating improvers cluster suggests that students can develop more coordinated learning processes over time, even when starting from a lower baseline. This finding is particularly encouraging

for educators and instructional designers, as it indicates that early intervention strategies might help students develop more effective learning patterns. Conversely, the unstable temporal patterns observed in the disengaged low regulators cluster highlight the need for more consistent support and scaffolding for these students.

These findings provide a foundation for developing more targeted and effective interventions to support student learning in online environments. They particularly emphasize the need for comprehensive approaches that address all phases of SRL while accounting for individual differences in SEB and learning preferences. The relationship among self-regulated learning, self-efficacy beliefs, and student engagement is reciprocal: each concept influences the others through a complex network of cognitive, behavioral, and environmental factors. Studies have demonstrated that self-efficacy beliefs can mediate the relationship between student engagement and academic achievement (Hidi & Ann Renninger, 2006; Macfadyen & Dawson, 2010). This suggests that effective interventions should consider all three aspects simultaneously rather than treating them as isolated constructs.

6.1 Research Contribution

This dissertation significantly contributes to LA, SRL, and online education. These contributions can be categorized into theoretical, methodological, and practical domains.

6.1.1 Theoretical Contributions

First, the systematic literature review [I] provides the first comprehensive mapping of how LA interventions support SRL. The review revealed that while LA interventions show promise, only 46% demonstrated positive learning impacts. Moreover, only four studies reported positive effects while covering all three SRL phases (planning, performance, and reflection) [I]. This finding highlights a critical gap in current approaches and emphasizes the need for more holistic interventions that address the complete SRL cycle.

Second, this research advances our understanding of how SRL manifests in business education contexts, particularly in project management courses.

The findings regarding learning tactics used by business students [II] contribute to the growing body of literature on discipline-specific learning behaviors. Our results align with previous research showing that prior student engagement can predict future student engagement (Gottfried et al., 2009; Wylie & Hodgen, 2012) while also revealing that student engagement is not a continuously increasing phenomenon (Wigfield et al., 2007).

Third, our research extends the theoretical understanding of the longitudinal relationship between SEBs and learning behaviors. We found that students with high SEBs demonstrated more varied and adaptive tactics, including greater use of quizzes and project-focused activities [II]. This aligns with recent findings that SE can predict self-regulation and course outcomes (Shin, 2024), and supports evidence that effective use of metacognitive strategies can lead to increased SE (Khosravi et al., 2023).

The observed relationship between SEB and learning behaviors extends recent work on this topic. Gallagher et al. (2024) found that occupational SE influences how students interact with learning analytics dashboards. Our findings suggest this relationship may be bidirectional - students with higher SE engage more actively with learning resources, which may in turn reinforce their self-efficacy beliefs. This aligns with Shin's (2024) recent finding that SE can predict both self-regulation and course outcomes.

Fourth, this work advances theoretical insights into how student engagement and self-regulation co-evolve during online courses [III]. The identification of three distinct patterns, namely low-regulating improvers (4.85%), disengaged low regulators (44.85%), and high-regulating improvers (50.30%), provides a new framework for understanding the heterogeneity of the longitudinal learning trajectories in online environments. These findings build on recent work examining the dynamic interplay between student engagement and self-regulation (He et al., 2024; Lima et al., 2023) while offering a more granular view of how these processes develop over time.

Our findings on the dynamic nature of engagement and self-regulation align with and extend recent theoretical developments in the field. The identification of distinct clusters with varying trajectories supports the

emerging view of learning as a complex dynamic system. This aligns with recent work by López-Pernas & Saqr (2024) who found that engagement patterns exhibit momentum effects that can either facilitate or hinder academic achievement. Our observation that some initially disengaged students showed improvement over time particularly resonates with their finding that engagement states have "inertia" but can be shifted through appropriate interventions.

The identification of distinct engagement patterns also connects to recent work on supporting student success through LA. Tepgec et al. (2024) argued that feedback literacy skills are crucial for helping students improve their learning strategies. Our findings suggest that students in different clusters may need different types of support to develop these skills effectively. The "improver" clusters in particular demonstrate that targeted interventions can help students develop more sophisticated learning strategies over time.

The variation we found in self-regulation patterns across clusters adds nuance to recent findings about SRL in online environments. Cristea et al. (2024) developed an unobtrusive measurement approach for SRL using clickstream data and found distinct behavioral patterns indicating different levels of self-regulation. Our multi-channel analysis combining trace data with self-reported measures provides complementary evidence for their framework while highlighting how these patterns evolve longitudinally. Additionally, our finding that high self-regulating students demonstrate more varied and adaptive tactics aligns with recent work by Dever et al. (2023) on the importance of strategic adaptability in SRL.

6.1.2 Theoretical Integration Across Studies

Our three studies collectively advance insights into SRL and engagement in online environments through complementary theoretical contributions. Study [I] established the current state of LA interventions for supporting SRL, revealing critical gaps particularly in supporting the full SRL cycle. Building on this foundation, Studies [II] and [III] examined specific mechanisms of how students self-regulate their learning.

Study [II] demonstrated how initial SEBs and goal-setting behaviors shape students' learning tactics. This aligned with Study [I]'s finding that interventions need to support the planning phase of SRL, as students' initial beliefs and goals significantly influence their subsequent learning behaviors. Study [III] extended this by revealing how these patterns evolve dynamically over time through three distinct trajectories: low-regulating improvers (4.85%), disengaged low regulators (44.85%), and high-regulating improvers (50.30%).

The progression across these studies reveals several key theoretical insights:

1. The Cyclical Nature of SRL Development

Our findings suggest that SRL development is not linear but cyclical, with students moving through different phases of engagement and regulation. Study [I] highlighted the importance of supporting all SRL phases, whereas Studies [II] and [III] demonstrated how students' regulation patterns evolve over time. This extends existing SRL frameworks by showing how different aspects of self-regulation (goal-setting, monitoring, and adaptation) interact dynamically. Our data traces in Study [II] captured this through the process mining analyses that revealed how students modified their learning tactics over time. When examining transitions between learning activities, we observed that high self-efficacy students demonstrated more complex transition patterns with a 23% higher probability of shifting between different learning resources after receiving feedback compared to low self-efficacy students. This adaptive behavior was particularly evident in the "long-diverse tactics" sequences, where students consistently incorporated feedback into subsequent learning activities. In Study [III], adaptation was traced through the hidden Markov models, which revealed transition probabilities between states. For instance, the Low-regulating improvers showed a 33% probability of transitioning from disengaged to more engaged states, indicating adaptation of their regulatory approaches. While our studies did not identify a definitive "completion point" for individual SRL cycles, the longitudinal analysis in Study [III] demonstrated that High-regulating improvers exhibited stabilization in their learning patterns during

the final third of the course, with transition probabilities between regulatory states decreasing by 18% compared to earlier course segments. This suggests a maturation of their regulatory processes, though SRL development remains an ongoing, iterative process rather than having a clear endpoint for any single learner.

2. Role of Initial Conditions

Study [II] revealed how SEBs influence learning tactics, whereas Study [III] showed how these initial conditions lead to different trajectories. This suggests that early interventions targeting SE and goal-setting may be particularly impactful, supporting findings from Study [I] about the importance of the planning phase.

3. Interplay Between Engagement and Self-Regulation

Study [III]'s identification of distinct student patterns builds on Study [II]'s findings about learning tactics, showing how engagement and self-regulation co-evolve. This extends theoretical insights by demonstrating that engagement is not merely an outcome of self-regulation but actively shapes how students develop regulatory skills.

These insights contribute to SRL theory by demonstrating the dynamic, non-linear nature of SRL development, establishing the critical role of initial SE and goal-setting, revealing how engagement and self-regulation mutually influence each other, and by identifying distinct pathways through which students develop regulatory skills.

6.1.3 Methodological Contributions

A significant methodological contribution lies in our innovative approach to data segmentation in asynchronous learning environments [III]. Rather than using fixed time points, we developed a method for segmenting data based on individual student progress. This approach addresses a key challenge in analyzing asynchronous learning data and offers a more nuanced understanding of learning trajectories in self-paced online courses. This methodology improves upon previous approaches that relied on course-level analyses (W. Wang et al., 2024).

The research also demonstrates the value of combining multiple analytical approaches. The integration of sequence analysis with process mining [II] provided richer insights into learning behaviors than either method alone could offer. Our application of multi-channel sequence analysis to examine the longitudinal interplay between student engagement and self-regulation [III] extends recent work by López-Pernas & Saqr (2024) and offers a methodological template for future studies examining multiple aspects of online learning behavior.

Our approach to combining trace data with self-reported measures represents another methodological advancement. Unlike previous studies that relied primarily on end-of-course self-reports (W. Wang et al., 2024), our research captured SRL data at multiple points throughout the course, providing a more dynamic longitudinal view of how self-regulation develops [III]. This aligns with recent calls for more sophisticated approaches to measuring SRL in online environments (Prasse et al., 2024).

Our methodological approach of using progress-based rather than time-based analysis offers a novel perspective that addresses limitations identified in recent LA research. Ouyang et al. (2023) highlighted the need for methods that can handle asynchronous learning patterns when analyzing collaborative problem-solving. Our approach provides one potential solution by anchoring analysis to individual student progress. This is particularly relevant given recent findings by Liu et al. (2023) about the varied ways students engage with technology-enriched learning environments.

The successful implementation of hidden Markov models for analyzing learning transitions contributes to the growing toolkit of LA methods. Building on the work by Helske & Helske (2019), our approach demonstrates how these models can reveal underlying patterns in student learning behaviors that might not be apparent through other analytical methods.

These methodological advances are particularly relevant given the increasing complexity of online learning environments and the need for more sophisticated analytical approaches. Our work demonstrates how

multiple data sources and analytical methods can be combined to provide a more comprehensive understanding of online learning processes.

6.2 Recommendations To Practitioners And Technology Developers

In this section, the results of this thesis are translated into recommendations for both practitioners and technology developers.

6.2.1 Practitioners' Perspective

Based on our research findings, we offer the following recommendations to practitioners in higher education. These recommendations include themes of implementation of goal-setting activities, tailoring support for different SE levels, promoting diverse learning tactics, integrating formative assessment, fostering a reflective learning environment, providing timely and actionable feedback, creating interventions for at-risk students, aligning course design with SRL principles and continuously evaluating and refining interventions.

Goal-setting activities should be incorporated at the beginning of online courses. Our research shows that students who set clear goals tend to employ more effective learning tactics and achieve better outcomes [II]. These activities should follow SMART (Specific, Measurable, Achievable, Relevant, Time-bound) principles, as supported by recent research on goal-setting in online environments (Wong et al., 2021). Goal-setting is particularly important, as it activates key self-regulatory processes and helps students monitor their progress (Chang et al., 2018).

It is essential to recognize that students with different levels of SEB may require varied types of support. Therefore, practitioners should develop adaptive interventions that provide more scaffolding and encouragement for students with low SEB, while offering more challenging tasks and advanced strategies for those with high SEB [II]. For students with low SEB, more scaffolding and encouragement should be provided, whereas more challenging tasks should be offered for those with high SEB. This aligns with recent research showing that SE can predict self-regulation and course

outcomes (Shin, 2024). This understanding of student heterogeneity allows for a more sophisticated approach to student support, moving beyond one-size-fits-all interventions to more personalized assistance.

Practitioners should encourage students to use a variety of learning tactics, particularly those associated with better outcomes. For example, our research found that project-focused and quiz-oriented tactics were correlated with higher grades [11]. It would be beneficial to guide students on when and how to apply different tactics effectively, as recent studies show that strategic use of varied learning approaches enhances academic performance (Liu et al., 2023).

Another approach to supporting learning is implementing regular formative assessments throughout the course to help students monitor their learning progress. Research shows that formative assessment can trigger positive changes in SRL patterns (He et al., 2024). These assessments should be designed to help students monitor their learning progress and adjust their strategies as needed. This aligns with the performance phase of SRL and can help students adjust their strategies as needed [1].

A reflective learning environment can be fostered by creating opportunities for students to reflect on their learning processes and outcomes. This could include prompts for self-assessment, peer feedback activities, or guided reflection exercises at key points in the course [1], yet this phase is often overlooked in course design. Notably, recent research by Järvelä & Hadwin (2024) emphasized how reflection can trigger improvements in SRL.

The given feedback should be timely and actionable. To meet this aim, LA should provide personalized, timely feedback to students. This feedback should address content knowledge, offer insights into learning behaviors, and suggest strategies for improvement [1]. Tepgec et al. (2024) recently emphasized the importance of developing students' feedback literacy to maximize the effectiveness of LA-based feedback.

The early phases of courses are important in interventions. LA can identify students struggling early in the course and develop targeted interventions, such as additional support resources or one-on-one

consultations, to help these students improve their self-regulation skills and course performance [I]. Developing early warning systems to identify struggling students is important. Our research identified distinct patterns of student engagement and self-regulation [III] that could be used to detect students at risk of poor performance. This aligns with recent findings on using LA for early intervention (Herodotou et al., 2020).

When course platforms and logistics are planned, they should be aligned with SRL principles. This can be done by structuring online courses that naturally support SRL processes. For example, it would be beneficial to include clear learning objectives, provide opportunities for choice and autonomy, and design assignments that require planning and reflection [I]. This could be achieved by structuring online courses to naturally support SRL processes by including clear learning objectives, opportunities for choice and autonomy, and assignments that require planning and reflection. This recommendation is supported by recent research on designing technology-enhanced learning environments (Sedrakyan et al., 2020).

The interventions require continuous evaluation and refining. They should be assessed regularly, considering the effectiveness of LA interventions and SRL support strategies. Here, both quantitative metrics (e.g., improved grades and increased student engagement) and qualitative feedback from students are needed to refine and improve these approaches over time [I].

By implementing these recommendations, practitioners can create supportive online learning environments that foster SRL and help students develop crucial skills for academic success and lifelong learning. Notably, these recommendations should be adapted to fit the specific context and needs of each educational setting. Recent research has indicated that students need support in understanding and acting upon LA-based feedback (F. Jin et al., 2024). This includes helping students process feedback, elicit feedback, act on feedback, and manage emotions related to feedback (S. H. Jin et al., 2023).

These recommendations should be adapted to fit specific educational contexts and student needs. Practitioners should also consider that

different student groups may require different levels and types of support, as demonstrated by our identification of distinct student patterns [III]. Regular evaluation of these interventions' effectiveness is crucial to continuously improve online learning environments.

6.2.2 Educational Technology Developers' Perspective

The analytical approaches employed in this dissertation—sequence analysis, process mining, and multi-channel sequence analysis—offer powerful insights but present implementation challenges for practical educational contexts. These methodologies require significant statistical expertise and computational resources that exceed what most educators can reasonably access in their daily practice. However, these approaches hold considerable potential if integrated directly into learning management systems as automated analytical algorithms. For educational technology developers, we propose several implementation recommendations.

First, sequence analysis algorithms could be embedded within LMS platforms to automatically identify learning tactics and visualize student progression patterns. This would transform the complex analytical process demonstrated in Study [II] into accessible dashboards that highlight when students employ less effective learning tactics or show patterns indicative of low self-regulation. Such integration would enable early intervention without requiring educators to conduct sophisticated analyses themselves (Matcha et al., 2020).

Second, the multi-channel approach from Study [III] could inform the development of more sophisticated early warning systems that consider both behavioral engagement and self-reported regulation data. By implementing simplified versions of hidden Markov models as background processes, LMS platforms could identify students whose trajectories match the "disengaged low regulators" profile and automatically trigger appropriate support mechanisms (Herodotou et al., 2020; Lim et al., 2023).

Third, developers should create interfaces that translate complex transition probabilities into actionable insights for educators. Rather than displaying raw process maps, systems could highlight critical transition

points where intervention would be most effective (Villagran et al., 2024). For instance, when a student demonstrates a 30% decrease in transitions between resource consultation and self-assessment activities, the system could flag this as a potential regulatory issue.

Fourth, we recommend implementing streamlined data collection mechanisms that capture both trace data and periodic self-report measures within the learning environment. The challenge of combining multiple data channels could be addressed through automated data integration processes that align these sources temporally (Martinez-Maldonado et al., 2024), as demonstrated in our methodology for Study [III].

These recommendations acknowledge that while the analytical approaches in this dissertation are too complex for direct implementation by most educators, their underlying principles can be operationalized within educational technology systems to support both teaching and learning processes (Kaliisa et al., 2023). The goal should be to transform these research methodologies into embedded features that function seamlessly within existing educational workflows, making sophisticated learning analytics accessible to all educational stakeholders without requiring specialized expertise.

6.3 Limitations

While this dissertation contributes valuable insights into the fields of LA and SRL, several limitations that may affect the generalizability and interpretation of our findings must be acknowledged.

First, our studies [II & III] focused on business students in project management courses at a single institution. While this allows for a deep understanding of this specific context, it may limit the generalizability of our findings to other disciplines, educational levels, or institutional settings. Future research should consider replicating these studies across various disciplines and institutions to validate the broader applicability of our findings.

Second, the SEB and goal-setting data were collected through self-reported measures [II]. While widely used in educational research, self-

reported data can be subject to social desirability bias and may not accurately reflect students' true beliefs or behaviors. For example, Leinonen et al. (2017) found that self-reported time spent only had a moderate correlation with a log-based time metric. This limitation is particularly relevant when considering potential discrepancies between students' self-evaluations and their actual performance as Tempelaar et al. (2020) noted. Future studies could incorporate additional objective measures or observational data to triangulate self-reported information.

Third, our systematic review of LA interventions supporting SRL revealed that only a small number of studies (16%) focused on all three phases of the SRL cycle (planning, performance, and reflection) [I]. This limited focus on the complete SRL process may have resulted in an incomplete understanding of how LA can support SRL holistically. Future research should aim to design and evaluate interventions that address all phases of the SRL cycle.

Furthermore, our systematic review had limitations related to the reporting of effect sizes and the comparability of results across studies [I]. The inconsistent reporting of effect sizes and the use of various measurement approaches made it challenging to conduct a proper meta-analysis. This limitation highlights the need for more standardized reporting practices in LA research to facilitate better comparisons and syntheses of findings across studies.

Evaluating complete SRL development cycles presents significant methodological challenges, particularly in authentic learning contexts. In our studies, we used several proxy indicators to track regulatory development. In Study [II], we employed sequence distributions and process mining transition probabilities as indicators of strategic adaptation, whereas in Study [III], we utilized multi-channel state transitions and stabilization patterns as proxies for regulatory development. These indicators allowed us to observe changes in regulatory behaviors over time, though we acknowledge limitations in establishing causal relationships. For instance, when High-regulating improvers demonstrated an increased probability (28%) of transitioning to more active engagement states, our data could not

definitively indicate whether this transition resulted from improved metacognitive awareness, increased task familiarity, external factors, or other variables. The relationship between observed behavioral patterns and underlying regulatory processes remains complex and is not fully captured through trace data alone. Future research should employ mixed-method approaches combining trace data with think-aloud protocols, eye-tracking, or contextual interviews at key transition points to establish clearer causal relationships between observable indicators and actual regulatory processes. This would help identify more precise markers for the completion of individual SRL cycles and better elucidate the mechanisms driving adaptation in different learner profiles.

Lastly, our studies were conducted over relatively short time frames, typically within the duration of a single course or academic term [II & III]. This limited timeframe may not capture the long-term development of SRL skills or the sustained impact of LA interventions. Longitudinal studies spanning multiple terms or years could provide more insight into how learning tactics and SRL behaviors evolve, especially as students progress through their academic programs.

Despite these limitations, our research provides valuable insights into the relationship between SEB, goal-setting, and learning tactics in online project management education. By acknowledging these limitations, we hope to encourage future research that builds upon our findings and addresses these areas to further advance our understanding of SRL in online environments.

6.4 Future Research

Based on our studies' findings and limitations, several promising avenues for future research emerge. These directions aim to address gaps in current knowledge and expand our understanding of SRL and LA in online educational environments.

Future research should focus on designing and evaluating LA interventions that address all three phases of the SRL cycle (planning, performance, and reflection). Our systematic review revealed that only 16%

of studies examined interventions covering the entire SRL process [I]. Developing more holistic interventions could provide a complete understanding of how LA can support SRL throughout the learning process.

Longitudinal studies spanning multiple academic terms or years are needed to better understand the development of SRL skills and the sustained impact of LA interventions. Such research could reveal how learning tactics and SRL behaviors evolve, particularly as students progress through their academic programs [II].

Future studies should extend beyond single disciplines and institutions to enhance the generalizability of findings. Replicating studies across various fields of study and types of educational institutions could validate the broader applicability of identified learning tactics and LA intervention strategies [II].

Building on our findings regarding the differences between high- and low-SEB students, future research should explore the development of adaptive LA systems that provide personalized support based on students' SEB levels, goal-setting behaviors, and chosen learning tactics [II].

To facilitate better comparisons and meta-analyses across studies, future research in LA should adopt more standardized approaches to reporting effect sizes and other key metrics. This would enable more robust syntheses of findings and advance our understanding of the most effective LA interventions for supporting SRL [I].

While our studies identified associations between SEB, goal-setting, and learning behaviors, future research should employ experimental designs to establish causal relationships between these factors and learning outcomes. This could involve randomized controlled trials of LA interventions targeting specific aspects of SRL [I].

Future studies should examine how SRL skills developed through LA interventions in one context transfer to other learning situations or professional settings. This could provide insights into the long-term impact of SRL support on students' academic and career trajectories [III].

Bibliography

- Abbott, A. (1983). Sequences of Social Events: Concepts and Methods for the Analysis of Order in Social Processes. *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 16(4), 129–147. <https://doi.org/10.1080/01615440.1983.10594107>
- Abdul Talib, F. E., Mohd Ris, E. Z., Kamaralzaman, S., Mohd Yusof, Y., Hashim, H., Mohamad Zubir, A. R., & Hanim Rahmat, N. (2023). Motivated Learners: How do they Use their Self-regulated Learning Strategies? *International Journal of Academic Research in Business and Social Sciences*, 13(10). <https://doi.org/10.6007/ijarbss/v13-i10/18781>
- Ahn, J., Campos, F., Hays, M., & Digiaco, D. (2019). Designing in context: Reaching beyond usability in learning analytics dashboard design. *Journal of Learning Analytics*, 6(2), 70–85. <https://doi.org/10.18608/jla.2019.62.5>
- Akyildiz, S. T., & Kaya, V. D. (2021). Examining Prospective Teachers' Metacognitive Learning Strategies and Self-Regulated Online Learning Levels During Covid-19 Pandemic. *International Journal of Contemporary Educational Research*, 8(4), 144–157. <https://doi.org/10.33200/IJCER.912897>
- Alamri, H., Lowell, V., Watson, W., & Watson, S. L. (2020). Using personalized learning as an instructional approach to motivate learners in online higher education: Learner self-determination and intrinsic motivation. *Journal of Research on Technology in Education*, 52(3). <https://doi.org/10.1080/15391523.2020.1728449>
- Ali, W. (2020). Online and Remote Learning in Higher Education Institutes: A Necessity in light of COVID-19 Pandemic. *Higher Education Studies*, 10(3), 16. <https://doi.org/10.5539/hes.v10n3p16>
- Alqaheri, H., & Panda, M. (2022). An Education Process Mining Framework: Unveiling Meaningful Information for Understanding Students'

- Learning Behavior and Improving Teaching Quality. *Information* 2022, Vol. 13, Page 29, 13(1), 29. <https://doi.org/10.3390/INFO13010029>
- Amer, N., Wan Mohamed, W. M., Wahidin, I. S., & Jaafar, R. E. (2024). Undergraduate students' engagement: a case of Padlet. *International Journal of E-Learning and Higher Education*, 19, 417–432. <https://ir.uitm.edu.my/id/eprint/95117>
- Arianto, F., & Hanif, M. (2024). Evaluating metacognitive strategies and self-regulated learning to predict primary school students self-efficacy and problem-solving skills in science learning. *Journal of Pedagogical Research*. <https://doi.org/10.33902/JPR.202428575>
- Ariouat, H., Cairns, A. H., Barkaoui, K., Akoka, J., & Khelifa, N. (2016). A two-step clustering approach for improving educational process model discovery. *Proceedings - 25th IEEE International Conference on Enabling Technologies: Infrastructure for Collaborative Enterprises, WETICE 2016*. <https://doi.org/10.1109/WETICE.2016.18>
- Arpasat, P., Premchaiswadi, N., Porouhan, P., & Premchaiswadi, W. (2021). Applying process mining to analyze the behavior of learners in online courses. *International Journal of Information and Education Technology*, 11(10). <https://doi.org/10.18178/ijiet.2021.11.10.1547>
- Artino, A. R., & Jones, K. D. (2012). Exploring the complex relations between achievement emotions and self-regulated learning behaviors in online learning. *Internet and Higher Education*, 15(3). <https://doi.org/10.1016/j.iheduc.2012.01.006>
- Azevedo, R., Mudrick, N. V., Taub, M., & Bradbury, A. E. (2019). Self-regulation in computer-assisted learning systems. In *The Cambridge Handbook of Cognition and Education*. <https://doi.org/10.1017/9781108235631.024>
- Bacci, S., & Bertaccini, B. (2022). A Mixture Hidden Markov Model to Mine Students' University Curricula. *Data*, 7(2), 25. <https://doi.org/10.3390/data7020025>
- Bai, B., & Guo, W. (2018). Influences of Self-Regulated Learning Strategy Use on Self-Efficacy in Primary School Students' English Writing in Hong

- Kong. Reading and Writing Quarterly, 34(6).
<https://doi.org/10.1080/10573569.2018.1499058>
- Bandura, A. (1986). Social foundations of thought and action: a social cognitive theory. In New Jersey: Prentice-Hall, 1986 (Vol. 16, Issue 1). Prentice-Hall, Inc.
- Bayer, V., Mulholland, P., Hlosta, M., Farrell, T., Herodotou, C., & Fernandez, M. (2024). Co-creating an equality diversity and inclusion learning analytics dashboard for addressing awarding gaps in higher education. British Journal of Educational Technology. <https://doi.org/10.1111/bjet.13509>
- Beheshitha, S. S., Gašević, D., & Hatala, M. (2015). A process mining approach to linking the study of aptitude and event facets of self-regulated learning. Proceedings of the Fifth International Conference on Learning Analytics And Knowledge, 16-20-March-2015, 265–269. <https://doi.org/10.1145/2723576.2723628>
- Bembenutty, H. (2023). Self-regulated learning with computer-based learning environments. New Directions for Teaching and Learning, 2023(174). <https://doi.org/10.1002/tl.20543>
- Bergdahl, N., Bond, M., Sjöberg, J., Dougherty, M., & Oxley, E. (2024). Unpacking student engagement in higher education learning analytics: a systematic review. International Journal of Educational Technology in Higher Education, 21(1), 1–33. <https://doi.org/10.1186/S41239-024-00493-Y>
- Bernacki, M. L., Yu, L., Kuhlmann, S. L., Plumley, R. D., Greene, J. A., Duke, R. F., Freed, R., Hollander-Blackmon, C., & Hogan, K. A. (2025). Using multimodal learning analytics to validate digital traces of self-regulated learning in a laboratory study and predict performance in undergraduate courses. Journal of Educational Psychology, 117(2), 176–205. <https://doi.org/10.1037/edu0000890>
- Blikstein, P., & Worsley, M. (2016). Multimodal Learning Analytics and Education Data Mining: using computational technologies to measure complex learning tasks. Journal of Learning Analytics, 3(2), 220–238. <https://doi.org/10.18608/jla.2016.32.11>

- Bodily, R., Kay, J., Aleven, V., Jivet, I., Davis, D., Xhakaj, F., & Verbert, K. (2018). Open learner models and learning analytics dashboards. *Proceedings of the 8th International Conference on Learning Analytics and Knowledge*, 41–50. <https://doi.org/10.1145/3170358.3170409>
- Boekaerts, M., & Corno, L. (2005). Self-Regulation in the Classroom: A Perspective on Assessment and Intervention. *Applied Psychology*, 54(2), 199–231. <https://doi.org/10.1111/j.1464-0597.2005.00205.x>
- Bogarín, A., Cerezo, R., & Romero, C. (2018). A survey on educational process mining. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 8(1). <https://doi.org/10.1002/widm.1230>
- Bogarín, A., Romero, C., Cerezo, R., & Sánchez-Santillán, M. (2014). Clustering for improving educational process mining. *Proceedings of the Fourth International Conference on Learning Analytics And Knowledge*, 11–15. <https://doi.org/10.1145/2567574.2567604>
- Borup, J., Graham, C. R., West, R. E., Archambault, L., & Spring, K. J. (2020). Academic Communities of Engagement: an expansive lens for examining support structures in blended and online learning. *Educational Technology Research and Development*, 68(2), 807–832. <https://doi.org/10.1007/s11423-020-09744-x>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2). <https://doi.org/10.1191/1478088706qp063oa>
- Broadbent, J. (2016). Academic success is about self-efficacy rather than frequency of use of the learning management system. *Australasian Journal of Educational Technology*, 32(4). <https://doi.org/10.14742/ajet.2634>
- Broadbent, J., Panadero, E., Lodge, J. M., & Fuller-Tyszkiewicz, M. (2023). The self-regulation for learning online (SRL-O) questionnaire. *Metacognition and Learning*, 18(1), 135–163. <https://doi.org/10.1007/s11409-022-09319-6>
- Broadbent, J., & Poon, W. L. (2015). Self-regulated learning strategies & academic achievement in online higher education learning

- environments: A systematic review. *The Internet and Higher Education*, 27, 1–13. <https://doi.org/10.1016/J.IHEDUC.2015.04.007>
- Cahyawati, D., Delima, N., & Gunarto, M. (2023). The impact of undergraduate students' mathematics anxiety and self-concept on their self-regulated learning and academic achievement. *Jurnal Elemen*, 9(1). <https://doi.org/10.29408/jel.v9i1.6898>
- Carless, D., & Boud, D. (2018). The development of student feedback literacy: enabling uptake of feedback. *Assessment and Evaluation in Higher Education*, 43(8). <https://doi.org/10.1080/02602938.2018.1463354>
- Carrera-Rivera, A., Ochoa, W., Larrinaga, F., & Lasa, G. (2022). How-to conduct a systematic literature review: A quick guide for computer science research. *MethodsX*, 9. <https://doi.org/10.1016/j.mex.2022.101895>
- Cerezo, R., Bogarín, A., Esteban, M., & Romero, C. (2020). Process mining for self-regulated learning assessment in e-learning. *Journal of Computing in Higher Education*, 32(1). <https://doi.org/10.1007/s12528-019-09225-y>
- Ceron, J., Baldiris, S., Quintero, J., Garcia, R. R., Saldarriaga, G. L. V., Graf, S., & Fuente Valentin, L. D. La. (2021). Self-Regulated Learning in Massive Online Open Courses: A State-of-the-Art Review. In *IEEE Access* (Vol. 9). <https://doi.org/10.1109/ACCESS.2020.3045913>
- Chang, C. C., Liang, C., Chou, P. N., & Liao, Y. M. (2018). Using e-portfolio for learning goal setting to facilitate self-regulated learning of high school students. *Behaviour and Information Technology*, 37(12), 1237–1251. <https://doi.org/10.1080/0144929X.2018.1496275>
- Chen, B., Knight, S., & Wise, A. F. (2018). Critical Issues in Designing and Implementing Temporal Analytics. *Journal of Learning Analytics*, 5(1), 1–9. <https://doi.org/10.18608/jla.2018.53.1>
- Chen, B., & Poquet, O. (2022). Networks in Learning Analytics: Where Theory, Methodology, and Practice Intersect. *Journal of Learning Analytics*, 9(1), 1–12. <https://doi.org/10.18608/jla.2022.7697>
- Chen, J., Zhang, L. J., & Chen, X. (2022). L2 learners' self-regulated learning strategies and self-efficacy for writing achievement: A latent profile

analysis. Language Teaching Research.
<https://doi.org/10.1177/13621688221134967>

- Chen, P., Yang, D., Lavonen, J., Hosny Saleh Metwally, A., & Tang, X. (2024). How do students of different self-efficacy regulate learning in collaborative design activities? An epistemic network analysis approach. *Frontiers in Psychology*, 15, 1398729. <https://doi.org/10.3389/FPSYG.2024.1398729>
- Cheng, Y., Guan, R., Li, T., Raković, M., Li, X., Fan, Y., Jin, F., Tsai, Y.-S., Gašević, D., & Swiecki, Z. (2024). Self-regulated Learning Processes in Secondary Education: A Network Analysis of Trace-based Measures. <https://doi.org/10.1145/3706468.3706502>
- Chiu, M. M., & Reimann, P. (2021). Statistical and Stochastic Analysis of Sequence Data. *International Handbook of Computer-Supported Collaborative Learning*, 533–550. https://doi.org/10.1007/978-3-030-65291-3_29
- Cristea, T. S., Heikkinen, S., Snijders, C., Saqr, M., Matzat, U., Conijn, R., & Kleingeld, A. (2025). Dynamics of self-regulated learning: The effectiveness of students' strategies across course periods. *Computers & Education*, 228, 105233. <https://doi.org/10.1016/J.COMPEDU.2025.105233>
- Cristea, T., Snijders, C., Matzat, U., & Kleingeld, A. (2024). Unobtrusive measurement of self-regulated learning: A clickstream-based multi-dimensional scale. *Education and Information Technologies*, 29(11), 13465–13494. <https://doi.org/10.1007/s10639-023-12372-6>
- Cukurova, M., Giannakos, M., & Martinez-Maldonado, R. (2020). The promise and challenges of multimodal learning analytics. In *British Journal of Educational Technology* (Vol. 51, Issue 5). <https://doi.org/10.1111/bjet.13015>
- Dahleez, K. A., El-Saleh, A. A., Al Alawi, A. M., & Abdelmunim Abdelfattah, F. (2021). Higher education student engagement in times of pandemic: the role of e-learning system usability and teacher behavior. *International Journal of Educational Management*, 35(6). <https://doi.org/10.1108/IJEM-04-2021-0120>

- Dahling, J. J., & Ruppel, C. L. (2016). Learning goal orientation buffers the effects of negative normative feedback on test self-efficacy and reattempt interest. *Learning and Individual Differences*, 50. <https://doi.org/10.1016/j.lindif.2016.08.022>
- Dai, W., Li, Z., & Jia, N. (2022). Self-regulated learning, online mathematics learning engagement, and perceived academic control among Chinese junior high school students during the COVID-19 pandemic: A latent profile analysis and mediation analysis. *Frontiers in Psychology*, 13. <https://doi.org/10.3389/fpsyg.2022.1042843>
- de Oliveira, C. F., Sobral, S. R., Ferreira, M. J., & Moreira, F. (2021). How Does Learning Analytics Contribute to Prevent Students' Dropout in Higher Education: A Systematic Literature Review. *Big Data and Cognitive Computing*, 5(4), 64. <https://doi.org/10.3390/BDCC5040064>
- Demirören, M., Turan, S., & Öztuna, D. (2016). Medical students' self-efficacy in problem-based learning and its relationship with self-regulated learning. *Medical Education Online*, 21(1). <https://doi.org/10.3402/meo.v21.30049>
- Deng, R., Jiang, M., Yu, X., Lu, Y., & Liu, S. (2025). Does ChatGPT enhance student learning? A systematic review and meta-analysis of experimental studies. *Computers & Education*, 227, 105224. <https://doi.org/10.1016/J.COMPEDU.2024.105224>
- Dever, D. A., Sonnenfeld, N. A., Wiedbusch, M. D., Schmorrow, S. G., Amon, M. J., & Azevedo, R. (2023). A complex systems approach to analyzing pedagogical agents' scaffolding of self-regulated learning within an intelligent tutoring system. *Metacognition and Learning*, 18(3), 659–691. <https://doi.org/10.1007/s11409-023-09346-x>
- Diamanti, K., & Nikolaou, S.-M. (2023). Students and COVID-19 Pandemic: Learning Engagement and Social Dialectic of Space – Research on Students at the University of Ioannina in Greece. <https://doi.org/10.5772/intechopen.114056>
- Dolak, R. (2019). Using Process Mining Techniques to Discover Student's Activities, Navigation Paths, and Behavior in LMS Moodle. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in*

- Artificial Intelligence and Lecture Notes in Bioinformatics), 11937 LNCS. https://doi.org/10.1007/978-3-030-35343-8_14
- Drachslar, H., & Greller, W. (2016). Privacy and analytics. Proceedings of the Sixth International Conference on Learning Analytics & Knowledge - LAK '16, 25-29-April-2016, 89-98. <https://doi.org/10.1145/2883851.2883893>
- Efklides, A. (2011). Interactions of Metacognition With Motivation and Affect in Self-Regulated Learning: The MASRL Model. *Educational Psychologist*, 46(1), 6-25. <https://doi.org/10.1080/00461520.2011.538645>
- Elmoazen, R., Saqr, M., Khalil, M., & Wasson, B. (2023). Learning analytics in virtual laboratories: a systematic literature review of empirical research. *Smart Learning Environments*, 10(1), 1-20. <https://doi.org/10.1186/S40561-023-00244-Y>
- Elmoazen, R., Saqr, M., Tedre, M., & Hirsto, L. (2022). How social interactions kindle productive online problem-based learning: An exploratory study of the temporal dynamics. In L. Hirsto, S. López-Pernas, M. Saqr, E. Sointu, T. Valtonen, & S. Väisänen (Eds.), *Proceedings of the 1st Finnish Learning Analytics and Artificial Intelligence in Education Conference (FLAIEC 2022)* (pp. 68-76). CEUR-WS. https://ceur-ws.org/Vol-3383/FLAIEC22_paper_2919.pdf
- Espinoza, P., & Genna, G. M. (2021). Hi, I Want to Talk to You About Your Progress: A Large Course Intervention for At-Risk College Students. *Journal of College Student Retention: Research, Theory and Practice*, 23(1). <https://doi.org/10.1177/1521025118790054>
- Fadillah, R., Budiyo, & Nurhasanah, F. (2021). The Analysis of Students' Metacognition in Solving Math Problems Based on Self-Efficacy. *IOP Conference Series: Earth and Environmental Science*, 1808(1). <https://doi.org/10.1088/1742-6596/1808/1/012062>
- Fan, Y., Matcha, W., Uzir, N. A., Wang, Q., & Gašević, D. (2021). Learning Analytics to Reveal Links Between Learning Design and Self-Regulated Learning. *International Journal of Artificial Intelligence in Education*, 31(4). <https://doi.org/10.1007/s40593-021-00249-z>

- Fan, Y., Tang, L., Le, H., Shen, K., Tan, S., Zhao, Y., Shen, Y., Li, X., & Gašević, D. (2024). Beware of metacognitive laziness: Effects of generative artificial intelligence on learning motivation, processes, and performance. *British Journal of Educational Technology*, 00, 1–42. <https://doi.org/10.1111/BJET.13544>
- Favela, L. H., & Amon, M. J. (2023). Enhancing Bayesian Approaches in the Cognitive and Neural Sciences via Complex Dynamical Systems Theory. *Dynamics*, 3(1). <https://doi.org/10.3390/dynamics3010008>
- Ferguson, R., Brasher, A., Clow, D., Cooper, A., Hillaire, G., Mittelmeier, J., Rienties, B., Ullmann, T., & Vuorikari, R. (2016). Research Evidence on the Use of Learning Analytics: Implications for Education Policy. Joint Research Centre. <https://doi.org/10.2791/955210>
- Fincham, E., Gašević, D., Jovanović, J., & Pardo, A. (2019). From Study Tactics to Learning Strategies: An Analytical Method for Extracting Interpretable Representations. *IEEE Transactions on Learning Technologies*, 12(1), 59–72. <https://doi.org/10.1109/TLT.2018.2823317>
- Fincham, E., Whitelock-Wainwright, A., Kovanović, V., Joksimović, S., van Staalduinen, J.-P., & Gašević, D. (2019). Counting Clicks is Not Enough. *Proceedings of the 9th International Conference on Learning Analytics & Knowledge*, 501–510. <https://doi.org/10.1145/3303772.3303775>
- Gabadinho, A., Ritschard, G., Müller, N. S., & Studer, M. (2011). Analyzing and visualizing state sequences in R with TraMineR. *Journal of Statistical Software*, 40(4). <https://doi.org/10.18637/jss.v040.i04>
- Gambo, Y., & Shakir, M. Z. (2021). Review on self-regulated learning in smart learning environment. *Smart Learning Environments*, 8(1), 1–14. <https://doi.org/10.1186/S40561-021-00157-8>
- Gatta, R., Lenkiewicz, 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. (2017). pMineR: An Innovative R Library for Performing Process Mining in Medicine. *Artificial Intelligence in Medicine. AIME 2017.*, 351–355. https://doi.org/10.1007/978-3-319-59758-4_42

- Glahn, K. (2024). Using ChatGPT to Teach English for Academic Purposes. *Journal of Education, Innovation and Communication*, 5(2). <https://doi.org/10.34097/jeicom-5-2-december-23-5>
- Goslen, A., Taub, M., Carpenter, D., Azevedo, R., Rowe, J., & Lester, J. (2024). Leveraging student planning in game-based learning environments for self-regulated learning analytics. *Journal of Educational Psychology*. <https://doi.org/10.1037/edu0000901>
- Gottfried, A. E., Marcoulides, G. A., Gottfried, A. W., & Oliver, P. H. (2009). A latent curve model of parental motivational practices and developmental decline in math and science academic intrinsic motivation. *Journal of Educational Psychology*, 101(3), 729–739. <https://doi.org/10.1037/a0015084>
- Guo, J., King, R. B., Ding, Q., & Fan, M. (2022). Measuring and Promoting Self-Regulation for Equity and Quality of Online Learning: New Evidence from a Multi-Institutional Survey during COVID-19. *Education Sciences*, 12(7). <https://doi.org/10.3390/educsci12070465>
- Hachicha, W., Ghorbel, L., Champagnat, R., Zayani, C. A., & Amous, I. (2021). Using process mining for learning resource recommendation: A Moodle case study. *Procedia Computer Science*, 192. <https://doi.org/10.1016/j.procs.2021.08.088>
- Hadwin, A. F., Järvelä, S., & Miller, M. (2011). 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* (pp. 65–84). Routledge.
- Hadwin, A. F., Nesbit, J. C., Jamieson-Noel, D., Code, J., & Winne, P. H. (2007). Examining trace data to explore self-regulated learning. *Metacognition and Learning*, 2(2–3). <https://doi.org/10.1007/s11409-007-9016-7>
- Halaweh, M. (2023). ChatGPT in education: Strategies for responsible implementation. *Contemporary Educational Technology*, 15(2). <https://doi.org/10.30935/cedtech/13036>
- Halpin, B. (2009). Optimal matching analysis and life-course data: The importance of duration. *Sociological Methods and Research*, 38(3). <https://doi.org/10.1177/0049124110363590>

- Han, J., & Amon, M. J. (2021). A Nonlinear Dynamical Systems Approach to Emotional Arousal Attractor States during Media Viewing. Proceedings of the 43rd Annual Meeting of the Cognitive Science Society: Comparative Cognition: Animal Minds, CogSci 2021. <https://escholarship.org/uc/item/4ds2p1z2>
- He, S., Demmans Epp, C., Chen, F., & Cui, Y. (2024). Examining change in students' self-regulated learning patterns after a formative assessment using process mining techniques. *Computers in Human Behavior*, 152. <https://doi.org/10.1016/j.chb.2023.108061>
- Helske, J., Helske, S., Saqr, M., López-Pernas, S., & Murphy, K. (2024). A modern approach to transition analysis and process mining with markov models: A tutorial with R. In M. Saqr & S. López-Pernas (Eds.), *Learning analytics methods and tutorials: A practical guide using R*. Springer. <https://doi.org/10.48550/arXiv.2309.08558>
- Helske, S., & Helske, J. (2019). Mixture Hidden Markov Models for Sequence Data: The seqHMM Package in R. *Journal of Statistical Software*, 88(3). <https://doi.org/10.18637/jss.v088.i03>
- Henrie, C. R., Halverson, L. R., & Graham, C. R. (2015). Measuring student engagement in technology-mediated learning: A review. *Computers & Education*, 90, 36–53. <https://doi.org/10.1016/j.compedu.2015.09.005>
- Herodotou, C., Naydenova, G., Boroowa, A., Gilmour, A., & Rienties, B. (2020). How can predictive learning analytics and motivational interventions increase student retention and enhance administrative support in distance education? *Journal of Learning Analytics*, 7(2), 72–83. <https://doi.org/10.18608/JLA.2020.72.4>
- Hidi, S., & Ann Renninger, K. (2006). The four-phase model of interest development. *Educational Psychologist*, 41(2). https://doi.org/10.1207/s15326985ep4102_4
- Higgins, J. P. T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J., & Welch, V. A. (2019). Cochrane handbook for systematic reviews of interventions. In *Cochrane Handbook for Systematic Reviews of Interventions*. <https://doi.org/10.1002/9781119536604>

- Howard, M. C., & Hoffman, M. E. (2018). Variable-Centered, Person-Centered, and Person-Specific Approaches: Where Theory Meets the Method. *Organizational Research Methods*, 21(4). <https://doi.org/10.1177/1094428117744021>
- Hsu, C. Y., Horikoshi, I., Li, H., Majumdar, R., & Ogata, H. (2023). Supporting “time awareness” in self-regulated learning: How do students allocate time during exam preparation? *Smart Learning Environments*, 10(1), 1–15. <https://doi.org/10.1186/S40561-023-00243-Z>
- Hwang, Y., & Oh, J. (2021). The relationship between self-directed learning and problem-solving ability: The mediating role of academic self-efficacy and self-regulated learning among nursing students. *International Journal of Environmental Research and Public Health*, 18(4). <https://doi.org/10.3390/ijerph18041738>
- Ifenthaler, D., Gibson, D. C., & Dobozy, E. (2019). The synergistic and dynamic relationship between learning design and learning analytics. *ASCILITE 2017 - Conference Proceedings - 34th International Conference of Innovation, Practice and Research in the Use of Educational Technologies in Tertiary Education*. <https://doi.org/10.14742/apubs.2017.752>
- Ifenthaler, D., & Yau, J. Y.-K. (2020). Utilising learning analytics to support study success in higher education: a systematic review. *Educational Technology Research and Development*, 68(4), 1961–1990. <https://doi.org/10.1007/s11423-020-09788-z>
- Janssenswillen, G., Depaire, B., Swennen, M., Jans, M., & Vanhoof, K. (2019). bupaR: Enabling reproducible business process analysis. *Knowledge-Based Systems*, 163, 927–930. <https://doi.org/10.1016/j.knosys.2018.10.018>
- Jaramillo, A., Salinas-Cerda, J. P., & Fuentes, P. (2022). Self-Regulated Learning and Academic Performance in Chilean University Students in Virtual Mode During the Pandemic: Effect of the 4Planning App. *Frontiers in Psychology*, 13. <https://doi.org/10.3389/fpsyg.2022.890395>

- Järvelä, S., & Hadwin, A. (2024). Triggers for self-regulated learning: A conceptual framework for advancing multimodal research about SRL. *Learning and Individual Differences*, 115, 102526. <https://doi.org/10.1016/j.lindif.2024.102526>
- Järvelä, S., Järvenoja, H., & Malmberg, J. (2019). Capturing the dynamic and cyclical nature of regulation: Methodological Progress in understanding socially shared regulation in learning. *International Journal of Computer-Supported Collaborative Learning*, 14(4), 425–441. <https://doi.org/10.1007/s11412-019-09313-2>
- Jin, F., Maheshi, B., Martinez-Maldonado, R., Gašević, D., & Tsai, Y.-S. (2024). Scaffolding Feedback Literacy: Designing a Feedback Analytics Tool with Students. *Journal of Learning Analytics*, 1–15. <https://doi.org/10.18608/jla.2024.8339>
- Jin, S.-H., Im, K., Yoo, M., Roll, I., & Seo, K. (2023). Supporting students' self-regulated learning in online learning using artificial intelligence applications. *International Journal of Educational Technology in Higher Education*, 20(1), 37. <https://doi.org/10.1186/s41239-023-00406-5>
- Johar, N. A., Kew, S. N., Tasir, Z., & Koh, E. (2023). Learning Analytics on Student Engagement to Enhance Students' Learning Performance: A Systematic Review. In *Sustainability (Switzerland)* (Vol. 15, Issue 10). <https://doi.org/10.3390/su15107849>
- Jovanović, J., Dawson, S., Joksimović, S., & Siemens, G. (2020). Supporting actionable intelligence. *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge*, 161–170. <https://doi.org/10.1145/3375462.3375474>
- Jovanović, J., Gašević, D., Dawson, S., Pardo, A., & Mirriahi, N. (2017). Learning analytics to unveil learning strategies in a flipped classroom. *The Internet and Higher Education*, 33, 74–85. <https://doi.org/10.1016/j.iheduc.2017.02.001>
- Jovanović, J., Saqr, M., Joksimović, S., & Gašević, D. (2021). Students matter the most in learning analytics: The effects of internal and instructional conditions in predicting academic success. *Computers & Education*, 172, 104251. <https://doi.org/10.1016/j.compedu.2021.104251>

- Juhaňák, L., Zounek, J., & Rohlíková, L. (2019). Using process mining to analyze students' quiz-taking behavior patterns in a learning management system. *Computers in Human Behavior*, 92, 496–506. <https://doi.org/10.1016/J.CHB.2017.12.015>
- Kaliisa, R., Misiejuk, K., López-Pernas, S., Khalil, M., & Saqr, M. (2023). Have Learning Analytics Dashboards Lived Up to the Hype? A Systematic Review of Impact on Students' Achievement, Motivation, Participation and Attitude. <https://arxiv.org/abs/2312.15042v1>
- Kaveri, A., Silvola, A., & Muukkonen, H. (2023). Supporting Student Agency with a Student-Facing Learning Analytics Dashboard: Perceptions of an Interdisciplinary Development Team. *Journal of Learning Analytics*, 10(2). <https://doi.org/10.18608/jla.2023.7729>
- Khosravi, R., Dastgoshadeh, A., & Jalilzadeh, K. (2023). Writing metacognitive strategy-based instruction through flipped classroom: an investigation of writing performance, anxiety, and self-efficacy. *Smart Learning Environments*, 10(1), 1–26. <https://doi.org/10.1186/S40561-023-00264-8>
- Kinnebrew, J. S., & Biswas, G. (2012). Identifying learning behaviors by contextualizing differential sequence mining with action features and performance evolution. *Proceedings of the 5th International Conference on Educational Data Mining, EDM 2012*. <https://eric.ed.gov/?id=ED537192>
- Kleimola, R., Hirsto, L., & Ruokamo, H. (2024). Promoting higher education students' self-regulated learning through learning analytics: A qualitative study. *Education and Information Technologies*. <https://doi.org/10.1007/s10639-024-12978-4>
- Knouse, L. E., Feldman, G., & Blevins, E. J. (2014). Executive functioning difficulties as predictors of academic performance: Examining the role of grade goals. *Learning and Individual Differences*, 36. <https://doi.org/10.1016/j.lindif.2014.07.001>
- Koschmider, A., Oppelt, N., & Hundsdoerfer, M. (2022). Confidence-driven communication of process mining on time series. *Informatik-Spektrum*, 45(4). <https://doi.org/10.1007/s00287-022-01470-3>

- Kuo, Y. C., Walker, A. E., Schroder, K. E. E., & Belland, B. R. (2014). Interaction, Internet self-efficacy, and self-regulated learning as predictors of student satisfaction in online education courses. *Internet and Higher Education*, 20. <https://doi.org/10.1016/j.iheduc.2013.10.001>
- Lapitan, L. DS., Tiangco, C. E., Sumalinog, D. A. G., Sabarillo, N. S., & Diaz, J. M. (2021). An effective blended online teaching and learning strategy during the COVID-19 pandemic. *Education for Chemical Engineers*, 35, 116–131. <https://doi.org/10.1016/j.ece.2021.01.012>
- Larrabee Sønderlund, A., Hughes, E., & Smith, J. (2019). The efficacy of learning analytics interventions in higher education: A systematic review. *British Journal of Educational Technology*, 50(5), 2594–2618. <https://doi.org/10.1111/bjet.12720>
- Lee, D., Watson, S. L., & Watson, W. R. (2019). Systematic literature review on self-regulated learning in massive open online courses. *Australasian Journal of Educational Technology*, 35(1), 28–41. <https://doi.org/10.14742/AJET.3749>
- Leinonen, J., Leppänen, L., Ihantola, P., & Hellas, A. (2017). Comparison of Time Metrics in Programming. *Proceedings of the 2017 ACM Conference on International Computing Education Research*, 200–208. <https://doi.org/10.1145/3105726.3106181>
- Li, Y., & Lerner, R. M. (2011). Trajectories of School Engagement During Adolescence: Implications for Grades, Depression, Delinquency, and Substance Use. *Developmental Psychology*, 47(1). <https://doi.org/10.1037/a0021307>
- Liao, T. F., Bolano, D., Brzinsky-Fay, C., Cornwell, B., Fasang, A. E., Helske, S., Piccarreta, R., Raab, M., Ritschard, G., Struffolino, E., & Studer, M. (2022). Sequence analysis: Its past, present, and future. *Social Science Research*, 107. <https://doi.org/10.1016/j.ssresearch.2022.102772>
- Lim, L., Bannert, M., van der Graaf, J., Singh, S., Fan, Y., Surendrannair, S., Rakovic, M., Molenaar, I., Moore, J., & Gašević, D. (2023). Effects of real-time analytics-based personalized scaffolds on students' self-regulated learning. *Computers in Human Behavior*, 139. <https://doi.org/10.1016/j.chb.2022.107547>

- Lima, G. D. O., Costa, J. A. R., Araújo, R. D., & Dorça, F. A. (2023). Exploring the Relationship between Students Engagement and Self-Regulated Learning: A Case Study using OULAD Dataset and Machine Learning Techniques. *Anais Do XXXIV Simpósio Brasileiro de Informática Na Educação* (SBIE 2023), 1154–1165. <https://doi.org/10.5753/sbie.2023.234344>
- Liu, M., Cai, Y., Han, S., & Shao, P. (2023). Understanding Middle School Students' Self-Efficacy and Performance in a Technology-Enriched Problem-Based Learning Program: A Learning Analytics Approach. *Journal of Educational Technology Systems*, 51(4), 513–543. <https://doi.org/10.1177/00472395231174034>
- Lodge, J. M., Panadero, E., Broadbent, J., & de Barba, P. G. (2018). Supporting self-regulated learning with learning analytics. In Lodge, Horvath, & Corrin (Eds.), *Learning Analytics in the Classroom* (Issue October, pp. 45–55). Routledge. <https://doi.org/10.4324/9781351113038-4>
- López-Pernas, S., Heikkinen, S., Tlili, A., Majumdar, R., & Saqr, M. (2025). Unpacking learning in the age of AI: Bridging AI, Complexity, and Precision Education. In S. Mohammed & M. López-Pernas (Eds.), *Advanced learning analytics methods - AI, Precision and Complexity*. Springer. <https://lamethods.org/book2/chapters/ch01-intro/ch01-intro.html>
- López-Pernas, S., & Saqr, M. (2021). Bringing Synchrony and Clarity to Complex Multi-Channel Data: A Learning Analytics Study in Programming Education. *IEEE Access*, 9, 166531–166541. <https://doi.org/10.1109/ACCESS.2021.3134844>
- López-Pernas, S., & Saqr, M. (2024). How the dynamics of engagement explain the momentum of achievement and the inertia of disengagement: A complex systems theory approach. *Computers in Human Behavior*, 153, 108126. <https://doi.org/10.1016/j.chb.2023.108126>
- López-Pernas, S., Saqr, M., Gordillo, A., & Barra, E. (2023). A learning analytics perspective on educational escape rooms. *Interactive Learning*

- Environments, 31(10), 6509–6525.
<https://doi.org/10.1080/10494820.2022.2041045>
- López-Pernas, S., Saqr, M., Helske, S., & Murphy, K. (2024). Multi-Channel Sequence Analysis in Educational Research: An Introduction and Tutorial with R. In M. Saqr & S. López-Pernas (Eds.), *Learning Analytics Methods and Tutorials* (pp. 429–465). Springer Nature Switzerland.
https://doi.org/10.1007/978-3-031-54464-4_13
- López-Pernas, S., Saqr, M., & Viberg, O. (2021). Putting It All Together: Combining Learning Analytics Methods and Data Sources to Understand Students' Approaches to Learning Programming. *Sustainability*, 13(9), 4825. <https://doi.org/10.3390/su13094825>
- Lorenz, E. (2000). The Butterfly Effect. In *World Scientific Series on Nonlinear Science Series A*.
- Macfadyen, L. P., & Dawson, S. (2010). Mining LMS data to develop an “early warning system” for educators: A proof of concept. *Computers and Education*, 54(2), 588–599.
<https://doi.org/10.1016/j.compedu.2009.09.008>
- Malmberg, J., Saqr, M., Järvenoja, H., & Järvelä, S. (2022). How the Monitoring Events of Individual Students Are Associated With Phases of Regulation: A Network Analysis Approach. *Journal of Learning Analytics*, 9(1), 77–92. <https://doi.org/10.18608/jla.2022.7429>
- Martin, F., & Borup, J. (2022). Online learner engagement: Conceptual definitions, research themes, and supportive practices. *Educational Psychologist*, 57(3), 162–177.
<https://doi.org/10.1080/00461520.2022.2089147>
- Martinez-Maldonado, R., Echeverria, V., Fernandez-Nieto, G., Yan, L., Zhao, L., Alfredo, R., Li, X., Dix, S., Jaggard, H., Wotherspoon, R., Osborne, A., Shum, S. B., & Gašević, D. (2024). Lessons Learnt from a Multimodal Learning Analytics Deployment In-the-Wild. *ACM Transactions on Computer-Human Interaction*, 31(1), 1–41.
<https://doi.org/10.1145/3622784>
- Matcha, W., Gašević, D., Ahmad Uzir, N., Jovanović, J., Pardo, A., Lim, L., Maldonado-Mahauad, J., Gentili, S., Pérez-Sanagustín, M., & Tsai, Y.-S.

- (2020). Analytics of Learning Strategies: Role of Course Design and Delivery Modality. *Journal of Learning Analytics*, 7(2), 45–71. <https://doi.org/10.18608/jla.2020.72.3>
- Matcha, W., Gašević, D., Ahmad Uzir, N., Jovanović, J., Pardo, A., Maldonado-Mahauad, J., & Pérez-Sanagustín, M. (2019). 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)*: Vol. 11722 LNCS (pp. 525–540). https://doi.org/10.1007/978-3-030-29736-7_39
- Matcha, W., Uzir, N. A., Gašević, D., & Pardo, A. (2020). A Systematic Review of Empirical Studies on Learning Analytics Dashboards: A Self-Regulated Learning Perspective. *IEEE Transactions on Learning Technologies*, 13(2), 226–245. <https://doi.org/10.1109/TLT.2019.2916802>
- Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., Shekelle, P., Stewart, L. A., Estarli, M., Barrera, E. S. A., Martínez-Rodríguez, R., Baladia, E., Agüero, S. D., Camacho, S., Buhring, K., Herrero-López, A., Gil-González, D. M., Altman, D. G., Booth, A., ... Whitlock, E. (2016). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Revista Española de Nutrición Humana y Dietética*, 20(2). <https://doi.org/10.1186/2046-4053-4-1>
- Molenaar, I., & Chiu, M. M. (2014). Dissecting sequences of regulation and cognition: Statistical discourse analysis of primary school children's collaborative learning. *Metacognition and Learning*, 9(2). <https://doi.org/10.1007/s11409-013-9105-8>
- Molenaar, I., & Wise, A. F. (2022). Chapter 7: Temporal Aspects of Learning Analytics-Grounding Analyses in Concepts of Time. In *Handbook of Learning Analytics* (Vol. 2, pp. 66–76). <https://doi.org/10.18608/hla22.007>
- Monroy, C., Rangel, V. S., Bell, E. R., & Whitaker, R. (2015). A learning analytics approach to characterize and analyze inquiry-based pedagogical

- processes. Proceedings of the Fifth International Conference on Learning Analytics And Knowledge, 16-20-March-2015, 398–399. <https://doi.org/10.1145/2723576.2723658>
- Murphy, K., & Murphy, T. B. (2020). Gaussian parsimonious clustering models with covariates and a noise component. *Advances in Data Analysis and Classification*, 14(2), 293–325. <https://doi.org/10.1007/s11634-019-00373-8>
- Nafasa, P., Waspada, I., Bahtiar, N., & Wibowo, A. (2019). Implementation of Alpha Miner Algorithm in Process Mining Application Development for Online Learning Activities Based on MOODLE Event Log Data. *ICICOS 2019 - 3rd International Conference on Informatics and Computational Sciences: Accelerating Informatics and Computational Research for Smarter Society in The Era of Industry 4.0*, Proceedings. <https://doi.org/10.1109/ICICoS48119.2019.8982384>
- Nakamura, K., Ishihara, M., Horikoshi, I., & Ogata, H. (2024). Uncovering insights from big data: change point detection of classroom engagement. *Smart Learning Environments*, 11(1), 1–19. <https://doi.org/10.1186/S40561-024-00317-6/TABLES/3>
- Namoun, A., & Alshanqiti, A. (2020). Predicting Student Performance Using Data Mining and Learning Analytics Techniques: A Systematic Literature Review. *Applied Sciences* 2021, Vol. 11, Page 237, 11(1), 237. <https://doi.org/10.3390/APP11010237>
- Natthaphatwirata, R., & Matcha, W. (2024). Exploration of the online learners' actions: A sequence mining approach. *Online Journal of Communication and Media Technologies*, 14(4), e202446. <https://doi.org/10.30935/ojcmt/14957>
- Newman, M. (2018). *Networks* (2nd ed., Vol. 1). Oxford University Press. <https://doi.org/10.1093/oso/9780198805090.001.0001>
- Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., Breckler, S. J., Buck, S., Chambers, C. D., Chin, G., Christensen, G., Contestabile, M., Dafoe, A., Eich, E., Freese, J., Glennerster, R., Goroff, D., Green, D. P., Hesse, B., Humphreys, M., ... Yarkoni, T. (2015). Promoting an open

- research culture. In *Science* (Vol. 348, Issue 6242).
<https://doi.org/10.1126/science.aab2374>
- Ouyang, F., Xu, W., & Cukurova, M. (2023). An artificial intelligence-driven learning analytics method to examine the collaborative problem-solving process from the complex adaptive systems perspective. *International Journal of Computer-Supported Collaborative Learning*, 18(1), 39–66. <https://doi.org/10.1007/s11412-023-09387-z>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*, 372. <https://doi.org/10.1136/BMJ.N71>
- Pajares, F. (1996). Self-efficacy beliefs in academic settings. *Review of Educational Research*, 66(4).
<https://doi.org/10.3102/00346543066004543>
- Pan, M., Lai, C., & Guo, K. (2025). Effects of GenAI-empowered interactive support on university EFL students' self-regulated strategy use and engagement in reading. *The Internet and Higher Education*, 65, 100991. <https://doi.org/10.1016/J.IHEDUC.2024.100991>
- Pan, X. (2022). Exploring the multidimensional relationships between educational situation perception, teacher support, online learning engagement, and academic self-efficacy in technology-based language learning. *Frontiers in Psychology*, 13.
<https://doi.org/10.3389/fpsyg.2022.1000069>
- Panadero, E. (2017). A Review of Self-regulated Learning: Six Models and Four Directions for Research. *Frontiers in Psychology*, 8(APR).
<https://doi.org/10.3389/fpsyg.2017.00422>
- Pardo, A., & Siemens, G. (2014). Ethical and privacy principles for learning analytics. *British Journal of Educational Technology*, 45(3).
<https://doi.org/10.1111/bjet.12152>

- Park, E. H. (2024). Exploring the Differences in Learning Strategy Use Between Online and Offline Classes. *STEM Journal*, 25(1). <https://doi.org/10.16875/stem.2024.25.1.54>
- Park, E., Ifenthaler, D., & Clariana, R. B. (2023). Adaptive or adapted to: Sequence and reflexive thematic analysis to understand learners' self-regulated learning in an adaptive learning analytics dashboard. *British Journal of Educational Technology*, 54(1). <https://doi.org/10.1111/bjet.13287>
- Pati, D., & Lorusso, L. N. (2018). How to Write a Systematic Review of the Literature. *Health Environments Research and Design Journal*, 11(1). <https://doi.org/10.1177/1937586717747384>
- Pausch, J. (2023). SRL-O and gamification—the connection between the two elements in online learning platforms supporting learning goals. *MAP Education and Humanities*, 3(1). <https://doi.org/10.53880/2744-2373.2023.3.1.64>
- Pechenizkiy, M., Trčka, N., Vasilyeva, E., Van Der Aalst, W., & De Bra, P. (2009). Process mining online assessment data. EDM'09 - Educational Data Mining 2009: 2nd International Conference on Educational Data Mining.
- Peeters, W., Saqr, M., & Viberg, O. (2020). Applying Learning Analytics to Map Students ' Self-Regulated Learning Tactics in an Academic Writing Course. *Proceedings of the 28th International Conference on Computers in Education*, August, 245–254. <https://library.apsce.net/index.php/ICCE/article/view/3926>
- Pérez-Álvarez, R., Maldonado-Mahauad, J., & Pérez-Sanagustín, M. (2018). Tools to Support Self-Regulated Learning in Online Environments: Literature Review. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 11082 LNCS(January), 16–30. https://doi.org/10.1007/978-3-319-98572-5_2
- 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. <https://doi.org/10.1016/B978-012109890-2/50043-3>
- Pintrich, P., Smith, D., Garcia, T., & McKeachie, W. (1991). A manual for the use of the Motivated Strategies for Learning Questionnaire (MSLQ). In Ann Arbor. Michigan (Vol. 48109, Issue August 2016). <https://doi.org/ED338122>
- Polat, C., & Eroğlu, E. B. (2021). Learning Analytics and an Evaluation Specific to University Libraries. *Journal of Balkan Libraries Union*, 8(1). <https://doi.org/10.16918/jblu.937396>
- Poquet, O., Jovanović, J., & Pardo, A. (2023). Student Profiles of Change in a University Course: A Complex Dynamical Systems Perspective. *LAK23: 13th International Learning Analytics and Knowledge Conference*, 197–207. <https://doi.org/10.1145/3576050.3576077>
- Pozdniakov, S., Martinez-Maldonado, R., Tsai, Y.-S., Cukurova, M., Bartindale, T., Chen, P., Marshall, H., Richardson, D., & Gasevic, D. (2022). The Question-driven Dashboard: How Can We Design Analytics Interfaces Aligned to Teachers' Inquiry? *LAK22: 12th International Learning Analytics and Knowledge Conference*, 175–185. <https://doi.org/10.1145/3506860.3506885>
- Prasetya, R. E., & Raya, J. C. (2023). The interplay between self-regulated learning behavioral factors and students' performance in English language learning through Moodle. *ELT Forum: Journal of English Language Teaching*, 12(3), 145–156. <https://doi.org/10.15294/ELT.V12I3.66613>
- Prasse, D., Webb, M., Deschênes, M., Parent, S., Aeschlimann, F., Goda, Y., Yamada, M., & Raynault, A. (2024). Challenges in Promoting Self-Regulated Learning in Technology Supported Learning Environments: An Umbrella Review of Systematic Reviews and Meta-Analyses. *Technology, Knowledge and Learning*. <https://doi.org/10.1007/s10758-024-09772-z>
- Prinsloo, P., & Slade, S. (2013). An evaluation of policy frameworks for addressing ethical considerations in learning analytics. *Proceedings of*

- the Third International Conference on Learning Analytics and Knowledge, 240–244. <https://doi.org/10.1145/2460296.2460344>
- Prinsloo, P., & Slade, S. (2016). Student Vulnerability, Agency and Learning Analytics: An Exploration. *Journal of Learning Analytics*, 3(1). <https://doi.org/10.18608/jla.2016.31.10>
- Raab, M., & Struffolino, E. (2023). *Sequence Analysis*. SAGE Publications, Inc. <https://doi.org/10.4135/9781071938942>
- Ramaswami, G., Susnjak, T., & Mathrani, A. (2023). Effectiveness of a Learning Analytics Dashboard for Increasing Student Engagement Levels. *Journal of Learning Analytics*, 10(3). <https://doi.org/10.18608/jla.2023.7935>
- Redmond, P., Heffernan, A., Abawi, L., Brown, A., & Henderson, R. (2018). An Online Engagement Framework for Higher Education. *Online Learning*, 22(1), 183–204. <https://doi.org/10.24059/olj.v22i1.1175>
- Reschly, A. L., & Christenson, S. L. (2022). Jingle-Jangle Revisited: History and Further Evolution of the Student Engagement Construct. In *Handbook of Research on Student Engagement: Second Edition*. https://doi.org/10.1007/978-3-031-07853-8_1
- Rienties, B., Cross, S., & Zdrahal, Z. (2017). Implementing a Learning Analytics Intervention and Evaluation Framework: What Works? In B. K. Daniel (Ed.), *Big Data and Learning Analytics in Higher Education* (pp. 147–166). Springer International Publishing. https://doi.org/10.1007/978-3-319-06520-5_10
- Rincon-Flores, E. G., Castano, L., Guerrero Solis, S. L., Olmos Lopez, O., Rodríguez Hernández, C. F., Castillo Lara, L. A., & Aldape Valdés, L. P. (2024). Improving the learning-teaching process through adaptive learning strategy. *Smart Learning Environments*, 11(1), 1–27. <https://doi.org/10.1186/S40561-024-00314-9>
- Ritschard, G. (2023). Measuring the Nature of Individual Sequences. *Sociological Methods & Research*, 52(4), 2016–2049. <https://doi.org/10.1177/00491241211036156>
- Romero, C., Cerezo, R., Bogarín, A., & Sánchez-Santillán, M. (2016). Educational process mining: A tutorial and case study using moodle

- data sets. In *Data Mining And Learning Analytics: Applications in Educational Research*. <https://doi.org/10.1002/9781118998205.ch1>
- Romero, C., & Ventura, S. (2020). Educational data mining and learning analytics: An updated survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 10(3), e1355. <https://doi.org/10.1002/WIDM.1355>
- Rui, Y., & Liu, T. (2023). The effect of online English learners' perceived teacher support on self-regulation mediated by their self-efficacy. *Porta Linguarum*, 2023(40). <https://doi.org/10.30827/portalin.vi40.26930>
- Salas-Pilco, S. Z., Yang, Y., & Zhang, Z. (2022). Student engagement in online learning in Latin American higher education during the COVID-19 pandemic: A systematic review. In *British Journal of Educational Technology* (Vol. 53, Issue 3). <https://doi.org/10.1111/bjet.13190>
- Salmela-Aro, K., Moeller, J., Schneider, B., Spicer, J., & Lavonen, J. (2016). Integrating the light and dark sides of student engagement using person-oriented and situation-specific approaches. *Learning and Instruction*, 43, 61–70. <https://doi.org/10.1016/J.LEARNINSTRUC.2016.01.001>
- Samuel, T. S., & Warner, J. (2021). "I Can Math!": Reducing Math Anxiety and Increasing Math Self-Efficacy Using a Mindfulness and Growth Mindset-Based Intervention in First-Year Students. *Community College Journal of Research and Practice*, 45(3). <https://doi.org/10.1080/10668926.2019.1666063>
- Saqr, M., & Alamro, A. (2019). The role of social network analysis as a learning analytics tool in online problem based learning. *BMC Medical Education*, 19(1), 1–11. <https://doi.org/https://doi.org/10.1186/s12909-019-1599-6>
- Saqr, M., Dever, D., López-Pernas, S., Gernigon, C., Marchand, G., & Kaplan, A. (2024). Complex Dynamic Systems in Education: Beyond the Static, the Linear and the Causal Reductionism. <http://arxiv.org/abs/2501.10386>

- Saqr, M., & López-Pernas, S. (2021). The longitudinal trajectories of online engagement over a full program. *Computers & Education*, 175(September), 104325. <https://doi.org/10.1016/j.compedu.2021.104325>
- Saqr, M., & López-Pernas, S. (2022). How CSCL roles emerge, persist, transition, and evolve over time: A four-year longitudinal study. *Computers & Education*, 189, 104581. <https://doi.org/10.1016/j.compedu.2022.104581>
- Saqr, M., & López-Pernas, S. (2023). The temporal dynamics of online problem-based learning: Why and when sequence matters. *International Journal of Computer-Supported Collaborative Learning*, 1–27. <https://doi.org/10.1007/S11412-023-09385-1>
- Saqr, M., & López-Pernas, S. (2024). Mapping the self in self-regulation using complex dynamic systems approach. *British Journal of Educational Technology*, 55(4), 1376–1397. <https://doi.org/10.1111/bjet.13452>
- Saqr, M., & López-Pernas, S. (2025). Advanced learning analytics methods - AI, Precision and Complexity (M. Saqr & S. López-Pernas, Eds.). Springer. <https://lamethods.github.io/>
- Saqr, M., López-Pernas, S., Helske, S., Durand, M., Murphy, K., Studer, M., & Ritschard, G. (2024). Sequence Analysis in Education: Principles, Technique, and Tutorial with R. In M. Saqr & S. López-Pernas (Eds.), *Learning Analytics Methods and Tutorials* (pp. 321–354). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-54464-4_10
- Saqr, M., López-Pernas, S., Helske, S., & Hrastinski, S. (2023). The longitudinal association between engagement and achievement varies by time, students' profiles, and achievement state: A full program study. *Computers & Education*, 199, 104787. <https://doi.org/10.1016/J.COMPEDU.2023.104787>
- Saqr, M., López-Pernas, S., Jovanović, J., & Gašević, D. (2023). Intense, turbulent, or wallowing in the mire: A longitudinal study of cross-course online tactics, strategies, and trajectories. *The Internet and Higher Education*, 57, 100902. <https://doi.org/10.1016/j.iheduc.2022.100902>

- Saqr, M., López-Pernas, S., & Tikka, S. (2025). Mapping Relational Dynamics with Transition Network Analysis: A Primer and Tutorial. In M. Saqr & S. López-Pernas (Eds.), *Advanced Learning Analytics Methods: AI, Precision and Complexity*. Springer. <https://lamethods.org/book2/chapters/ch15-tna/ch15-tna.html>
- Saqr, M., López-Pernas, S., Törmänen, T., Kaliisa, R., Misiejuk, K., & Tikka, S. (2024). Transition Network Analysis: A Novel Framework for Modeling, Visualizing, and Identifying the Temporal Patterns of Learners and Learning Processes. <http://arxiv.org/abs/2411.15486>
- Saqr, M., & Nouri, J. (2020). High resolution temporal network analysis to understand and improve collaborative learning. *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge*, 314–319. <https://doi.org/10.1145/3375462.3375501>
- Saqr, M., & Peeters, W. (2022). Temporal networks in collaborative learning: A case study. *British Journal of Educational Technology*, 53(5). <https://doi.org/10.1111/bjet.13187>
- Schmitz, B. (2006). Advantages of studying processes in educational research. *Learning and Instruction*, 16(5), 433–449. <https://doi.org/10.1016/j.learninstruc.2006.09.004>
- Schneider, B., Davis, R., Martinez-Maldonado, R., Biswas, G., Worsley, M., & Rummel, N. (2024). Stepping Outside the Ivory Tower: How Can We Implement Multimodal Learning Analytics in Ecological Settings, and Turn Complex Temporal Data Sources into Actionable Insights? *CSCL 2024 Proceedings*, 323–330. <https://doi.org/10.22318/cscl2024.259119>
- Schunk, D. H. (1991). Self-Efficacy and Academic Motivation. *Educational Psychologist*, 26(3–4), 207–231. <https://doi.org/10.1080/00461520.1991.9653133>
- Schunk, D. H. (2011). *Handbook of Self-Regulation of Learning and Performance*. Routledge. <https://doi.org/10.4324/9780203839010>
- Schunk, D. H., & Zimmerman, B. J. (2023). Self-Regulation of Learning And Performance: Issues and Educational Applications. In D. H. Schunk & B. J. Zimmerman (Eds.), *Self-Regulation of Learning and Performance*

- Issues and Educational Applications. Taylor and Francis.
[https://doi.org/https://doi.org/10.4324/9780203763353](https://doi.org/10.4324/9780203763353)
- Schwarz, G. (2007). Estimating the Dimension of a Model. *The Annals of Statistics*, 6(2). <https://doi.org/10.1214/aos/1176344136>
- Sedrakyan, G., Malmberg, J., Verbert, K., Järvelä, S., & Kirschner, P. A. (2020). Linking learning behavior analytics and learning science concepts: Designing a learning analytics dashboard for feedback to support learning regulation. *Computers in Human Behavior*, 107(2014), 105512. <https://doi.org/10.1016/j.chb.2018.05.004>
- Shafiq, D. A., Marjani, M., Habeeb, R. A. A., & Asirvatham, D. (2022). Student Retention Using Educational Data Mining and Predictive Analytics: A Systematic Literature Review. *IEEE Access*, 10, 72480–72503. <https://doi.org/10.1109/ACCESS.2022.3188767>
- Shen, D., Cho, M. H., Tsai, C. L., & Marra, R. (2013). Unpacking online learning experiences: Online learning self-efficacy and learning satisfaction. *Internet and Higher Education*, 19. <https://doi.org/10.1016/j.iheduc.2013.04.001>
- Shin, D. D. (2024). Curiosity promotes self-regulated learning and achievement in online courses for students with varying self-efficacy levels. *Educational Psychology*, 1–20. <https://doi.org/10.1080/01443410.2024.2372302>
- Siadaty, M., Gašević, D., & Hatala, M. (2016). Associations between technological scaffolding and micro-level processes of self-regulated learning: A workplace study. *Computers in Human Behavior*, 55, 1007–1019. <https://doi.org/10.1016/j.chb.2015.10.035>
- Siemens, G. (2013). Learning Analytics. *American Behavioral Scientist*, 57(10), 1380–1400. <https://doi.org/10.1177/0002764213498851>
- Siemens, G., & Long, P. (2011). Penetrating the Fog: Analytics in Learning and Education. *EDUCAUSE Review*, 46(5). <https://er.educause.edu/articles/2011/9/penetrating-the-fog-analytics-in-learning-and-education>
- Silvola, A., Sjöblom, A., Näykki, P., Gedrimiene, E., & Muukkonen, H. (2023). Learning analytics for academic paths: student evaluations of two

- dashboards for study planning and monitoring. *Frontline Learning Research*, 11(2). <https://doi.org/10.14786/flr.v11i2.1277>
- Slade, S., & Prinsloo, P. (2013). Learning Analytics: Ethical Issues and Dilemmas. *American Behavioral Scientist*, 57(10). <https://doi.org/10.1177/0002764213479366>
- Song, Y., Zhang, R., Li, Z., Xin, C., & Ding, N. (2024). Unveiling the link between self-regulated learning and academic success: a longitudinal study on Chinese medical students. *Advances in Health Sciences Education*, 1–17. <https://doi.org/10.1007/S10459-024-10400-9>
- Stoyanov, S., & Kirschner, P. A. (2023). Text analytics for uncovering untapped ideas at the intersection of learning design and learning analytics: Critical interpretative synthesis. *Journal of Computer Assisted Learning*. <https://doi.org/10.1111/jcal.12775>
- Studer, M. (2013). *WeightedCluster Library Manual: A practical guide to creating typologies of trajectories in the social sciences with R*. (LIVES Working Papers, 24). <https://doi.org/10.12682/lives.2296-1658>
- Su, Y., Li, Y., Liang, J. C., & Tsai, C. C. (2019). Moving literature circles into wiki-based environment: the role of online self-regulation in EFL learners' attitude toward collaborative learning. *Computer Assisted Language Learning*, 32(5–6). <https://doi.org/10.1080/09588221.2018.1527363>
- Sulistiawati, S., Kusumah, Y. S., Dahlan, J. A., Juandi, D., & Vos, H. (2023). A Bibliometric Analysis: Trend of Studies in Self-Regulated Learning Over The Past Three Decades. *Indonesian Journal on Learning and Advanced Education (IJOLAE)*, 5(2). <https://doi.org/10.23917/ijolae.v5i2.21381>
- Talsma, K., Schüz, B., & Norris, K. (2019). Miscalibration of self-efficacy and academic performance: Self-efficacy ≠ self-fulfilling prophecy. *Learning and Individual Differences*, 69, 182–195. <https://doi.org/10.1016/J.LINDIF.2018.11.002>
- Tan, T. K., & Samavedham, L. (2022). The learning process matter: A sequence analysis perspective of examining procrastination using learning management system. *Computers and Education Open*, 3, 100112. <https://doi.org/10.1016/j.caeo.2022.100112>

- Tao, X., Hanif, H., & Ebrahim, N. A. (2023). Emerging Trends of Self-regulated Learning: A Comprehensive Bibliometric Analysis. *World Journal of English Language*, 13(6). <https://doi.org/10.5430/WJEL.V13N6P252>
- Tempelaar, D., Rienties, B., Giesbers, B., & Nguyen, Q. (2023). Modelling Temporality in Person-and Variable-Centred Approaches. *Journal of Learning Analytics*, 10(2). <https://doi.org/10.18608/jla.2023.7841>
- Tempelaar, D., Rienties, B., & Nguyen, Q. (2020). Subjective data, objective data and the role of bias in predictive modelling: Lessons from a dispositional learning analytics application. *PLOS ONE*, 15(6), e0233977. <https://doi.org/10.1371/JOURNAL.PONE.0233977>
- Teng, L. S., & Zhang, L. J. (2020). Empowering learners in the second/foreign language classroom: Can self-regulated learning strategies-based writing instruction make a difference? *Journal of Second Language Writing*, 48. <https://doi.org/10.1016/j.jslw.2019.100701>
- Tepgec, M., Heil, J., & Ifenthaler, D. (2024). Feedback literacy matters: unlocking the potential of learning analytics-based feedback. *Assessment & Evaluation in Higher Education*, 1–17. <https://doi.org/10.1080/02602938.2024.2367587>
- Thomas, D. R. (2006). A General Inductive Approach for Analyzing Qualitative Evaluation Data. *American Journal of Evaluation*, 27(2). <https://doi.org/10.1177/1098214005283748>
- Tinto, V. (2022). Exploring the Character of Student Persistence in Higher Education: The Impact of Perception, Motivation, and Engagement. In *Handbook of Research on Student Engagement: Second Edition*. https://doi.org/10.1007/978-3-031-07853-8_17
- Törmänen, T., Järvenoja, H., Saqr, M., Malmberg, J., & Järvelä, S. (2022). A Person-Centered Approach to Study Students' Socio-Emotional Interaction Profiles and Regulation of Collaborative Learning. *Frontiers in Education*, 7. <https://doi.org/10.3389/feduc.2022.866612>
- Törmänen, T., Järvenoja, H., Saqr, M., Malmberg, J., & Järvelä, S. (2023). Affective states and regulation of learning during socio-emotional interactions in secondary school collaborative groups. *British Journal*

- of Educational Psychology, 93(S1), 48–70.
<https://doi.org/10.1111/bjep.12525>
- Tsai, C. L., Cho, M. H., Marra, R., & Shen, D. (2020). The Self-Efficacy Questionnaire for Online Learning (SeQoL). *Distance Education*, 41(4).
<https://doi.org/10.1080/01587919.2020.1821604>
- Tsai, Y.-S., Whitelock-Wainwright, A., & Gašević, D. (2020). The privacy paradox and its implications for learning analytics. *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge*, 230–239. <https://doi.org/10.1145/3375462.3375536>
- Ucar, H., & Ugurhan, Y. Z. C. (2023). The Role of E-Learning Readiness on Self-Regulation in Open and Distance Learning. *Turkish Online Journal of Distance Education*, 24(4). <https://doi.org/10.17718/tojde.1231705>
- Uzir, N. A., Gašević, D., Jovanović, J., Matcha, W., Lim, L. A., & Fudge, A. (2020). Analytics of time management and learning strategies for effective online learning in blended environments. *Proceedings of the 10th International Conference on Learning Analytics and Knowledge (LAK '20)*, March 23-27, Frankfurt, Germany.
<https://doi.org/10.1145/3375462.3375493>
- van Alten, D. C. D., Phielix, C., Janssen, J., & Kester, L. (2020). Effects of self-regulated learning prompts in a flipped history classroom. *Computers in Human Behavior*, 108, 106318.
<https://doi.org/10.1016/J.CHB.2020.106318>
- Vardal-Ocakli, S., & Ok, A. (2024). Personal and social facilitators of student engagement: transition into higher education. *Frontiers in Education*, 9, 1286867. <https://doi.org/10.3389/FEDUC.2024.1286867/BIBTEX>
- Vartiainen, H., López-Pernas, S., Saqr, M., Kahila, J., Parkki, T., Tedre, M., & Valtonen, T. (2022). Mapping students' temporal pathways in a computational thinking escape room. *CEUR Workshop Proceedings*, 3383. https://ceur-ws.org/Vol-3383/FLAIEC22_paper_9625.pdf
- Vergnes, J. N., Marchal-Sixou, C., Nabet, C., Maret, D., & Hamel, O. (2010). Ethics in systematic reviews. *Journal of Medical Ethics*, 36(12).
<https://doi.org/10.1136/jme.2010.039941>

- Viberg, O., Hatakka, M., Bälter, O., & Mavroudi, A. (2018). The current landscape of learning analytics in higher education. *Computers in Human Behavior*, 89(October 2017), 98–110. <https://doi.org/10.1016/j.chb.2018.07.027>
- Viberg, O., Khalil, M., & Baars, M. (2020). Self-regulated learning and learning analytics in online learning environments. *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge*, March, 524–533. <https://doi.org/10.1145/3375462.3375483>
- Villagran, I., Hernández, R., Schuit, G., Neyem, A., Fuentes, J., Larrondo, L., Margozzini, E., Hurtado, M. T., Iriarte, Z., Miranda, C., Varas, J., & Hilliger, I. (2024). Enhancing Feedback Uptake and Self-Regulated Learning in Procedural Skills Training. *Journal of Learning Analytics*, 1–19. <https://doi.org/10.18608/jla.2024.8195>
- Villalobos, E. (2024). Developing a Learning Analytics framework to understand the temporal behavior of students in Blended Learning Environments [PhD Dissertation]. Université de Toulouse.
- Villalobos, E., Perez-Sanagustin, M., Azevedo, R., Sanza, C., & Broisin, J. (2024). Exploring Manifestations of Learners' Self-Regulated Tactics and Strategies Across Blended Learning Courses. *IEEE Transactions on Learning Technologies*. <https://doi.org/10.1109/TLT.2024.3385641>
- Villalobos, E., Pérez-Sanagustín, M., & Broisin, J. (2024). From Learning Actions to Dynamics: Characterizing Students' Individual Temporal Behavior with Sequence Analysis. 3–17. https://doi.org/10.1007/978-3-031-64302-6_1
- Villalobos, E., Pérez-Sanagustin, M., Sanza, C., Tricot, A., & Broisin, J. (2022). Supporting Self-regulated Learning in BL: Exploring Learners' Tactics and Strategies. In *Educating for a New Future: Making Sense of Technology-Enhanced Learning Adoption: 17th European Conference on Technology Enhanced Learning, EC-TEL 2022, Toulouse, France, September 12–16, 2022, Proceedings*. (pp. 407–420). https://doi.org/10.1007/978-3-031-16290-9_30

- Wager, E., & Wiffen, P. J. (2011). Ethical issues in preparing and publishing systematic reviews. In *Journal of Evidence-Based Medicine* (Vol. 4, Issue 2). <https://doi.org/10.1111/j.1756-5391.2011.01122.x>
- Wambsganss, T., & Schmitt, A. (2024). Enhancing Personalized Learning Through Process Mining: A Taxonomy and Design Patterns Approach. *Business and Information Systems Engineering*, 1–24. <https://doi.org/10.1007/S12599-024-00901-7>
- Wang, M., & Degol, J. (2014). Staying Engaged: Knowledge and Research Needs in Student Engagement. *Child Development Perspectives*, 8(3), 137–143. <https://doi.org/10.1111/cdep.12073>
- Wang, W., Lee, H., Lin, C., Li, P., Huang, Y., & Wu, T. (2024). Enhancing students' learning outcomes in self-regulated virtual reality learning environment with learning aid mechanisms. *British Journal of Educational Technology*. <https://doi.org/10.1111/bjet.13512>
- Wigfield, A., & Eccles, J. S. (2000). Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology*, 25(1). <https://doi.org/10.1006/ceps.1999.1015>
- Wigfield, A., Eccles, J. S., Schiefele, U., Roeser, R. W., & Davis-Kean, P. (2007). Development of Achievement Motivation. In *Handbook of Child Psychology*. Wiley. <https://doi.org/10.1002/9780470147658.chpsy0315>
- Winne, P. H. (2017). Learning Analytics for Self-Regulated Learning. In *Handbook of Learning Analytics* (pp. 241–249). <https://doi.org/10.18608/hla17.021>
- Winne, P. H. (2022). Modeling self-regulated learning as learners doing learning science: How trace data and learning analytics help develop skills for self-regulated learning. *Metacognition and Learning*, 17(3), 773–791. <https://doi.org/10.1007/s11409-022-09305-y>
- Winne, P. H., & Hadwin, A. F. (1998). Studying as self-regulated engagement in learning. In D. Hacker, J. Dunlosky, & A. Graesser (Eds.), *Metacognition in Educational Theory and Practice* (pp. 277–304). Erlbaum.

- Winne, P. H., & Marzouk, Z. (2019). Learning Strategies and Self-Regulated Learning. *The Cambridge Handbook of Cognition and Education*, 696–715. <https://doi.org/10.1017/9781108235631.028>
- Wise, A. F. (2014). Designing Pedagogical Interventions to Support Student Use of Learning Analytics. *Proceedings of the Fourth International Conference on Learning Analytics And Knowledge*. <https://doi.org/10.1145/2567574>
- Wong, J., Baars, M., Davis, D., Van Der Zee, T., Houben, G. J., & Paas, F. (2019). Supporting Self-Regulated Learning in Online Learning Environments and MOOCs: A Systematic Review. *International Journal of Human-Computer Interaction*, 35(4–5), 356–373. <https://doi.org/10.1080/10447318.2018.1543084>
- Wong, J., Baars, M., He, M., de Koning, B. B., & Paas, F. (2021). Facilitating goal setting and planning to enhance online self-regulation of learning. *Computers in Human Behavior*, 124. <https://doi.org/10.1016/j.chb.2021.106913>
- Worsley, M., Anderson, K., Melo, N., & Jang, J. (2021). Designing analytics for collaboration literacy and student empowerment. *Journal of Learning Analytics*, 8(1). <https://doi.org/10.18608/JLA.2021.7242>
- Wu, Y., Xu, X., Xue, J., & Hu, P. (2023). A cross-group comparison study of the effect of interaction on satisfaction in online learning: The parallel mediating role of academic emotions and self-regulated learning. *Computers and Education*, 199. <https://doi.org/10.1016/j.compedu.2023.104776>
- Wylie, C., & Hodgen, E. (2012). Trajectories and Patterns of Student Engagement: Evidence from a Longitudinal Study. In *Handbook of Research on Student Engagement* (pp. 585–599). Springer US. https://doi.org/10.1007/978-1-4614-2018-7_28
- Xu, J., Li, J., & Yang, J. (2024). Self-regulated learning strategies, self-efficacy, and learning engagement of EFL students in smart classrooms: A structural equation modeling analysis. *System*, 125, 103451. <https://doi.org/10.1016/J.SYSTEM.2024.103451>

- Xu, J., & Zhou, S. (2024). Unraveling the interplay between social-emotional need satisfaction and self-regulated learning: Longitudinal evidence from university students in China's EMI programs. <https://doi.org/10.1177/13621688241304875>.
- Yan, L., Martinez-Maldonado, R., Swiecki, Z., Zhao, L., Li, X., & Gašević, D. (2024). Dissecting the temporal dynamics of embodied collaborative learning using multimodal learning analytics. *Journal of Educational Psychology*. <https://doi.org/10.1037/edu0000905>
- Ye, D., & Pennisi, S. (2022). Using trace data to enhance Students' self-regulation: A learning analytics perspective. *Internet and Higher Education*, 54. <https://doi.org/10.1016/j.iheduc.2022.100855>
- Yin, D., & Luo, L. (2024). The influence of perceived teacher support on online English learning engagement among Chinese university students: a cross-sectional study on the mediating effects of self-regulation. *Frontiers in Psychology*, 15. <https://doi.org/10.3389/fpsyg.2024.1246958>
- You, S., & Sharkey, J. (2009). Testing a developmental-ecological model of student engagement: a multilevel latent growth curve analysis. *Educational Psychology*, 29(6), 659–684. <https://doi.org/10.1080/01443410903206815>
- Yundayani, A., Abdullah, F., Tandiana, S. T., & Sutrisno, B. (2021). Students' cognitive engagement during emergency remote teaching: Evidence from the Indonesian EFL milieu. *Journal of Language and Linguistic Studies*, 17(1). <https://doi.org/10.52462/jlls.2>
- Zhang, Y., & Paquette, L. (2023). Sequential pattern mining in educational data: the application context, potential, strengths, and limitations. In A. Peña-Ayala (Ed.), *Educational Data Science: Essentials, Approaches, and Tendencies: Proactive Education Based on Empirical Big Data Evidence* (1st ed., pp. 219–254). Springer. <https://doi.org/10.1007/978-981-99-0026-8>
- Zhen, R., Liu, R. De, Wang, M. Te, Ding, Y., Jiang, R., Fu, X., & Sun, Y. (2020). Trajectory patterns of academic engagement among elementary

- school students: The implicit theory of intelligence and academic self-efficacy matters. *British Journal of Educational Psychology*, 90(3). <https://doi.org/10.1111/bjep.12320>
- Zheng, J., Jiang, N., & Dou, J. (2020). Autonomy support and academic stress: A relationship mediated by self-regulated learning and mastery goal orientation. *New Waves Educational Research & Development*, 23(summer). <https://eric.ed.gov/?id=EJ1264274>
- Zheng, Y., & Xiao, A. (2023). A structural equation model of online learning: investigating self-efficacy, informal digital learning, self-regulated learning, and course satisfaction. *Frontiers in Psychology*, 14. <https://doi.org/10.3389/fpsyg.2023.1276266>
- Zhou, W., Zhao, L., & Kaabar, M. (2022). The Effect of Teachers' Support on Learners' Online Self-Regulated Learning: Mediating Analysis Based on Self-Efficacy. *International Journal of Emerging Technologies in Learning*, 17(17). <https://doi.org/10.3991/ijet.v17i17.34027>
- Zimmerman, B. J. (2000). Attaining Self-Regulation. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of Self-Regulation* (pp. 13–39). Elsevier. <https://doi.org/10.1016/B978-012109890-2/50031-7>
- Zimmerman, B. J., Bandura, A., & Martinez-Pons, M. (1992). Self-Motivation for Academic Attainment: The Role of Self-Efficacy Beliefs and Personal Goal Setting. *American Educational Research Journal*, 29(3). <https://doi.org/10.3102/00028312029003663>
- Zimmerman, B., & Kitsantas, A. (2007). Reliability and Validity of Self-Efficacy for Learning Form (SELF) Scores of College Students. *Zeitschrift Für Psychologie / Journal of Psychology*, 215(3), 157–163. <https://doi.org/10.1027/0044-3409.215.3.157>
- Zimmerman, W. A., & Kulikowich, J. M. (2016). Online Learning Self-Efficacy in Students With and Without Online Learning Experience. *American Journal of Distance Education*, 30(3), 180–191. <https://doi.org/10.1080/08923647.2016.1193801>

Appendices

APPENDIX 1. RESEARCH PERMIT FOR STUDY 2

Tutkimuslupahakemus

Vastaaja:

-

Vastaus:

01.09.2021, 10:49 - 01.09.2021, 10:51

1. Hakijan yhteystiedot *

Etunimi	Sami
Sukunimi	Heikkinen
Matkapuhelin	0447081820
Sähköposti	sami.heikkinen@lab.fi
Osoite	
Postinumero	
Postitoimipaikka	Lahti
Yritys/organisaatio	LAB ammattikorkeakoulu

2. Tutkimuksen nimi tai nimiluonnos *

Prosessinlouhinta itsesääntyvän oppimisen tukena

3. Tutkimuksen tarkoitus *

Tutkimus on osa väitöskirjaprosessia. Väitöskirjassa on tavoitteena luoda malli, jonka avulla tunnistetaan opintojaksoilla hyvin ja heikosti menestyvät opiskelijat prosessin louhintaa hyödyntäen mahdollisimman varhaisessa vaiheessa. Tämän tutkimusluvan kohteena olevan tutkimuksen tarkoituksena on tunnistaa opintojaksolta opiskelijat, jotka suoriutuvat opintojaksolla keskimääräistä paremmin ja saavuttavat korkeammat arvosanat. Prosessin louhinnan avulla analysoidaan Moodlen lokidataa ja tunnistetaan sen avulla tapoja, joilla hyvin menestyvät opiskelijat toimivat saavuttaakseen tavoitteensa. Tutkimuksen myöhemmissä vaiheissa tämän datan perusteella kehitetään interventioita, joilla heikosti menestyviksi ennustettujen opiskelijoiden kohdalla pyritään vaikuttamaan heidän opiskelutapoihinsa verkko-oppimisympäristössä ja näin parannetaan heidän opintojakson läpäisyyään sekä saavuttamia oppimistuloksia.

4. Tutkimusmenetelmätyökalu *

Opintojakson päättymisen jälkeen opintojakson osallistujien lokidata yhdistetään opintojaksolta saavutettuun arvosanaan. Yhdistämisen jälkeen osallistujien tiedot anonymisoidaan. Tutkimusmenetelmänä käytetään prosessinlouhintaa, jolla analysoidaan Moodlen lokidataa tunnistetaan erilaisia oppimisprosesseja, jotka erottavat opintojaksolla hyvin ja heikommin menestyneet opiskelijat.

5. Aikataulu *

Lokidataa kerätään opiskelijoista lukuvuoden 2021-2022 ajan. Tämän jälkeen tehdään tiedon analysointi. Aiheesta syntyvä artikkeli on tavoitteena saada julkaistuksi vuoden 2022 aikana.

6. Laajuus *

Opiskelijoilta ei kerätä tietoa esim. kyselyjen tai haastattelujen avulla. Tutkimusaineisto syntyy opiskelijoiden normaalin opiskelun aikana Moodlen lokitietoineen sekä suoritustiedoiksi opintorekisteriin. Tietoja kerätään kahdelta

Tutkimuslupa, hakijana: / 6066ba2b-83ab-44de-8024-9a487720756e / sivu 1/25

opintojaksototeutukselta, joissa opiskelijoita on kaikkiaan arviolta n. 150.

7. Kohderyhmä ja osallistujien määrän tavoite *

Tutkimus kohdistuu Liiketalouden ja matkailun alan koulutusten opintojaksoon Projektinhallinta ja työelämäprojekti, jonka monimuoto-opiskelijoille suunnattuihin verkkototeutuksiin osallistuu lukuvuoden 2021-2022 aikana arviolta n. 150 opiskelijaa.

8. Hakemuksen liitteet

Tutkimussuunnitelma (pakollinen)

Tietosuojailmoitus (pakollinen) mallipohja löytyy osoitteesta www.lab.fi

9. Menettelytavat *

Sitoudun noudattamaan Tutkimuseettisen neuvottelukunnan julkaisemia hyvän tieteellisen käytännön menettelytapoja. (www.tenk.fi --> HTK-ohje)

Tämä dokumentti on allekirjoitettu sähköisesti LUT Sign-järjestelmällä

This document has been electronically signed using LUT Sign

Päiväys / Date: 01.09.2021 15:38:06 (UTC +0300)

Milla Laisi-Wessman

LAB-ammattikorkeakoulu

yksikön johtaja

Organisation varmennema (LUT käyttäjätunnus)

Certified by organization (LUT user account)

Certified by organization

APPENDIX 2. RESEARCH PERMIT FOR STUDY 3

Research permit for P3

Tutkimuslupahakemus

Vastaaja:

*

Vastaus:

19.04.2023, 15:35 - 19.04.2023, 15:37

1. Hakijan yhteystiedot *

Etunimi	Sami
Sukunimi	Heikkinen
Matkapuhelin	0447081820
Sähköposti	sami.heikkinen@lab.fi
Osoite	
Postinumero	
Postitoimipaikka	Lahti
Yritys/organisaatio	LAB-ammattikorkeakoulu

2. Tutkimuksen nimi tai nimiluonnos *

Prosessinlouhinta itsesäätyvän oppimisen tukena

3. Tutkimuksen tarkoitus *

Tämä tutkimus on osa Sami Heikkisen väitöskirjatutkimusta. Väitöskirjassa on tavoitteena luoda malli, jonka avulla tunnistetaan opintojaksoilla hyvin ja heikosti menestyvät opiskelijat sekvenssianalyysia ja prosessin louhintaa hyödyntäen mahdollisimman varhaisessa vaiheessa ja löytää malleja opiskelijoiden oppimisen tukemiseen. Tämän tutkimusluvan kohteena olevan tutkimuksen tarkoituksena on mitata opiskelijoiden itsesäätyvän oppimisen taitoja ja niissä tapahtuvia muutoksia opintojakson aikana sekä tunnistaa opiskelijat, jotka tarvitsevat tukea opintojakson loppuun suorittamisessa ja/tai omien tavoitteiden saavuttamisessa. Kerätyn tiedon avulla kohdistetaan oppimista tukevia toimia niitä tarvitseville opiskelijoille.

4. Tutkimusmenetelmä/työkalu

*

Opintojakson päättymisen jälkeen opintojakson osallistujien lokidata ja opintojakson aikana kerätty kyselyaineisto yhdistetään opintojaksolta saavutettuun arvosanaan. Yhdistämisen jälkeen osallistujien tiedot anonymisoidaan eli yksittäisen opiskelijan henkilöllisyys ei paljastu tietojen käsittelyn aikana. Tutkimusmenetelmänä käytetään sekvenssianalyysia ja prosessinlouhintaa, joilla analysoidaan Moodlen lokidataa tunnistetaan oppimisprosessien eroja, mitkä erottavat opintojaksolla eri tavoin menestyvät opiskelijat. Kyselyaineistoa analysoidaan erilaisten itsesäätyvän oppimisen taitojen tunnistamiseksi ja sen perusteella tukitoimien kohdistamiseksi niitä tarvitseville opiskelijoille.

5. Aikataulu *

Aineisto kerätään opiskelijoista ajanjaksolla kevät 2023 – kesä 2024. Tämän jälkeen tehdään tiedon analysointi. Aiheesta syntyvä artikkeli on tavoitteena saada julkaistuksi vuoden 2024 aikana.

6. Laajuus *

Tutkimusaineisto syntyy opiskelijoiden normaalin opiskelun aikana Moodlen lokitietoihin ja kyselytietona tenttävapahtumusten yhteydessä sekä suoritustiedoiksi opintorekisteriin. Tietoja kerätään neljältä opintojaksototeutukselta,

7. Kohderyhmä ja osallistujien määrän tavoite *

Tutkimus kohdistuu Liiketalouden ja matkailun alan koulutusten opintojaksoon Projektinhallinta ja työelämäprojekti, jonka verkkototeutuksiin osallistuu aikataulun mukaisena ajanjaksona arviolta n. 300 opiskelijaa.

8. Hakemuksen liitteet

Tutkimussuunnitelma (pakollinen)

Tietosuojailmoitus (pakollinen) mallipohja löytyy osoitteesta www.lab.fi

9. Menettelytavat *

Sitoudun noudattamaan Tutkimuseettisen neuvottelukunnan julkaisemia hyvän tieteellisen käytännön menettelytapoja. (www.tenk.fi --> HTK-ohje)

Tämä dokumentti on allekirjoitettu sähköisesti LUT Sign-järjestelmällä

This document has been electronically signed using LUT Sign

Päiväys / Date: 24.04.2023 17:30:38 (UTC +0300)

Nora Ahtola

LAB-ammattikorkeakoulu

Yksikön johtaja

Organisaation varmentama (LUT käyttäjätunnus)
Certified by organization (LUT user account)

Certified by organization

ARTICLE I

© Springer 2023. Reprinted, with permission from:

Heikkinen, S., Saqr, M., Malmberg, J., & Tedre, M. (2023). Supporting self-regulated learning with learning analytics interventions – a systematic literature review. *Educ Inf Technol* 28, 3059–3088.

<https://doi.org/10.1007/s10639-022-11281-4>

Supporting self-regulated learning with learning analytics interventions – a systematic literature review

Sami Heikkinen^{1,2}  · Mohammed Saqr¹  · Jonna Malmberg³  · Matti Tedre¹ 

Received: 9 June 2022 / Accepted: 4 August 2022 / Published online: 8 September 2022
© The Author(s) 2022

Abstract

During the past years scholars have shown an increasing interest in supporting students' self-regulated learning (SRL). Learning analytics (LA) can be applied in various ways to identify a learner's current state of self-regulation and support SRL processes. It is important to examine how LA has been used to identify the need for support in different phases of SRL cycle, which channels are used to mediate the intervention and how efficient and impactful the intervention is. This will help the learners to achieve the anticipated learning outcomes. The systematic literature review followed PRISMA 2020 statement to examine studies that applied LA interventions to enhance SRL. The search terms used for this research identified 753 papers in May 2021. Of these, 56 studies included the elements of LA, SRL, and intervention. The reviewed studies contained various LA interventions aimed at supporting SRL, but only 46% of them revealed a positive impact of an intervention on learning. Furthermore, only four studies reported positive effects for SRL and covered all three SRL phases (planning, performance, and reflection). Based on the findings of this literature review, the key recommendation is for all phases of SRL to be considered when planning interventions to support learning. In addition, more comparative research on this topic is needed to identify the most effective interventions and to provide further evidence on the effectiveness of interventions supporting SRL.

Keywords Learning analytics · Self-regulated learning · Intervention · Systematic literature review

✉ Sami Heikkinen
sami.heikkinen@lab.fi

¹ School of Computing, The University of Eastern Finland, Yliopistonkatu 2, 80100 Joensuu, Finland

² LAB University of Applied Sciences, Mukkulankatu 19, 15210 Lahti, Finland

³ Department of Educational Sciences and Teacher Education, The University of Oulu, P.O. Box 2000, 90014 Oulu, Finland

1 Introduction

Self-regulated learning (SRL) is often described as a process that includes the setting of goals, the monitoring of progress, and the regulating of learning (Järvelä et al., 2019; Pintrich, 2000). Empirical studies have continuously demonstrated that students do not use efficient self-regulating strategies when learning in virtual environments (Akyildiz & Kaya, 2021; Azevedo et al., 2019; Broadbent & Poon, 2015; Pedrotti & Nistor, 2019). Therefore, learners may not achieve the expected learning outcome (Lee et al., 2019; Shyr & Chen, 2018; van Alten et al., 2020). This situation has motivated both teachers and researchers to design ways to support it.

Panadero (2017) reviewed six different SRL models and concluded that a common feature of the models is the division of SRL into three phases: planning, performance, and reflection. This three-phase model was initially introduced by (Puustinen & Pulkkinen, 2001). Although the SRL models reviewed by Panadero share commonalities, they vary depending on their focus. For example, Boekaerts and Corno's (2005) model focuses on motivational aspects of regulated learning, while Efklides's (2011) model focuses on metacognitive aspects of regulated learning, such as metacognitive experiences of students' learning regulation. Winne and Hadwin (1998) built their model based on information processing theory and focused on the effective use of cognitive strategies to enhance learning. Hadwin et al.'s (2011) model not only considers individual regulation, but also the ways regulation of learning can be shared between group members. Finally, the models of Pintrich (2000) and Zimmerman (2000) can be described as general models of regulated learning since they consider motivation and metacognition to be central to SRL.

SRL phases comprise several subprocesses that differ by model but are still conceptually close. For instance, Boekaerts and Corno's (2005) model includes identification, interpretation, primary and secondary appraisal, and goal setting in the planning phase of SRL. Efklides's (2011) model includes task presentation in this phase. Hadwin et al. (2011) have claimed that planning is included in this phase. Pintrich (2000) has stated forethought, planning, and activation belonging to this phase. Winne and Hadwin (1998) have allocated task definition, goal setting, and planning to this phase. Finally, Zimmerman (2000) has located forethought (task analysis and self-motivation) in the planning phase of SRL. These subprocesses identified by the researchers are actions learners take at the early phase of the learning process, and they involve estimating how the learning task can be accomplished.

After planning, the learner continues to the performance phase and implements their planned actions. Depending on the model used, the subprocesses of this phase may include goal-striving (Boekaerts & Corno, 2005), cognitive processing, performance (Efklides, 2011), monitoring, control (Hadwin et al., 2011; Pintrich, 2000), applying tactics and strategies (Winne & Hadwin, 1998), and performance (with self-control and self-observation Zimmerman, 2000). After the performance phase, it is time to reflect on the learning process. Within

this reflection phase, subprocesses may include performance feedback (Boekaerts & Corno, 2005), regulating (Hadwin et al., 2011; Pintrich, 2000; Winne & Hadwin, 1998), reaction and reflection (Pintrich, 2000), adapting metacognition (Winne & Hadwin, 1998), and self-reflection (including self-judgement and self-reaction; Zimmerman, 2000). According to Efklides's (2011) model, the reflection phase includes no explicit subprocesses.

Many studies have suggested that SRL could be enhanced by interventions (Ceron et al., 2021; Lee et al., 2019; Wong et al., 2019a). Interventions can assist at-risk learners in improving their learning performance (Espinoza & Genna, 2021) or can help reduce undesirable learning behaviour (Lodge et al., 2019). Furthermore, interventions can be used to increase students' mastery goals (Zheng et al., 2020), self-efficacy (Samuel & Warner, 2021), and perceived value of a course (Alamri et al., 2020). At the same time, this intervention has been found to decrease test anxiety among students (Putwain & von der Embse, 2021). However, numerous SRL subprocesses need different kinds of interventions performed at the right time of the learning process. These needs set the requirements for methods used to create effective interventions.

The rest of this paper follows the following organization. In the second section, the way how LA is applied to enhance SRL is explained. The previous reviews focusing on the themes of this review are explained in the third section. This is followed by the fourth section, where we explain the method used to conduct this review. In the fifth section, we report the results of our study, which is followed by the sixth section discussing the results. Finally, the last section concludes this article.

2 Learning analytics as a tool to support SRL

Learning analytics (LA) is “the measurement, collection, analysis, and reporting of data about learners and their contexts for the purposes of understanding and optimising learning and the environments in which it occurs” (Siemens, 2013 p. 1382), whereas LA methods are referred in this paper as a set of methods used to support the measuring, collecting, analysing and reporting the data. The reporting of the data is mediated to learners in different means which are referred as channels.

LA can be used to analyse and report learner performance and to highlight the parts of learning where SRL processes could be improved (Lodge et al., 2019; Winne, 2022). LA for SRL has two elements, calculation and recommendation (Winne, 2017). These two elements are built within the LA model's four stages: data generation, tracking, analysis, and action (Romero & Ventura, 2020). Data is generated when the learner does different activities in an online learning environment. The trace data from the activities can be used as an input for calculation, analysis, and tracking of the learner's process. Next, the action stage is when the instructor executes interventions to support learners' activities, such as SRL subprocesses. These LA interventions support changes a learner can take in their learning process to achieve better outcomes (Ifenthaler & Yau, 2020).

LA intervention is a surrounding frame of activity used to mediate data and reports created with the help of analytics (Wise, 2014). LA dashboard is a typical channel used for LA intervention (Jivet et al., 2018) and according to Sønderslund et al. (2019) most interventions are aimed to identify at-risk students (Shafiq et al., 2022). Sønderslund et al. (2019) have stated that early LA methods relied on fixed factors and could populate predictions only within limited settings, whereas contemporary dynamic models can be used in more complex learning contexts. However, Rienties et al. (2017) have argued that implementing an intervention remains the challenging part of the LA cycle, and the design of intervention should be well established to succeed in its endeavour to support learning (Sedrakyan et al., 2020; Wise, 2014).

3 Previous literature reviews

A number of literature reviews have already mapped the state of the art in research on SRL, LA, and interventions. Bodily and Verbert (2017) studied the LA dashboards and recommender systems used to intervene in students' learning processes. However, this study did not include the perspective of SRL. Another study focusing on LA dashboards was conducted by Matcha et al. (2020) who studied the LA dashboard interventions aimed to support SRL. This study did not include the different channels besides the LA dashboards. Valle et al. (2021) have studied learner-facing LA dashboards and found out that the intended outcomes of LA dashboards do not meet with their evaluation in terms of effective measures.

The literature review by Pérez-Álvarez et al. (2018) explored tools supporting SRL in online environments. Their study found that most of the tools were used in higher education or general education. The most common means of supporting SRL the researchers observed was visualisation, and the most used indicators were either action-related or content-related. The most supported SRL strategies were goal setting, self-evaluation, help-seeking, and organisation. The studies examined by the researchers rarely explained the relationships between SRL strategies and tools. Their review found that studies measure the value of the tools in terms of usability (easiness of usage), usefulness (appropriateness of use), user satisfaction (acceptance by stakeholders), and learning outcomes.

Wong et al., (2019a) reviewed 35 records supporting SRL in online learning environments and MOOCs. Their review studied the effectiveness of approaches to support SRL strategies and whether human factors influenced the approaches. The study did not include LA. The most used intervention types to support SRL identified by the researchers were prompts (14 studies), integrated support systems (10 studies), and prompts with feedback (four studies). The rest of the reviewed studies used different, rarely used approaches. Despite the investigation, how well different approaches support SRL remained unclear.

Sønderslund et al. (2019) explored the efficacy of LA interventions in higher education but did not include the perspective of SRL within its scope. Specifically, they reviewed studies examining LA interventions' efficacy for student retention and academic success. The results from the research they found were promising,

as the studies identified that LA interventions increased grades and led to higher retention. However, the researchers found only few studies critically assessing the effectiveness of LA interventions for student success and retention.

A literature review by Viberg et al. (2020) focused on studies of SRL and LA in online learning environments. The researchers examined the current state of LA applications and how it measures and supports students' SRL in online learning. Their results showed that most studies focused on the forethought and performance phase of SRL and focused less on the reflection phase. Viberg et al. (2020) concluded that LA was not used to support SRL in their sample of records. They argued that there was little evidence of LA improving learning outcomes or improving support for learning and teaching. Instead of supporting SRL, they found that LA research was focused on measuring SRL. Viberg et al. (2020) thus recommended LA support mechanisms be exploited to foster student SRL within online learning.

Araka et al. (2020) examined the research trends regarding measurement and intervention tools for supporting SRL in online learning environments between 2008 and 2018. The review findings indicated that current tools simultaneously measure and implement interventions to support SRL. Araka et al. (2020) identified LA as one way to measure the efficiency of interventions. However, the LA methods in use were not extensively explored, and this perspective remains unstudied.

The previous literature reviews have examined LA, SRL, and interventions, but none of the studies connected the three concepts. Researchers have been focusing on tools to support SRL (Pérez-Álvarez et al., 2018) and learning analytics interventions in online education (Sønderlund et al., 2019). Still, combination of these elements remains unstudied. This study connects these concepts for the first time. First, it is important to examine how LA has been used to identify the need for SRL support. Then, the channels to mediate the intervention has to be selected to target the phases of SRL cycle. Finally, the efficiency and the impact of the intervention have to be evaluated to find out how the selection of LA methods and channels for LA interventions foster SRL. This will help the learners to achieve the anticipated learning outcomes (Azevedo & Hadwin, 2005; Lee et al., 2019; Shyr & Chen, 2018; van Alten et al., 2020). The research questions for this research are the following:

RQ1: What learning analytics methods have been applied to identify the needs for SRL support?

RQ2: What channels have been applied in LA interventions to foster SRL?

RQ3: Which phases of the SRL cycle were targeted by LA interventions?

RQ4: How do studies evaluate the SRL support efficiency and impact, and what kind of results were achieved?

4 Method

This research followed the PRISMA 2020 statement. PRISMA was chosen because it is widely used and suitable for studies reviewing educational interventions. The PRISMA 2020 is also eligible for systematic reviews including qualitative and

Table 1 Systematic literature review search terms

Topic and cluster	Search terms
self-regulated learning	“self-regulat*” OR “self regulat*” AND
learning analytics	“educational data mining” OR “learning analytics”

quantitative methods, thus allowing a meta-analysis of effect estimates and their variances when they are available (Page et al., 2021.)

We selected a list of keywords (Table 1) to reduce the possibility of reviewer bias. While doing preliminary searches, we found that only a few studies used the exact term *LA intervention* or *intervention* to identify the supporting acts for SRL. Instead, these acts were described using various terms or keywords when presented as a study’s experiment. Therefore, we decided to perform a broad search and include all records about LA and SRL. After the records were collected, they were checked manually for the evidence of an intervention since it is hard to find a reasonable keyword or combination of keywords that would ensure that all intervention related records would be captured. To reduce the author bias, we discussed the records in meetings to the unanimous opinions on which records to be included in the review.

The keywords used to find relevant research identified records focusing on SRL and learning analytics. There are two ways to write self-regulation (i.e., with or without a hyphen), and both ways were used to identify the records including this topic. “Self-regulat*” was used to include all papers that have the keyword “self-regulation”; “self-regulating”; “self-regulated”; “self-regulate”; “self-regulates”; and so forth. The term “self regulat*” was used to include variants of the previous terms written without a hyphen. For learning analytics, we decided to include “educational data mining” and “learning analytics” due to the close relationship between the two often overlapping concepts (see Romero & Ventura, 2020).

The inclusion criteria for the eligible record were the following:

1. The manuscript reported a learning analytics analysis or an intervention hypothesised to support or enhance learners’ self-regulated learning.
2. The full text was available.
3. The record was written in English.
4. The record was a peer-reviewed empirical paper.
5. The record was published prior to the end of 2021.

We conducted comprehensive searches on Scopus (390 documents) and Web of Science (164 documents) in May 2022. Additionally, we used Google Scholar to search for records published by the *Journal of Learning Analytics* (142), the *Journal of Educational Data Mining* (35), and the International Conference on Educational Data Mining (22). The *Journal of Learning Analytics*, *Journal of Educational Data Mining*, and International Conference of Educational Data Mining are not fully indexed in the Scopus or Web of Science databases. Therefore, we checked these publication channels manually because they have been

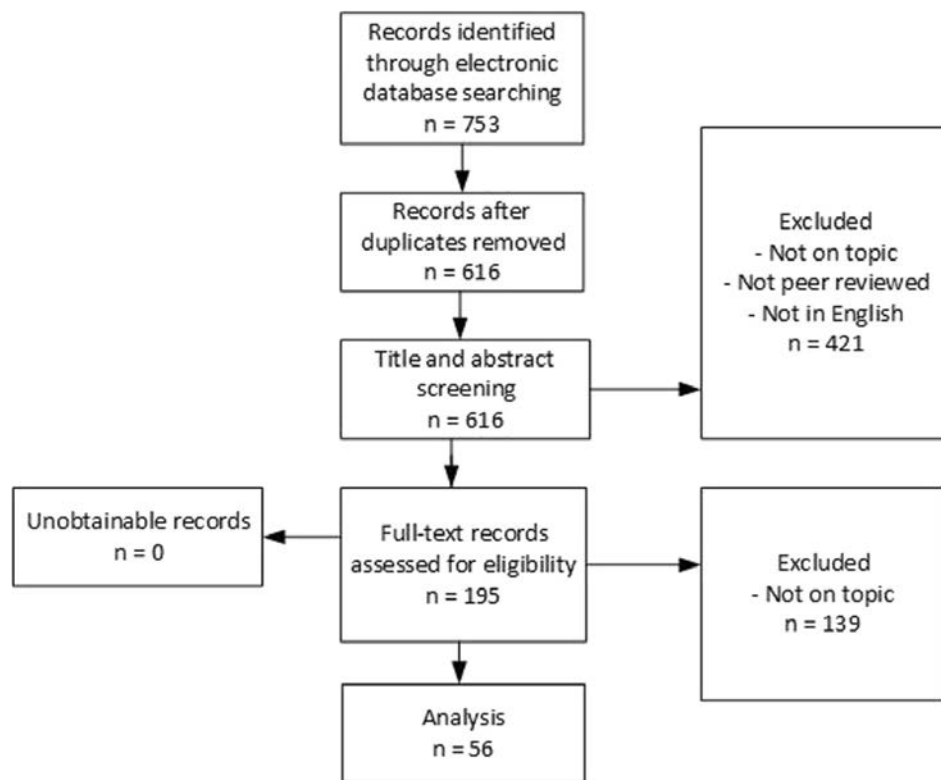


Fig. 1 Search strategy and process of literature review

identified as having the most specific focus on LA and EDM (Romero & Ventura, 2020), and we expected to find studies focusing on the topic of the review.

We carried out the selection strategy and the literature review process (Fig. 1) following the inclusion criteria. The records obtained from the search were reviewed in three phases. In every phase, authors had meetings to discuss the records to find the unanimous opinions of whether to include or exclude the records. The first phase of examination focused on the title and the abstract of the studies, and the first author of the current review excluded records not meeting the inclusion criteria based on these elements. In the second phase, the first author read the full texts. If a record did not include the three elements of the research question, it was excluded. There were 47 records left (see Appendix Table 7) after phase two. For the last phase, the first author examined each record in detail and recorded the information from their studies. The first and the second authors met repeatedly to agree on a coding scheme. Then the first author and the second author coded papers independently, followed by a meeting to reach an agreement regarding the disagreements and after consensus, the first author completed the coding. Coding was used to describe the details of the studies and enable mapping of the LA methods, the interventions applied, the phases of SRL that were focused on, the evaluation methods used to assess the outcome of LA interventions, and the impact of a study's intervention methods on learning process or learning outcome. Table 2 shows the codification summary of the research questions.

Table 2 Codification by research question (1 – 4)

RQ1: What learning analytics methods have been applied to identify the needs for SRL support?	
Coding	Categories
LA methods	Causal mining, clustering, discovery with models, distillation of data for human judgment, knowledge tracing, non-negative matrix factorisation, outlier detection, prediction, process mining, recommendation, relationship mining, statistics, social network analysis, text mining, visualization
RQ2: What channels have been applied in LA interventions to foster SRL?	
Coding	Categories
Channels used for intervention	Learning analytics dashboard (LAD), embedded, prompts, phone call, text message, e-mail
RQ3: Which phases of the SRL cycle were targeted by LA interventions?	
Coding	Categories
SRL phase	Planning, performance, reflection
RQ4: How do studies evaluate the SRL support efficiency and impact, and what kind of results were achieved?	
Coding	Categories
Form of impact	Improved self-regulation, better learning outcomes, engagement with course materials, improved course completion, improved student retention

We synthesised the results of previous studies to answer research questions and reported the findings. There are diverse LA methods that have been used to measure, collect, analyse, and report the data about learners to understand and optimise learning (Siemens, 2013). The categorisation of each record's methods in this literature review followed the categories presented by Romero and Ventura (2020). They have suggested division of the methods into the following categories: causal mining (CM), clustering (C), discovery with models (DM), distillation of data for human judgement (DD), knowledge tracing (KT), non-negative matrix factorisation (NMF), outlier detection (OD), prediction (P), process mining (PM), recommendation (R), relationship mining (RM), statistics (S), social network analysis (SNA), text mining (TM), and visualisation (V). Almost all the methods of the records were identified and could be categorised using the definitions of Romero and Ventura. Two studies (Wise et al., 2014, 2016) applied methods that did not fit into these categories. Within these studies there are integrated analytics used to express learners the expected quantity, quality, and timing of discussions (Wise et al., 2014) and a conceptualization including LA as a part of SRL process (Wise et al., 2016). Therefore, we allocated these records into the category of other methods (O).

We coded the SRL subprocesses to phases: planning, performance, and reflection following the division presented by Panadero (2017). This was done by examining the full texts using “subprocesses” as a search term to identify which subprocesses had been investigated in the studies and organised the subprocesses under the three phases of SRL.

The impact of LA interventions can be measured in several ways, including retention, course completion, and dropout rates. LA Interventions that promote a better retention rate improve the learner's continuous study in a course (e.g., Freitas et al., 2015), whereas LA interventions that create greater completion rates reflect learners' successful completion of individual courses and thesis processes (e.g., Nouri et al., 2019). A reduced dropout rate from an LA intervention results in fewer learners failing to finish coursework (e.g., Cambruzzi et al., 2015). Other positive effects can also include better engagement with a course (Chen et al., 2020), usage of learning management system (LMS) resources (e.g., Jovanović et al., 2021), and eventually improved course grades (e.g., Wong, 2017). Finally, the outcome from a LA intervention can be seen in form of enhanced SRL processes (e.g., Roll & Winne, 2015). For the analysis, we considered these aspects as measurable impacts of LA interventions.

The reviewed studies used different effect sizes to measure the impact of their LA interventions. Due to insufficient number of papers with similar research settings, we were not able to execute proper meta-analysis. Instead, we decided to use both qualitative and quantitative studies to illustrate what kind of outcomes are achieved in the reviewed studies. To prevent bias due to differences in how reviewed records analysed different tasks, procedures, methods, and effect sizes, we did not convert the effect sizes to create an equal scale (e.g., Cohen's d Cohen, 1992). Instead, we reported the effect sizes as presented in the original records. To enable the comparison between different effect sizes, we used the interpretation guidelines given by (Maher et al., 2013) to synthesise the results. The results of this approach are presented in Table 3. There are other effect sizes (hazard ratio, β , F , and X^2) used in the studies, but they were not comparable with the interpretation approach. These effect sizes are displayed in the results without interpretations.

5 Results

There were 56 studies included in the final analysis. The first studies about designing LA interventions were conducted by (Bouchet et al., 2013) and (Inventado et al., 2013). In their work, the researchers of these studies formed clusters of students, characterised the students belonging to different clusters (Bouchet et al., 2013), and identified effective long-term learning behaviours (Inventado et al., 2013). Interest in the topic has increased since 2013, and nine studies from 2021 focusing on the topic were found (Fig. 2). These more recent studies focus on implementing and evaluating the tools designed for LA interventions. The reviewed studies primarily examined

Table 3 Effect sizes and their interpretation

	Effect sizes			
	Small	Medium	Large	Very large
Cohen's d	0.20	0.50	0.80	1.30
correlation	0.10	0.30	0.50	0.70
η^2	0.01	0.06	0.14	-

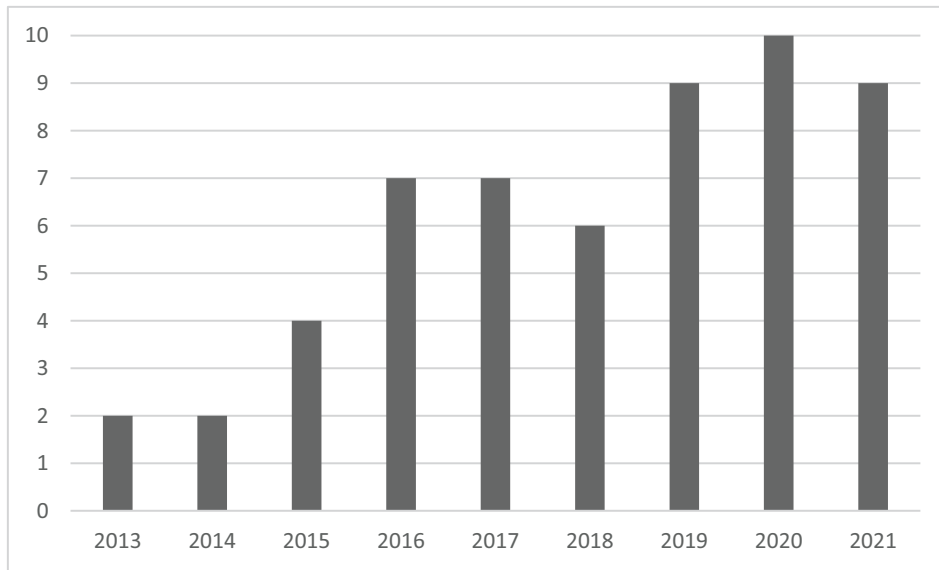


Fig. 2 Number of records/year

higher education settings (46 studies; 82% of studies) with undergraduate or graduate students. Less frequently, the studies concerned doctoral students (two studies; 4%), workplaces (three studies; 5%), or public schools (one study; 2%). In (Jivet et al., 2021) there were higher education degree students and students in professional development process included. Four studies did not report the population that they researched. The sizes of the samples involved in the studies varied. Some studies had less than 10 participants (Cha & Park, 2019; Inventado et al., 2013; Wise et al., 2014, 2016), whereas the largest population was made of 33,726 learners (Davis et al., 2017). The mean population of the reviewed studies was 1,398 participants and the median was 102 participants. Most studies used single courses as their learning context (27 studies; 48%), but some studies conducted their LA intervention across nine (Hardebolle et al., 2020) or ten (Haynes, 2020; Kia et al., 2020) courses. This trend of aiming to generalise the LA interventions is relatively recent.

There were different phases and perspectives in the development of the LA interventions. Some studies explained the early phases of their LA intervention design from a technological point of view (Frey et al., 2016; Manso-Vázquez & Llamas-Nistal, 2015; Siadaty et al., 2016b; Winne et al., 2019). Other studies examined the next phases with the perspective of user's experience on the interface of the intended LA intervention (Cha & Park, 2019; Rohloff et al., 2019; Sonnenberg & Bannert, 2016), and the rest of the studies tested the effectiveness of the LA intervention in action.

5.1 RQ1: What learning analytics methods have been applied to identify the needs for SRL support?

While some of the studies ($n=14$; 25%) focused on supporting SRL with one LA method, it was more common to see a study use two ($n=16$; 29%) or three ($n=18$;

Table 4 Summary of methods used in reviewed studies

Learning analytics methods	Number of studies
Visualisation	32 (57%)
Statistics	32 (57%)
Recommendation	16 (29%)
Prediction	12 (21%)
Discovery with models	8 (14%)
Knowledge tracing	7 (13%)
Clustering	7 (13%)
Process mining	6 (11%)
Relationship mining	6 (11%)
Non-negative matrix factorisation	2 (4%)
Outlier detection	2 (4%)
Other methods	2 (4%)
Causal mining	1 (2%)
Distillation of data for human judgement	1 (2%)
Social network analysis	1 (2%)
Text mining	1 (2%)

32%) methods. Some scholars even used four (Jivet et al., 2020; Matcha et al., 2019; Russell et al., 2020) or five methods (Afzaal et al., 2021; Sonnenberg & Bannert, 2015, 2016, 2019; Yu et al., 2018) to achieve their goals. The methods used in the reviewed studies are summarised in Table 4. The first column shows the learning analytics methods. As multiple methods can be applied, the overall number of learning analytics methods surpassed the number of reviewed studies. The number of studies using a method is displayed in the second column.

Statistical approach was often used to provide content for visualisations ($n = 19$; 34%). Four studies used a combination of the three most used methods (Balavijendran & Burnie, 2018; Jivet et al., 2020; Siadaty et al., 2016a; Yu et al., 2018). In these studies, a learner's trace data was analysed using statistical methods and compared to their peer learners. The analysis result was then visualised for the learner and accompanied by recommendations on how to improve their learning processes.

5.2 RQ2: What channels have been applied in LA interventions to foster SRL?

Table 5 presents the different channels used for LA interventions. The first column displays the different channels used for LA interventions and the second column shows the number of studies applying these channels.

An LAD displays an LA intervention using a visual form. Embedded systems may include elements directly connected to the learning materials or assignments. For example, (Bouchet et al., 2013) used a virtual pedagogical assistant to give hints, and (Frey et al., 2016) implemented a system that helped learners rehearse question-asking skills for question-based dialogue. Prompting can provide directions to a learner at different parts of the learning process and help them to improve their learning.

Table 5 Summary of channels used in reviewed studies

Channels	Number of studies
Learning analytics dashboard	27 (48%)
Embedded in LMS	12 (21%)
Prompt	11 (20%)
E-mail	5 (9%)
Foldout	2 (4%)
Phone call	1 (2%)
Text message	1 (2%)

The LA interventions mentioned used automation. However, there were also LA interventions where the amount of individual human touch was high. These LA interventions include e-mails (Lim et al., 2021; Lim et al., 2020; Matcha et al., 2019; Menchaca et al., 2018; Nikolayeva et al., 2020), foldouts (Ott et al., 2015; Sedraz Silva et al., 2018), phone calls, and text messages (Herodotou et al., 2020). E-mail is an excellent tool for providing individual guidance and helping a student develop their learning process (Matcha et al., 2019). Foldouts (Ott et al., 2015; Sedraz Silva et al., 2018) also provide learners with information about their learning process.

5.3 RQ3: Which phases of the SRL cycle were targeted by LA interventions?

It is most common to foster the SRL subprocesses of the performance phase (e.g., giving hints on how to approach assignments). In total, 44 studies (79%) focused on the performance subprocesses. 25 studies (45%) targeted the subprocesses of the planning phase (e.g., setting goals, time management), and 17 studies (30%) targeted subprocesses of the reflection phase (e.g., evaluation learning outcomes and identifying how to improve performance). Five studies (9%) did not specify the subprocesses they focused on, and therefore those studies could not be categorised. Figure 3 presents the division of the studies into different phases.

Notably, there were only nine studies (16%, Bouchet et al., 2013; Hardebolle et al., 2020; Jivet et al., 2021; Lallé et al., 2017; Lim et al., 2021; Nikolayeva et al., 2020; Siadaty et al., 2016b; Wise et al., 2014; Wong et al., 2019b) that focused on the entire SRL cycle. Manganello et al. (2021) applied the all the phases of 4C model, which is used in professional contexts. Connecting only two phases was more common. For example, studies typically connected the performance with either the planning phase (12 studies; 21%) or with the reflection phase (four studies; 7%). Connecting the planning and reflection phases was only done by (Lu et al., 2017).

5.4 RQ4: How do studies evaluate the SRL support efficiency and impact, and what kind of results were achieved?

The studies evaluated the efficiency of an LA intervention in various ways. Studies focused on their LA intervention design tended to use qualitative methods, such as

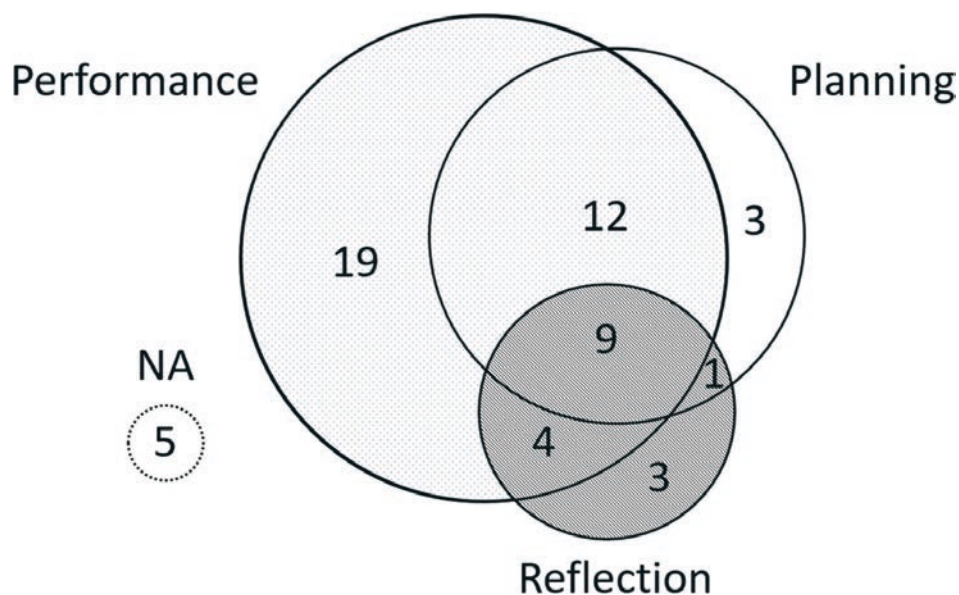


Fig. 3 Division of LA interventions on SRL phases

interviews, the think-aloud method, and focus groups, to verify the usability, usefulness, and learners' perception of an LA intervention. Interviews (Cha & Park, 2019; Corrin & de Barba, 2014; Wise et al., 2014, 2016) were used to discover how learners perceived the functionalities and features of LA interventions and to find approaches that learners would be willing to use in their learning processes. Sonnenberg and Bannert (2016, 2019) used the think-aloud method to gain insight into these learner perceptions and to identify how useful a design was for a learner. In other cases, some researchers used focus groups (Balavijendran & Burnie, 2018; Lim et al., 2021; Lim et al., 2020), observations of facial expressions (Bouchet et al., 2013), eye-tracking (Bouchet et al., 2013; Lallé et al., 2017; Munshi & Biswas, 2019), and physiological sensors (Lallé et al., 2017) to evaluate the effects of an LA intervention during its early developmental phase.

Quantitative methods were used by researchers to evaluate the efficiency of an LA intervention during the execution phase. 34 studies (61%) used trace data for this purpose, making it the most used method. Pre- and post-intervention surveys were popular evaluation methods as well. There were 23 studies (41%) that used surveys, either alone or with trace data, to evaluate the efficiency of a LA intervention. The actual outcomes of participants' learning processes were evaluated in terms of grade distribution (Afzaal et al., 2021; Cody et al., 2020; Kia et al., 2020; Lim et al., 2021; Lu et al., 2017; Manganello et al., 2021; Ott et al., 2015; Sedrakyan & Snoeck, 2017; van Horne et al., 2018) and course completion (Davis et al., 2017; Herodotou et al., 2020).

26 studies (46%) reported a positive, measurable impact on learning from their LA intervention. In Table 6, the impact of the LA interventions and the measurements of the impacts are displayed. In the first column, the types of impact are denoted. The second

Table 6 LA intervention impacts and the measurements of impact

The type of impact	Study	Impacted variable	Effect size	Interpretation
Improved self-regulation	Aguilar et al., 2021	Improved SRL	NA	
	Balavijendran & Burnie, 2018	Perceived improvement in SRL	NA	
	Chen et al., 2020	Regulation of cognition	$\beta=0.393$	
	Corrin & De Barba, 2014	Perceived better progress during the performance phase	NA	
	Domínguez et al., 2021	Self-assessment	NA	
	Lim et al., 2020	Goal setting and procrastination	NA	
	Lim et al., 2021	Improved SRL	NA	
	Lim et al., 2021	Improves SRL	$\eta_p^2=0.22$	Large
	Matcha et al., 2019	High engagement strategy	NA	
	Siadaty et al., 2016c	Task analysis	$r=0.667$	Large
		Goal setting	$r=0.778$	Very large
		Personal planning	$r=0.740$	Very large
		Working on task	$r=0.781$	Very large
		Applying strategy changes	$r=0.745$	Very large
		Evaluation	$r=0.685$	Large
		Reflection	$r=0.682$	Large
	Sonnenberg & Bannert, 2015	Metacognitive learning activities	$d=0.44$	Small
	Sonnenberg & Bannert, 2016	Metacognitive learning activities	$d=2.00$	Very large
	Sonnenberg & Bannert, 2019	Task analysis	$\eta_p^2=0.082$	Medium
		Evaluation	$\eta_p^2=0.063$	Medium

Table 6 (continued)

The type of impact	Study	Impacted variable	Effect size	Interpretation
Better learning outcomes	Afzaal et al., 2021	Improvement in grades	$r = 0.86$	Very large
	Cody et al., 2020	Faster performance	$F = 5.24$	
	Davis et al., 2016	Improvement in grades	NA	
	Lim et al., 2021	Improvement in grades	$d = 0.47$	Small
	Lu et al., 2017	Improvement in grades	$F = 19.71$	
	Manganello et al., 2021	Improvement in grades	$d = 1.10$	Large
	Menchaca et al., 2018	Defining targets Task definition Communication and collaboration	$d = 0.90$ $d = 0.78$ $d = 0.35$	Large Medium Small
	Schumacher & Ifenthaler 2021	Declarative knowledge	NA	
	Van Horne et al., 2018	Improvement in grades	NA	
	Davis et al., 2016	Timely submissions and number of submissions	NA	
Better engagement with course materials	Nikolayeva et al., 2020	More completed quizzes	$d = 0.31$	Small
	Siadaty et al., 2016a	Perceived learning paths useful	NA	
	Wong et al., 2019b	Increased interaction with activities	NA	
Improved course completion	Davis et al., 2017	Improved completion rate	$\chi^2 = 5.87$	
	Russell et al., 2020	Lower dropout rate	$HR = 0.74$	
Improved student retention	Herodotou et al., 2020	Improved course completion	NA	

Table 6 (continued)

The type of impact	Study	Impacted variable	Effect size	Interpretation
Not promising	Bouchet et al., 2013	Effect on SRL	NA	
	Haynes, 2020	Perceived usefulness of the LA intervention	NA	
	Kia et al., 2020	The connection between SRL and academic achievement	$\chi^2 = 2.43$	
	Li et al., 2017a	Improvement on dropout rates	NA	
	Li et al., 2017b	Learning effectiveness	NA	
	Ott et al., 2015	Attendance Submission rates Exam performance	NA	
	Tabuenca et al., 2021	Time management	NA	

column identifies the studies, and the third column shows the impacted variable. The fourth column presents the effect sizes measured in the studies, and the last column interprets effect size following the principles introduced in Table 2.

In 13 studies (23%), the impact of an LA intervention was seen in the form of improved self-regulation. Nine (16%) studies achieved better learning outcomes. LA interventions resulted in better engagement with course materials in four (7%) studies. Davis et al. (2017) and Russell et al. (2020) found improved course completion (4%), and Herodotou et al., (2020; 2%) reported an improvement in student retention in response to their respective LA interventions. Seven studies (13%; Bouchet et al., 2013; Haynes, 2020; Kia et al., 2020; Li et al., 2017a, b; Ott et al., 2015; Tabuenca et al., 2021) claimed the outcomes they observed were not promising or stated there were significant limitations for practical implications. Four (7%) studies (Frey et al., 2016; Manso-Vázquez & Llamas-Nistal, 2015; Munshi & Biswas, 2019; Siadaty et al., 2016b; Winne et al., 2019) were not able to evaluate the impact of their LA intervention (e.g., due to the early stage of development of the LA intervention).

The reporting of effect sizes varied among studies. There were in total 16 (29%) studies where effect sizes were reported. Very large effect sizes were found by Afzaal et al. (2021) regarding the grades, by Sonnenberg and Bannert (2016) the number of metacognitive learning activities ($d=2.0$) and by Siadaty et al. (2016b) for the subprocesses of goal setting ($r=0.778$), personal planning ($r=0.740$), working on task ($r=0.781$), and applying strategy changes ($r=0.745$). A Large effect sizes were found by Manganello et al. (2021) on improvement in grades ($d=1.10$), Lim et al. (2021) on SRL improvement ($\eta_p^2=0.22$), Siadaty et al. (2016b) on task analysis ($r=0.667$), evaluation ($r=0.685$), reflection ($r=0.682$), and by Menchaca et al. (2018) on defining targets ($d=0.90$). The LA intervention of Sonnenberg and Bannert (2019) had a medium effect on evaluation ($\eta_p^2=0.063$) and task analysis ($\eta_p^2=0.082$). Small effect sizes were reported for grades ($d=0.47$, L. A. Lim et al., 2021), prompts ($d=0.44$, Sonnenberg & Bannert, 2015), communication and collaboration ($d=0.35$, Menchaca et al., 2018), and the number of quizzes completed ($d=0.31$, Nikolayeva et al., 2020).

6 Discussion

The goal of this paper was to systematically review how LA has been applied to enhance SRL in. Many of the reviewed studies achieved promising results. Many others, however, were not able to show an impact on learning. Despite the promising results, it is still unclear whether there are features that could predict the success of a LA intervention. Some studies report how learners perceive the impact of intervention, whereas some studies report the changes in the students' learning process or learning outcomes. The previous studies have shown, that even though students' self-evaluation would suggest the improvement in SRL, the impact seen in trace data, doesn't necessarily show similar improvement due to e.g., response styles or confidence measures (Tempelaar et al., 2020).

Almost two-thirds of the studies reviewed applied two or three LA methods. Nevertheless, the increase in the number of applied LA methods does not seem to improve the effectiveness of a LA intervention. For instance, (Afzaal et al., 2021; Sonnenberg &

Bannert, 2015, 2016, 2019; Yu et al., 2018) applied five LA methods within all their studies, and the effectiveness of their LA interventions ranged from small to very large. If the number of LA methods is not the key for a successful LA intervention, then perhaps the combination of methods could be. Four studies used a combination of the three common LA methods: statistics, visualisation, and recommendation (Balavijendran & Burnie, 2018; Jivet et al., 2020; Siadaty et al., 2016a; Yu et al., 2018). A positive impact was reported only by Balavijendran and Burnie (2018), whose students reported that LA interventions improved their learning. In contrast, the studies of (Jivet et al., 2020; Siadaty et al., 2016a; Yu et al., 2018) were not able to find an impact on learning, even though they included the four stages of the LA cycle (Khalil & Ebner, 2015; Siemens, 2013) and used a combination of methods.

LAD was the most used LA intervention among the reviewed studies, though only 43% of the LAD interventions positively impacted learning. Success rates suggest that the implementation of LAD is not sufficient to support SRL. Instead, how LAD should be used seems to be an unknown element affecting SRL. This can be due to the absence of support in terms of dashboard design from learning sciences with regard to core concepts of learning processes and feedback mechanisms (Sedrakyan et al., 2020). The features and characteristics of LADs differ (Bodily & Verbert, 2017) and the effective elements should be considered as principles of LAD design. A closer look should also be put on the different aspects of e.g., demographics of the sample, duration of the study and the studied subjects to find underlying variables affecting the unknown elements of LAD success.

Only nine studies focused on the entire SRL cycle. These findings are similar to those of Viberg et al. (2020), who found that the studies they reviewed focused less on the reflection phase. Araka et al. (2020) identified LA as one way to measure the effectiveness of LA interventions but found that the methods used for this purpose remained unexplored. Furthermore, Viberg et al. (2020) has argued that there is little evidence of LA improving learning outcomes or learning support. In this current study, only four out of the nine studies reported positive impacts on SRL (Lim et al., 2021; Nikolayeva et al., 2020; Siadaty et al., 2016b; Wong et al., 2019b). Channels used for LA intervention and the LA methods differed studies. Nikolayeva et al. (2020) and Lim et al. (2021) performed their LA intervention using e-mail. Siadaty et al. (2016b) used LAD to give recommendations to learners. Nikolayeva et al., Siadaty et al., and Lim et al. applied statistical methods to support their LA interventions. Wong et al., (2019b) used a discovery with methods accompanied by relationship mining and visualisation to prompt learners. The common feature for all four studies can be seen in learning outcomes: all interventions increase the learning actions e.g., working on task (Siadaty et al., 2016b), completed quizzes (Nikolayeva et al., 2020), interaction with activities (Wong et al., 2019b), and improved grades (L. A. Lim et al., 2021). Lim et al. (2021) was the only study that showed an impact in improved SRL. This is a promising result but requires further studies to verify results in different study contexts and with larger sample sizes. The results suggest that intervention focusing on planning and reflection can be seen in the outcomes of the performance phase. However, the limited number of studies focusing on the entire SRL cycle, sets a need to design more interventions covering the entire SRL cycle and to execute comparative research to find the most effective interventions as a basis for intervention design principles.

The reviewed studies reported the impact of their LA interventions on learning in various ways. Both qualitative and quantitative methods were used at different stages of an LA intervention's development. While 26 studies (46%) reported a positive impact on learning, only 14 studies reported enough details to make the results interpretable and enable proper comparison. A similar problem was found by S nderlund et al. (2019). In their review of research, only three studies critically assessed the effectiveness of LA interventions for student success and retention. The research of (S nderlund et al., 2019) explored the efficacy of LA interventions in higher education, and they evaluated LA interventions from the perspectives of student retention and academic success, which is a rather narrow perspective for evaluating the impact of a LA intervention on learning. The current study broadened the framework of impact by adding the aspects of improved SRL processes and engagement with the course materials. This helps the future studies to assess the impact of LA interventions.

Different designs of LA interventions have different impacts on learning. For example, the LA interventions by (Siadaty et al., 2016b) and (Sonnenberg & Bannert, 2019) were aimed at supporting the subprocess of task analysis in SRL. The former study found their LA intervention had a large impact on learning, whereas the latter found it had a medium impact. It is also interesting to note that an initial LA intervention can be further developed, thus leading to an improved impact. This kind of development can be seen in the consecutive studies by (Sonnenberg & Bannert, 2015, 2016), where the initial impact was improved from small to very large across their studies.

We argue that a successful LA intervention design requires the LA cycle to be supplemented with a sound pedagogical approach to provide feedback for both cognitive and behavioural processes (Sedrakyan et al., 2020). The initial LA intervention must be developed iteratively, as was done by Sonnenberg and Bannert (2015, 2016) to test different approaches and functionalities in order to find the most effective ways to enhance SRL. In SRL, regulation can focus on the cognitive, emotional, and motivational aspects of the learning process. In addition, regulation can also focus on different phases, such as planning, performance, or reflection. This is to say that it would not be reasonable to expect a LA intervention to support all phases simultaneously unless the subprocesses included in the different phases are considered during the design of an intervention. Also, LA interventions are typically conducted for a short period, despite the fact that only long-term LA interventions can potentially contribute to the development of SRL skills. The longevity of training and support must be considered during the design of a LA intervention to enable the students to master SRL skills (Zimmerman, 2000). Furthermore, research should explore the personalized ways of studying and supporting self-regulation i.e., idiographic methods. Idiographic methods rely on several data points collected from individual students to create precise insights based on student's own data e.g., (Saqr & L pez-Pernas, 2021). With the rise of multimodal learning analytics, research can tap into the fine grained within-person process and emotions and therefore, deliver highly personalized just-in-time insights and support (Malmberg et al., 2022; T rm nen et al., 2022).

6.1 Validity threats

The validity assessment includes the aspects of internal, external, construct, and conclusion validity (Zhou et al., 2016). In this review, we reduced the risk of validity threats by following the recommended mitigation actions proposed by (Ampatzoglou et al., 2019). Inappropriate or incomplete search terms can be a threat to construct validity and internal validity. LA interventions can be done in various ways, and the concept of LA intervention might not be used in some records. Within this study, *LA intervention* was not used as a keyword. Instead, we identified LA interventions qualitatively. The same applies to SRL. There were records screened focusing on interventions and learning, but SRL was not used as a concept in these records. These records may have been excluded even though the interventions could have been beneficial for learning. Therefore, there might be records that were relevant to this review, but they were not included because they used keywords not identified within this research. Incomplete research information in primary studies causes an external threat to validity. We have been mitigating this risk by identifying the missing information in Table 6. To reduce the risk of threat to conclusion validity, we have removed the duplicate studies. The discussions among authors have been used aiming for unanimous consensus to reduce subjective interpretations thus mitigating to internal validity and conclusion validity.

7 Conclusion

We have reviewed how studies use LA methods to create LA interventions supporting SRL. Based on the review, neither the number, popularity, nor combination of LA methods implemented in a learning environment provide a guarantee for improved SRL. Moreover, none of the LA intervention channels has shown proof of success in supporting SRL. Only seven studies covered all phases of SRL, and only 12 studies assessed the effect sizes of the LA interventions. Out of these studies, only three focused on the entire SRL cycle and showed results that LA interventions positively impacted learning. Furthermore, the LA methods and the LA intervention channels of the three studies varied. Therefore, we cannot say which combination of LA methods and LA intervention channels would ensure the successful support of SRL.

Interpretable and comparable LA intervention results would help scholars and practitioners to find LA interventions that are the most beneficial for SRL. Future research needs to focus on the pedagogical approach and how it may be used to create impactful LA interventions. Second, LA interventions should focus on different phases and processes of regulated learning. Future research should find the most impactful LA interventions capable of supporting different SRL phases and subprocesses. Third, there are a limited number of LA interventions covering all three SRL phases. Future research and LA intervention design should note the aspects of all SRL phases to support holistic learning. Finally, there are limitations related to how the reviewed studies reported their findings. To improve the interpretability and comparability of an LA intervention design, future research needs to put focus on providing comparable data with larger sample sizes to enable the comparison of different kinds and types of LA interventions.

Appendix

Table 7

Table 7 Summary of the reviewed studies

Publication	RQ1: LA Methods used	RQ2: Type of intervention	RQ3: SRL Phases	RQ4: Impact
Afzaal et al., 2021	P, DM, S, R, V	LAD	PER	Better learning outcomes
Aguilar et al., 2021	P, S, V	LAD	PER	Improved self-regulation
Arnold et al., 2017	R, V	LAD	PER	No impact
Balavijendran & Burnie, 2018	R, S, V	LAD	PER + REF	Improved self-regulation
Bouchet et al., 2013	C, KT, S	EMB, PRO	PLA + PER + REF	No impact
Cha & Park, 2019	V	LAD	PER	No impact
Chen et al., 2020	KT, S, V	LAD	PER	Improved self-regulation
Cody et al., 2020	P, R	EMB	PER	Better learning outcomes
Corrin & de Barba, 2014	V	LAD	PER	Improved self-regulation
Davis et al., 2016	OD, S, V	LAD	PLA + PER	Better learning outcomes, engagement with course materials
Davis et al., 2017	OD, S, V	LAD	PER	Improved course completion
Dominguez, 2021	C	EMB	REF	Improved self-assessment
Frey et al., 2016	V	EMB	REF	No impact
Hardebolle et al., 2020	R, V	LAD	PLA + PER + REF	No impact
Haynes, 2020	S, V	LAD	PLA + PER	No impact
Herodotou et al., 2020	P, S	PHONE, TEXT	PLA + PER	Improved student retention
Howell et al., 2018	NMF, S	PRO	PER	No impact
Inventado et al., 2013	DM	N/A	N/A	No impact
Jivet et al., 2020	P, R, S, V	LAD	PLA + PER	No impact
Jivet et al., 2021	P, S, V	LAD	PLA + PER + REF	No impact
Kia et al., 2020	C, RM, V	LAD	PLA + PER	No impact

Table 7 (continued)

Publication	RQ1: LA Methods used	RQ2: Type of intervention	RQ3: SRL Phases	RQ4: Impact
Lallé et al., 2017	KT, S	PRO	PLA + PER + REF	No impact
Li et al., 2017a	DD, S, V	LAD	PER	No impact
Li et al., 2017b	RM	EMB	N/A	No impact
Lim et al., 2020	R, SNA V	E-MAIL	N/A	Improved self-regulation
Lim et al., 2021	C	EMB	PER	Improved self-regulation
Lim et al., 2021	P, R, S	E-MAIL	PLA + PER + REF	Better learning outcomes, improved self-regulation
Lu et al., 2017	S, V	PRO	PLA + REF	Better learning outcomes, improved self-regulation
Manganello et al., 2021	S, V	LAD	NA	Better learning outcomes
Manso-Vázquez & Llamas-Nistal, 2015	KT	EMB	PLA + PER	No impact
Matcha et al., 2019	C, PM, S, V	LAD, E-MAIL	PER	Improved self-regulation
McKenna et al., 2019	R, V	LAD	PER	No impact
Menchaca et al., 2018	S	E-MAIL	PER	Better learning outcomes
Munshi & Biswas, 2019	DM, PM, RM	EMB	N/A	No impact
Naranjo et al., 2019	S, V	LAD	PER	No impact
Nikolayeva et al., 2020	S	E-MAIL	PLA + PER + REF	Engagement with course materials
Ott et al., 2015	P, S, V	FOLDDOUT	PLA + PER	No impact
Rohloff et al., 2019	R, V	LAD	PER + REF	No impact
Russell et al., 2020	DM, P, S, V	LAD	PER	Improved course completion
Schumacher & Ifenthaler, 2021	P, S	PRO	PLA, PER	Improved learning outcomes
Sedrakyan & Snoeck, 2017	CM, PM, S	EMB	PER + REF	No impact
Sedraz Silva et al., 2018	R, S	FOLDDOUT	PER	No impact
Siadaty et al., 2016a	R, S, V	LAD	PLA + PER	Engagement with course materials
Siadaty et al., 2016b	R	LAD	PLA + PER	No impact
Siadaty et al., 2016c	R, S	LAD	PLA + PER + REF	Improved self-regulation

Table 7 (continued)

Publication	RQ1 : LA Methods used	RQ2: Type of intervention	RQ3 : SRL Phases	RQ4: Impact
Sonnenberg & Bannert, 2015	C, DM, KT, PM, RM	PRO	PLA	Improved self-regulation
Sonnenberg & Bannert, 2016	DM, NMF, PM, S, V	PRO	PLA	Improved self-regulation
Sonnenberg & Bannert, 2019	DM, KT, PM, RM, S	PRO	PLA	Improved self-regulation
Tabuenca et al., 2015	C, S, V	LAD	PER	No impact
Tabuenca et al., 2021	S	PRO	REF	No impact
van Home et al., 2018	P, S, V	LAD, EMB	PER + REF	Better learning outcomes
Winne et al., 2019	KT, R	EMB	PLA + PER	No impact
Wise et al., 2016	O	LAD, EMB	PER	No impact
Wise et al., 2014	O	PRO	PLA + PER + REF	No impact
Wong et al., 2019b	DM, RM, V	PRO	PLA + PER + REF	Engagement with course materials
Yu et al., 2018	P, R, S, TM, V	N/A	PLA + PER	No impact

Acknowledgements The research reported in this paper was supported by the LAB University of Applied Sciences' LAB-D program.

Funding Open access funding provided by University of Eastern Finland (UEF) including Kuopio University Hospital. The paper is co-funded by the Academy of Finland for the project TOPEILA, Decision Number 350560 which was received by the second author.

Data availability No new data were created during the study.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Afzaal, M., Nouri, J., Zia, A., Papapetrou, P., Fors, U., Wu, Y., Li, X., & Weegar, R. (2021). Explainable AI for Data-Driven Feedback and Intelligent Action Recommendations to Support Students Self-Regulation. *Frontiers in Artificial Intelligence*, 4. <https://doi.org/10.3389/frai.2021.723447>
- Aguilar, Stephen J., Karabenick, Stuart A., Teasley, Stephanie D., & Baek, Clare. (2021). Associations between learning analytics dashboard exposure and motivation and self-regulated learning. *Computers & Education*, 162, 104085. <https://doi.org/10.1016/j.compedu.2020.104085>
- Akyildiz, S. T., & Kaya, V. D. (2021). Examining prospective teachers' metacognitive learning strategies and self-regulated online learning levels during Covid-19 pandemic. *International Journal of Contemporary Educational Research*, 8(4), 144–157. <https://doi.org/10.33200/IJCER.912897>
- Alamri, H., Lowell, V., Watson, W., & Watson, S. L. (2020). Using personalized learning as an instructional approach to motivate learners in online higher education: Learner self-determination and intrinsic motivation. *Journal of Research on Technology in Education*, 52(3). <https://doi.org/10.1080/15391523.2020.1728449>
- Ampatzoglou, A., Bibi, S., Avgeriou, P., Verbeek, M., & Chatzigeorgiou, A. (2019). Identifying, categorizing and mitigating threats to validity in software engineering secondary studies. *Information and Software Technology* (Vol. 106). <https://doi.org/10.1016/j.infsof.2018.10.006>
- Araka, E., Maina, E., Gitonga, R., & Oboko, R. (2020). Research trends in measurement and intervention tools for self-regulated learning for e-learning environments—systematic review (2008–2018). *Research and Practice in Technology Enhanced Learning*, 15(1). <https://doi.org/10.1186/s41039-020-00129-5>
- Arnold, K. E., Karcher, B., Wright, C. V., & McKay, J. (2017). Student empowerment, awareness, and self-regulation through a quantified-self student tool. *ACM International Conference Proceeding Series*, 526–527. <https://doi.org/10.1145/3027385.3029434>
- Azevedo, R., & Hadwin, A. F. (2005). Scaffolding Self-Regulated Learning and Metacognition-Implications for the Design of Computer-Based Scaffolds. *Instructional Science*, 33, 367–379. <https://doi.org/10.1007/s11251-005-1272-9>
- Azevedo, R., Mudrick, N. V., Taub, M., & Bradbury, A. E. (2019). Self-regulation in computer-assisted learning systems. *The Cambridge Handbook of Cognition and Education*. <https://doi.org/10.1017/9781108235631.024>
- Balavijendran, L., & Burnie, M. (2018). Ready to study: An online tool to measure learning and align university and student expectations via reflection and personalisation. *ASCILITE 2018 - Conference Proceedings - 35th International Conference of Innovation, Practice and Research in the Use of Educational Technologies in Tertiary Education: Open Oceans: Learning Without Borders*, 35–44.

- Bodily, R., & Verbert, K. (2017). Review of research on student-facing learning analytics dashboards and educational recommender systems. *IEEE Transactions on Learning Technologies*, 10(4), 405–418. <https://doi.org/10.1109/TLT.2017.2740172>
- Boekaerts, M., & Corno, L. (2005). Self-regulation in the classroom: A perspective on assessment and intervention. *Applied Psychology*, 54(2), 199–231. <https://doi.org/10.1111/j.1464-0597.2005.00205.x>
- Bouchet, F., Harley, J. M., Trevors, G. J., & Azevedo, R. (2013). Clustering and profiling students according to their interactions with an intelligent tutoring system fostering self-regulated learning. *JEDM - Journal of Educational Data Mining*, 5(1), 104–146. <https://doi.org/10.5281/zenodo.3554613>
- Broadbent, J., & Poon, W. L. (2015). Self-regulated learning strategies & academic achievement in online higher education learning environments: A systematic review. *The Internet and Higher Education*, 27, 1–13. <https://doi.org/10.1016/J.IHEDUC.2015.04.007>
- Cambruzzi, W., Rigo, S. J., & Barbosa, J. L. V. (2015). Dropout prediction and reduction in distance education courses with the learning analytics multitrail approach. *Journal of Universal Computer Science*, 21(1), 23–47. <https://doi.org/10.3217/JUCS-021-01-0023>
- Ceron, J., Baldiris, S., Quintero, J., Garcia, R. R., Saldarriaga, G. L. V., Graf, S., & Fuente Valentin, L. D. la. (2021). Self-regulated learning in massive online open courses: A state-of-the-art review. *IEEE Access* (Vol. 9). <https://doi.org/10.1109/ACCESS.2020.3045913>
- Cha, H. J., & Park, T. (2019). Applying and evaluating visualization design guidelines for a MOOC dashboard to facilitate self-regulated learning based on learning analytics. *KSII Transactions on Internet and Information Systems*, 13(6), 2799–2823. <https://doi.org/10.3837/tiis.2019.06.002>
- Chen, L., Lu, M., Goda, Y., Shimada, A., & Yamada, M. (2020). Factors of the use of learning analytics dashboard that affect metacognition. *17th International Conference on Cognition and Exploratory Learning in Digital Age, CELDA 2020, November*, 295–302. https://doi.org/10.33965/celda2020_202014i038
- Cody, C., Warren, D., Chi, M., & Barnes, T. (2020). Does autonomy help Help? The impact of unsolicited hints on help avoidance and performance. *Proceedings of the 13th International Conference on Educational Data Mining, EDM 2020, Edm*.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112, 155–159. <https://doi.org/10.1037/0033-2909.112.1.155>
- Corrin, L., & de Barba, P. (2014). Exploring students' interpretation of feedback delivered through learning analytics dashboards. *Proceedings of ASCILITE 2014 - Annual Conference of the Australian Society for Computers in Tertiary Education, February 2015*, 629–633.
- Davis, D., Chen, G., Jivet, I., Hauff, C., & Houben, G. J. (2016). Encouraging metacognition and self-regulation in MOOCs through increased learner feedback. *CEUR Workshop Proceedings*, 1596, 17–22.
- Davis, D., Chen, G., Jivet, I., Hauff, C., Kizilcec, R. F., & Houben, G. J. (2017). Follow the successful crowd: Raising MOOC completion rates through social comparison at scale? *ACM International Conference Proceeding Series*, 454–463. <https://doi.org/10.1145/3027385.3027411>
- de Freitas, S., Gibson, D. C., du Plessis, C., Halloran, P., Williams, E., Ambrose, M., Dunwell, I., & Arnab, S. (2015). Foundations of dynamic learning analytics: Using university student data to increase retention. *British Journal of Educational Technology*, 46(6), 1175–1188.
- Domínguez, C., García-Izquierdo, F. J., Jaime, A., Pérez, B., Rubio, Á. L., & Zapata, M. A. (2021). Using process mining to analyze time distribution of self-assessment and formative assessment exercises on an online learning tool. *IEEE Transactions on Learning Technologies*, 14(5), 709–722. <https://doi.org/10.13039/501100011033>
- Efklides, A. (2011). Interactions of metacognition with motivation and affect in self-regulated learning: The MASRL model. *Educational Psychologist*, 46(1), 6–25. <https://doi.org/10.1080/00461520.2011.538645>
- Espinoza, P., & Genna, G. M. (2021). Hi, I want to talk to you about your progress: A large course intervention for at-risk college students. *Journal of College Student Retention: Research, Theory and Practice*, 23(1). <https://doi.org/10.1177/1521025118790054>
- Frey, T. F., Gkotsis, G., & Mikroyannidis, A. (2016). Are you thinking what I'm thinking? Representing Metacognition with Question-based Dialogue Conference Item. *CEUR Workshop Proceedings*, 51–58.
- Hadwin, A. F., Järvelä, S., & Miller, M. (2011). 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* (pp. 65–84). Routledge.

- Hardebolle, C., Jermann, P., Pinto, F., & Tormey, R. (2020). Impact of a learning analytics dashboard on the practice of students and teachers. *SEFI 47th Annual Conference: Varietas Delectat... Complexity Is the New Normality, Proceedings*, 1622–1632.
- Haynes, C. C. (2020). The Role of Self-Regulated Learning in the Design, Implementation, and Evaluation of Learning Analytics Dashboards. *L@S 2020 - Proceedings of the 7th ACM Conference on Learning @ Scale*, 297–300. <https://doi.org/10.1145/3386527.3406732>
- Herodotou, C., Naydenova, G., Boroowa, A., Gilmour, A., & Rienties, B. (2020). How can predictive learning analytics and motivational interventions increase student retention and enhance administrative support in distance education? *Journal of Learning Analytics*, 7(2), 72–83. <https://doi.org/10.18608/JLA.2020.72.4>
- Howell, J. A., Roberts, L. D., & Mancini, V. O. (2018). Learning analytics messages: Impact of grade sender comparative information and message style on student affect and academic resilience. *Computers in Human Behavior*, 89, 8–15. <https://doi.org/10.1016/j.chb.2018.07.021>
- Ifenthaler, D., & Yau, J. Y. K. (2020). Utilising learning analytics to support study success in higher education: a systematic review. *Educational Technology Research and Development*, 68(4). <https://doi.org/10.1007/s11423-020-09788-z>
- Inventado, P. S., Legaspi, R., & Numao, M. (2013). Helping students manage personalized learning scenarios. *Proceedings of the 6th International Conference on Educational Data Mining, EDM 2013*, 2–4.
- Järvelä, S., Järvenoja, H., & Malmberg, J. (2019). Capturing the dynamic and cyclical nature of regulation: Methodological Progress in understanding socially shared regulation in learning. *International Journal of Computer-Supported Collaborative Learning*, 14(4). <https://doi.org/10.1007/s11412-019-09313-2>
- Jivet, I., Scheffel, M., Schmitz, M., Robbers, S., Specht, M., & Drachsler, H. (2020). From students with love: An empirical study on learner goals, self-regulated learning and sense-making of learning analytics in higher education. *Internet and Higher Education*, 47(July), 100758. <https://doi.org/10.1016/j.iheduc.2020.100758>
- Jivet, I., Wong, J., Scheffel, M., Valle Torre, M., Specht, M., & Drachsler, H. (2021). Quantum of choice: How learners' feedback monitoring decisions, goals and self-regulated learning skills are related. *ACM International Conference Proceeding Series*, 416–427. <https://doi.org/10.1145/3448139.3448179>
- Jivet, I., Specht, M., Scheffel, M., & Drachsler, H. (2018). License to evaluate: Preparing learning analytics dashboards for educational practice Citation for published version (APA). *Proceedings of the 8th International Conference on Learning Analytics and Knowledge : (LAK '18)*, 32–40. <https://doi.org/10.1145/3170358.3170421>
- Jovanović, J., Saqr, M., Joksimović, S., & Gašević, D. (2021). Students matter the most in learning analytics: The effects of internal and instructional conditions in predicting academic success. *Computers and Education*, 172. <https://doi.org/10.1016/j.compedu.2021.104251>
- Khalil, M., & Ebner, M. (2015). Learning Analytics: Principles and Constraints. *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2015*, 1326–1336. <https://doi.org/10.13140/RG.2.1.1733.2083>
- Kia, F. S., Teasley, S. D., Hatala, M., Karabenick, S. A., & Kay, M. (2020). How patterns of students dashboard use are related to their achievement and self-regulatory engagement. *ACM International Conference Proceeding Series*, 340–349. <https://doi.org/10.1145/3375462.3375472>
- Lallé, S., Conati, C., Azevedo, R., Mudrick, N., & Taub, M. (2017). On the influence on learning of student compliance with prompts fostering self-regulated learning. *Proceedings of the 10th International Conference on Educational Data Mining, EDM 2017*, 120–127.
- Lee, D., Watson, S. L., & Watson, W. R. (2019). Systematic literature review on self-regulated learning in massive open online courses. *Australasian Journal of Educational Technology*, 35(1), 28–41. <https://doi.org/10.14742/AJET.3749>
- Li, H., Ogata, H., Tsuchiya, T., Suzuki, Y., Uchida, S., Ohashi, H., & Konomi, S. (2017a). Using learning analytics to support computer-assisted language learning. *Proceedings of the 25th International Conference on Computers in Education, ICCE 2017a - Main Conference Proceedings, December*, 908–913.
- Li, I. H., Hwang, G. H., & Lin, Y. X. (2017b). Data mining on the prior knowledge and the effectiveness of the self-regulated learning. *2017b 4th International Conference on Industrial Engineering and Applications, ICIEA 2017b*, 275–278. <https://doi.org/10.1109/IEA.2017.7939221>

- Lim, L. A., Gentili, S., Pardo, A., Kovanović, V., Whitelock-Wainwright, A., Gašević, D., & Dawson, S. (2021). What changes, and for whom? A study of the impact of learning analytics-based process feedback in a large course. *Learning and Instruction*, 72, 101202. <https://doi.org/10.1016/j.learninstruc.2019.04.003>
- Lim, L.-A., Dawson, S., Gašević, D., Joksimović, S., Fudge, A., Pardo, A., & Gentili, S. (2020). Students' sense-making of personalised feedback based on learning analytics. *Australasian Journal of Educational Technology*, 36(6), 15–33. <https://doi.org/10.14742/ajet.6370>
- Lodge, J. M., Panadero, E., Broadbent, J., & de Barba, P. G. (2019). Supporting self-regulated learning with learning analytics. In Lodge, Horvath, & Corrin (Eds.), *Learning analytics in the classroom* (Issue October, pp. 45–55). Routledge.
- Lu, O. H. T., Huang, J. C. H., Huang, A. Y. Q., & Yang, S. J. H. (2017). Applying learning analytics for improving students engagement and learning outcomes in an MOOCs enabled collaborative programming course. *Interactive Learning Environments*, 25(2), 220–234. <https://doi.org/10.1080/10494820.2016.1278391>
- Maher, J. M., Markey, J. C., & Ebert-May, D. (2013). The Other Half of the Story: Effect Size Analysis in Quantitative Research. *Life Sciences Education*, 12, 345–351. <https://doi.org/10.1187/cbe.13-04-0082>
- Malmberg, J., Saqr, M., Järveloja, H., & Järvelä, S. (2022). How the Monitoring Events of Individual Students Are Associated With Phases of Regulation: A Network Analysis Approach. *Journal of Learning Analytics*, 9(1), 77–92. <https://doi.org/10.18608/jla.2022.7429>
- Manganello, F., Pozzi, F., Passarelli, M., Persico, D., & Dagnino, F. M. (2021). A dashboard to monitor self-regulated learning behaviours in online professional development. *International Journal of Distance Education Technologies*, 19(1), 18–34. <https://doi.org/10.4018/IJDET.2021010102>
- Manso-Vázquez, M., & Llamas-Nistal, M. (2015). A Monitoring System to Ease Self-Regulated Learning Processes. *Revista Iberoamericana De Tecnologías Del Aprendizaje*, 10(2), 52–59. <https://doi.org/10.1109/RITA.2015.2417952>
- Matcha, W., Gašević, D., Uzir, N. A., Jovanović, J., & Pardo, A. (2019). Analytics of learning strategies: Associations with academic performance and feedback. *ACM International Conference Proceeding Series*, March, 461–470. <https://doi.org/10.1145/3303772.3303787>
- Matcha, W., Uzir, N. A., Gasevic, D., & Pardo, A. (2020). A Systematic Review of Empirical Studies on Learning Analytics Dashboards: A Self-Regulated Learning Perspective. In *IEEE Transactions on Learning Technologies* (Vol. 13, Issue 2, pp. 226–245). Institute of Electrical and Electronics Engineers. <https://doi.org/10.1109/TLT.2019.2916802>
- Menchaca, I., Guenaga, M., & Solabarrieta, J. (2018). Learning analytics for formative assessment in engineering education. *International Journal of Engineering Education*, 34(3), 953–967.
- Munshi, A., & Biswas, G. (2019). Personalization in OELEs: Developing a data-driven framework to model and scaffold SRL processes. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 11626 LNAI, 354–358. https://doi.org/10.1007/978-3-030-23207-8_65
- Naranjo, D. M., Prieto, J. R., Moltó, G., & Calatrava, A. (2019). A visual dashboard to track learning analytics for educational cloud computing. *Sensors*, 19(13), 2952. <https://doi.org/10.3390/s19132952>
- Nikolayeva, I., Yessad, A., Laforge, B., & Luengo, V. (2020). Does an e-mail reminder intervention with learning analytics reduce procrastination in a blended university course? *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. LNCS, 12315, 60–73. https://doi.org/10.1007/978-3-030-57717-9_5
- Nouri, J., Larsson, K., & Saqr, M. (2019). Identifying Factors for Master Thesis Completion and Non-completion Through Learning Analytics and Machine Learning. *Transforming Learning with Meaningful Technologies Proceedings*, 28–39. https://doi.org/10.1007/978-3-030-29736-7_3
- Ott, C., Robins, A., Haden, P., & Shephard, K. (2015). Illustrating performance indicators and course characteristics to support students' self-regulated learning in CS1. *Computer Science Education*, 25(2), 174–198. <https://doi.org/10.1080/08993408.2015.1033129>
- Page, M. J., Moher, D., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ..., McKenzie, J. E. (2021). PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. *BMJ*, 372. <https://doi.org/10.1136/BMJ.N160>
- Panadero, E. (2017). A review of self-regulated learning: Six models and four directions for research. *Frontiers in Psychology*, 8(APR). <https://doi.org/10.3389/fpsyg.2017.00422>

- Pedrotti, M., & Nistor, N. (2019). How Students Fail to Self-regulate Their Online Learning Experience. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 11722 LNCS, 377–385. https://doi.org/10.1007/978-3-030-29736-7_28
- Pérez-Álvarez, R., Maldonado-Mahauad, J., & Pérez-Sanagustín, M. (2018). Tools to Support Self-Regulated Learning in Online Environments: Literature Review. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 11082 LNCS(January), 16–30. https://doi.org/10.1007/978-3-319-98572-5_2
- 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. 452–502). Academic Press.
- Putwain, D. W., & von der Embse, N. P. (2021). Cognitive-behavioral intervention for test anxiety in adolescent students: do benefits extend to school-related wellbeing and clinical anxiety. *Anxiety Stress & Coping*, 34(1), 22–36. <https://doi.org/10.1080/10615806.2020.1800656>
- Puustinen, M., & Pulkkinen, L. (2001). Models of self-regulated learning: A review. *Scandinavian Journal of Educational Research*, 45, 269–286. <https://doi.org/10.1080/00313830120074206>
- Rienties, B., Cross, S., & Zdrahal, Z. (2017). Implementing a learning analytics intervention and evaluation framework: What works? In B. K. Daniel (Ed.), *Big data and learning analytics in higher education* (pp. 147–166). Springer.
- Rohloff, T., Sauer, D., & Meinel, C. (2019). Student Perception of a Learner Dashboard in MOOCs to Encourage Self-Regulated Learning. *TALE 2019 - 2019 IEEE International Conference on Engineering, Technology and Education*. <https://doi.org/10.1109/TALE48000.2019.9225939>
- Roll, I., & Winne, P. H. (2015). Understanding, evaluating, and supporting self-regulated learning using learning analytics. *Journal of Learning Analytics*, 2(1), 7–12. <https://doi.org/10.18608/jla.2015.21.2>
- Romero, C., & Ventura, S. (2020). Educational data mining and learning analytics: An updated survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 10(3), e1355. <https://doi.org/10.1002/WIDM.1355>
- Russell, J. E., Smith, A., & Larsen, R. (2020). Elements of Success: Supporting at-risk student resilience through learning analytics. *Computers and Education*, 152(March), 103890. <https://doi.org/10.1016/j.compedu.2020.103890>
- Samuel, T. S., & Warner, J. (2021). “I Can Math!”: Reducing Math Anxiety and Increasing Math Self-Efficacy Using a Mindfulness and Growth Mindset-Based Intervention in First-Year Students. *Community College Journal of Research and Practice*, 45(3). <https://doi.org/10.1080/10668926.2019.1666063>
- Saqr, M., & López-Pernas, S. (2021). Idiographic learning analytics: A single student (N=1) approach using psychological networks. *CEUR Workshop Proceedings*, 2868.
- Schumacher, C., & Ifenthaler, D. (2021). Investigating prompts for supporting students’ self-regulation – A remaining challenge for learning analytics approaches? *The Internet and Higher Education*, 49, 100791. <https://doi.org/10.1016/j.iheduc.2020.100791>
- Sedrakyan, G., & Snoeck, M. (2017). Cognitive feedback and behavioral feedforward automation perspectives for modeling and validation in a learning context. *Communications in Computer and Information Science*, 692, 70–92. https://doi.org/10.1007/978-3-319-66302-9_4
- Sedrakyan, G., Malmberg, J., Verbert, K., Järvelä, S., & Kirschner, P. A. (2020). Linking learning behavior analytics and learning science concepts: Designing a learning analytics dashboard for feedback to support learning regulation. *Computers in Human Behavior*, 107(2014), 1–26. <https://doi.org/10.1016/j.chb.2018.05.004>
- Sedraz Silva, J. C., Zamboni, E., Rodrigues, R. L., Ramos, J. L. C., & da Fonseca De Souza, F. (2018). Effects of learning analytics on students’ self-regulated learning in flipped classroom. *International Journal of Information and Communication Technology Education*, 14(3), 91–107. <https://doi.org/10.4018/IJICTE.2018070108>
- Shafiq, D. A., Marjani, M., Habeeb, R. A. A., & Asirvatham, D. (2022). Student Retention Using Educational Data Mining and Predictive Analytics: A Systematic Literature Review. *IEEE Access*, 10, 72480–72503. <https://doi.org/10.1109/ACCESS.2022.3188767>
- Shyr, W. J., & Chen, C. H. (2018). Designing a technology-enhanced flipped learning system to facilitate students’ self-regulation and performance. *Journal of Computer Assisted Learning*, 34(1), 53–62. <https://doi.org/10.1111/JCAL.12213>

- Siadaty, M., Gašević, D., & Hatala, M. (2016a). Associations between technological scaffolding and micro-level processes of self-regulated learning: A workplace study. *Computers in Human Behavior*, 55, 1007–1019. <https://doi.org/10.1016/j.chb.2015.10.035>
- Siadaty, M., Gašević, D., & Hatala, M. (2016b). Measuring the impact of technological scaffolding interventions on micro-level processes of self-regulated workplace learning. *Computers in Human Behavior*, 59(June), 469–482. <https://doi.org/10.1016/j.chb.2016.02.025>
- Siadaty, M., Gasevic, D., & Hatala M. (2016c). Trace-based micro-analytic measurement of self-regulated learning processes. *Journal of Learning Analytics* 3(1). <https://doi.org/10.18608/jla.2016.31.11>
- Siemens, G. (2013). Learning Analytics: The Emergence of a Discipline. *American Behavioral Scientist*, 57(10), 1380–1400. <https://doi.org/10.1177/0002764213498851>
- Sønderlund, A. L., Hughes, E., & Smith, J. (2019). The efficacy of learning analytics interventions in higher education: A systematic review. *British Journal of Educational Technology*, 50(5), 2594–2618. <https://doi.org/10.1111/bjet.12720>
- Sonnenberg, C., & Bannert, M. (2015). Discovering the Effects of Metacognitive Prompts on the Sequential Structure of SRL-Processes Using Process Mining Techniques. *Journal of Learning Analytics*, 2(1), 72–100. <https://doi.org/10.18608/jla.2015.21.5>
- Sonnenberg, C., & Bannert, M. (2016). Evaluating the impact of instructional support using data mining and process mining: A micro-level analysis of the effectiveness of metacognitive prompts. *JEDM - Journal of Educational Data Mining*, 8(2), 51–83. <https://doi.org/10.5281/zenodo.3554597>
- Sonnenberg, C., & Bannert, M. (2019). Using Process Mining to examine the sustainability of instructional support: How stable are the effects of metacognitive prompting on self-regulatory behavior? *Computers in Human Behavior*, 96(June 2018), 259–272. <https://doi.org/10.1016/j.chb.2018.06.003>
- Tabuenca, B., Kalz, M., Drachsler, H., & Specht, M. (2015). Time will tell: The role of mobile learning analytics in self-regulated learning. *Computers & Education*, 89, 53–74. <https://doi.org/10.1016/j.compedu.2015.08.004>
- Tabuenca, B., Greller, W., & Verpoorten, D. (2021). Mind the gap smoothing the transition to higher education fostering time management skills. *Universal Access in the Information Society*. <https://doi.org/10.1007/s10209-021-00833-z>
- Tempelaar, D., Rienties, B., & Nguyen, Q. (2020). Subjective data, objective data and the role of bias in predictive modelling: Lessons from a dispositional learning analytics application. *PLoS ONE*, 15(6), e0233977. <https://doi.org/10.1371/JOURNAL.PONE.0233977>
- Törmänen, T., Järvenoja, H., Saqr, M., Malmberg, J., & Järvelä, S. (2022). Affective states and regulation of learning during socio-emotional interactions in secondary school collaborative groups. *British Journal of Educational Psychology*. <https://doi.org/10.1111/BJEP.12525>
- Valle, N., Antonenko, P., Dawson, K., & Huggins-Manley, A. C. (2021). Staying on target: A systematic literature review on learner-facing learning analytics dashboards. *British Journal of Educational Technology*, 52(4), 1724–1748. <https://doi.org/10.1111/BJET.13089>
- van Alten, D. C. D., Phielix, C., Janssen, J., & Kester, L. (2020). Effects of self-regulated learning prompts in a flipped history classroom. *Computers in Human Behavior*, 108, 106318. <https://doi.org/10.1016/j.chb.2020.106318>
- van Horne, S., Curran, M., Smith, A., VanBuren, J., Zahrieh, D., Larsen, R., & Miller, R. (2018). Facilitating Student Success in Introductory Chemistry with Feedback in an Online Platform. *Technology, Knowledge and Learning*, 23(1), 21–40. <https://doi.org/10.1007/s10758-017-9341-0>
- Viberg, O., Khalil, M., & Baars, M. (2020). Self-regulated learning and learning analytics in online learning environments: A review of empirical research. *ACM International Conference Proceeding Series, March*, 524–533. <https://doi.org/10.1145/3375462.3375483>
- Winne, P. H. (2022). Modeling self-regulated learning as learners doing learning science: How trace data and learning analytics help develop skills for self-regulated learning. *Metacognition and Learning*, 1–19. <https://doi.org/10.1007/S11409-022-09305-Y/TABLES/1>
- Winne, P. H., & Hadwin, A. F. (1998). Studying as self-regulated engagement in learning. In D. Hacker, J. Dunlosky, & A. Graesser (Eds.), *Metacognition in Educational Theory and Practice* (pp. 277–304). Erlbaum.
- Winne, P. H., Teng, K., Chang, D., Lin, M. P. C., Marzouk, Z., Nesbit, J. C., Patzak, A., Rakovic, M., Samadi, D., & Vytasek, J. (2019). NStudy: Software for learning analytics about learning processes and self-regulated learning. *Journal of Learning Analytics*, 6(2), 95–106. <https://doi.org/10.18608/jla.2019.62.7>

- Winne, P. H. (2017). Learning Analytics for Self-Regulated Learning. In *Handbook of Learning Analytics* (pp. 241–249). <https://doi.org/10.18608/hla17.021>
- Wise, A. F. (2014). Designing Pedagogical Interventions to Support Student Use of Learning Analytics. *Proceedings of the Fourth International Conference on Learning Analytics and Knowledge*. <https://doi.org/10.1145/2567574>
- Wise, A. F., Zhao, Y., & Hausknecht, S. (2014). Learning Analytics for Online Discussions: Embedded and Extracted Approaches. *Journal of Learning Analytics*, 1(2), 48–71. <https://doi.org/10.18608/jla.2014.12.4>
- Wise, A. F., Vytasek, J., Hausknecht, S., & Zhao, Y. (2016). Developing learning analytics design knowledge in the “middle space”: The student tuning model and align design framework for learning analytics use. *Online Learning Journal*, 20(2), 155–182. <https://doi.org/10.24059/olj.v20i2.783>
- Wong, B. T. M. (2017). Learning analytics in higher education: An analysis of case studies. *Asian Association of Open Universities Journal*, 12(1), 21–40. <https://doi.org/10.1108/aaouj-01-2017-0009>
- Wong, J., Baars, M., Davis, D., van der Zee, T., Houben, G. J., & Paas, F. (2019a). Supporting Self-Regulated Learning in Online Learning Environments and MOOCs: A Systematic Review. *International Journal of Human-Computer Interaction*, 35(4–5), 356–373. <https://doi.org/10.1080/10447318.2018.1543084>
- Wong, J., Khalil, M., Baars, M., de Koning, B. B., & Paas, F. (2019b). Exploring sequences of learner activities in relation to self-regulated learning in a massive open online course. *Computers and Education*, 140(June), 103595. <https://doi.org/10.1016/j.compedu.2019.103595>
- Yu, L. C., Lee, C. W., Pan, H. I., Chou, C. Y., Chao, P. Y., Chen, Z. H., Tseng, S. F., Chan, C. L., & Lai, K. R. (2018). Improving early prediction of academic failure using sentiment analysis on self-evaluated comments. *Journal of Computer Assisted Learning*, 34(4), 358–365. <https://doi.org/10.1111/jcal.12247>
- Zheng, J., Jiang, N., & Dou, J. (2020). Autonomy support and academic stress: A relationship mediated by self-regulated learning and mastery goal orientation. *New Waves Educational Research & Development*, 23(summer).
- Zhou, X., Jin, Y., Zhang, H., Li, S., & Huang, X. (2016). A map of threats to validity of systematic literature reviews in software engineering. *Proceedings - Asia-Pacific Software Engineering Conference, APSEC*. <https://doi.org/10.1109/APSEC.2016.031>
- Zimmerman, B. J. (2000). Attaining self-regulation: a social cognitive perspective. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of Self-Regulation* (pp. 13–40). Academic Press. <https://doi.org/10.1016/b978-012109890-2/50031-7>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

ARTICLE II

© Springer 2024. Reprinted, with permission from:

Heikkinen, S., Cristea, T., Saqr, M., Malmberg, J., Kleingeld, A., Snijders, C., Matzat, U., & Tedre, M. (2024). Sequence analysis and process mining perspectives to goal setting: What distinguishes business students with high and low self-efficacy beliefs?. *Smart Learn. Environ.* 11, 40.

<https://doi.org/10.1186/s40561-024-00327-4>

RESEARCH

Open Access

Sequence analysis and process mining perspectives to goal setting: What distinguishes business students with high and low self-efficacy beliefs?

Sami Heikkinen^{1,2*} , Tudor Cristea³, Mohammed Saqr¹, Jonna Malmberg⁴, Ad Kleingeld⁵, Chris Snijders³, Uwe Matzat³ and Matti Tedre¹

*Correspondence:
samp@uef.fi

¹ Faculty of Science and Forestry, School of Computing, University of Eastern Finland, Joensuu, Finland

² LAB University of Applied Sciences, Mikkulankatu 19, 15210 Lahti, Finland

³ Human-Technology Interaction Group, Department of Industrial Engineering and Innovation Sciences, Eindhoven University of Technology, Eindhoven, The Netherlands

⁴ Department of Educational Sciences and Teacher Education, University of Oulu, Oulu, Finland

⁵ Human Performance Management Group, Department of Industrial Engineering and Innovation Sciences, Eindhoven University of Technology, Eindhoven, The Netherlands

Abstract

This study investigates the relationship between students' self-efficacy beliefs, goal-setting, and learning tactics in an online business course. Using sequence analysis and process mining techniques, we analyzed log data from 209 students to identify distinct learning tactics and their association with self-efficacy beliefs, inferred from initial goal-setting and final grades. Four learning tactics were identified, with significant differences observed between students with high and low self-efficacy beliefs. High self-efficacy students demonstrated more varied and adaptive tactics, including greater use of quizzes and project-focused activities. In contrast, low self-efficacy students and those who didn't set goals showed less engagement and fewer monitoring activities. The project-focused tactic had the strongest correlation with final grades and goal achievement. Goal-setting at the course's start was linked to more effective learning behaviors and better outcomes. The study reveals how students' online learning behavior changes based on initial goal-setting and subsequent reflection. These findings contribute to research on self-regulated learning in online environments and offer practical implications for designing online courses and learning analytics interventions. Understanding these differences enables the development of targeted interventions to improve learning tactics and self-efficacy beliefs among students, ultimately enhancing their ability to achieve learning goals in online educational settings.

Keywords: Sequence analysis, Process mining, Self-efficacy beliefs, Learning analytics, Project management

Introduction

Self-Regulated learning (SRL) is a process including goal setting and planning, monitoring the progress of learning, and reflecting on outcomes and the process influencing learning outcomes (Pintrich, 2000). Goal-setting is a subprocess included in the first phase of SRL, namely as a part of the planning process, and depending on the goal quality it can lead to improved SRL (Boekaerts & Corno, 2005; Winne & Hadwin, 1998). This effect has been found in different levels of education, e.g., in primary level (Hirsto et al.,

2022) in vocational high schools (Chang et al., 2018) and in higher education MOOCs (Wong et al., 2021). Students who excel in goal-setting are likelier to complete coursework (Handoko et al., 2019), since goals help students to monitor their progress at the learning task. Decades of research has shown that SRL is not easy (McCardle & Hadwin, 2015). Especially in online studying, skills in SRL are particularly needed (Heikkinen et al., 2023).

Recent studies have further emphasized the importance of SRL and goal-setting in online learning environments. Motivation, cognition, metacognition, and self-efficacy are significant components in the SRL process (Gambo & Shakir, 2021). This underscores the multifaceted nature of SRL and the need for comprehensive approaches to support learners in online settings. The relationship between self-efficacy and SRL has been a focus of recent research. Shin (2024) found that self-efficacy can predict self-regulation and course outcomes, suggesting that students' beliefs in their capabilities play a crucial role in their ability to effectively manage their learning. The effective use of meta-cognitive strategies can lead to an increase in self-efficacy, indicating a potential reciprocal relationship between SRL skills and self-efficacy beliefs (Khosravi et al., 2023). In the context of online learning, recent studies have explored innovative approaches to support SRL. Hsu et al. (2023) investigated how students allocate time during exam preparation, finding that consistency in study time throughout the preparation period led to better performance. This highlights the importance of time management skills within the broader framework of SRL. Implementing adaptive learning strategies with diverse didactic techniques can lead to improved learning outcomes and achievement of both disciplinary and transversal competencies (Rincon-Flores et al., 2024).

The need for SRL is recognized in project management and business education, although ways to support SRL are still few (De Oliveira Fassbinder et al., 2017; Marcelino-Sádaba & Perez-Ezcurdia, 2020). The research of SRL in business education started more than a decade ago. For instance, Schloemer and Brennan (2006) looked at developing the SRL skills of accounting students. Within the study, the students received a list of competencies established by the university, upon which they set personal learning goals to achieve them. Instructors used the learning goals to perform prompting interventions to support learners in considering alternative strategies to achieve said goals better. Surveys were used to help students monitor their progress based on the amount of studying time in the past week. In the middle of the course, students did a mid-term evaluation to assess the need for change in learning strategy. The self-evaluation was accompanied by teacher feedback. The results showed that goal-setting and monitoring helped students to improve effective learning behaviors, such as class preparation and collaboration.

A proper learning goal must be specific, measurable, and challenging, but yet attainable (Zimmerman & Schunk, 2011), just like the definition of the project goal (Doran, 1981). In practice, goal-setting is an excellent task for students at the beginning of the course; it combines SRL and project management perspectives. There are two elements connected to goal setting which explain the students' ability to achieve the goal: self-efficacy beliefs and expectancy value. Self-efficacy beliefs are an important factor impacting students' motivation and academic achievement (Pajares, 1996). Educators play a key role in developing stronger self-efficacy beliefs with educational practices. Self-efficacy

beliefs have a relation with SRL, especially via self-assessment, and factors enhancing it include the amount of feedback and support for a self-assessment process (Panadero et al., 2017). The expectancy-value theory explains individuals' attitudes and behaviour based on their beliefs and evaluations of certain outcomes. It proposes that people's attitudes are influenced by two main factors: expectancy and value (Wigfield & Eccles, 2000). Expectancy refers to a student's subjective belief about the likelihood or probability of a particular outcome occurring because of learning activities. For example, if someone believes that studying hard will result in good grades, they have a high expectancy regarding the relationship between studying and academic success. Value refers to the subjective importance or desirability that a student attaches to a specific outcome or goal. This is to say, when students believe they can succeed and see the value in the learning task, they are more likely to persist in their efforts, leading to improved learning outcomes.

Learning analytics to support learning

Learning analytics (LA) has a rich collection of methods that can be used to analyse online learning processes. 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" (Siemens & Long, 2011). The data used in LA can come from clicks, various assessments such as quizzes, or log traces; these are usually gathered from a learning management system (LMS). Different ways to analyse this data include, e.g., measuring (Elmoazen et al., 2023) and predicting student performance (Namoun & Alshanqiti, 2020) or drop-out (de Oliveira et al., 2021), evaluating the elements of networking (Malmberg et al., 2022; Saqr & Alamro, 2019), and different learning patterns (Juhaňák et al., 2019).

Recent advancements in LA have expanded our understanding of student engagement and learning processes in online environments. Nakamura et al. (2024) employed change point detection algorithms to analyze student engagement in classrooms, demonstrating the potential of such techniques to provide valuable insights for improving teaching and learning processes. This approach offers new possibilities for identifying significant shifts in student behavior and engagement, potentially allowing for more timely and targeted interventions.

However, it's important to note that while LA interventions can improve online engagement, they do not necessarily lead to improved learning outcomes (Villalobos et al., 2024). This highlights the need for a more nuanced approach to implementing LA, one that goes beyond simply increasing engagement to focus on enhancing learning effectiveness. In this context, (Tepgec et al., 2024) argue for the importance of developing students' feedback literacy skills, emphasizing that the ability to understand and apply feedback is crucial for translating increased engagement into improved learning outcomes. Recent research has also explored the application of LA in problem-based learning (PBL) environments and revealed that students' levels of self-efficacy were related with their success rates in solving problems (Liu et al., 2023), while LA dashboards can significantly impact self-efficacy in virtual reality simulation-based training (Gallagher et al., 2024).

Learning patterns can be analysed as sequences to identify learning tactics students use while studying online (Hadwin et al., 2007). This sequence analysis approach uses students' trace data to construct sequences of the student's actions studying online. The differences in the ways students act in LMS can be interpreted as the learning tactic students use (Jovanovic et al., 2020). This approach has been used to analyse learning tactics of business students (Heikkinen et al., 2022). There is a well-established connection between the usage of learning tactics and strategies, and the outcomes of learning processes (Jovanović et al., 2017; Siadaty et al., 2016). Another perspective to study patterns is process mining, which can be used, e.g., to cluster sequences to show the students' interactions with the content and activities in LMS (Bogarín et al., 2014). An understanding of students' learning processes enables the development of teaching processes and learning designs (Alqaheri & Panda, 2022; Stoyanov & Kirschner, 2023).

Building on these approaches, recent research has explored more personalized and adaptive learning systems. Takami et al. (2023) proposed a personality-based tailored explainable recommendation system for trustworthy smart learning. This innovative approach suggests that by considering individual differences in personality and learning preferences, we can create more effective and engaging learning experiences. Such personalized systems could potentially address the varying needs of students with different levels of self-efficacy and goal-setting behaviors, offering targeted support and interventions.

Sequence analysis methods can be used to understand the needs of students with low self-efficacy beliefs—the ones who believe that higher grades are either too hard to be achieved or require too much work. We do know that sometimes students are struggling with their assignments (López-Pernas & Saqr, 2021), but we are not sure how we can help them. It is also known that students have different goals, i.e. not all students want to achieve the highest grades (e.g., Knouse et al., 2014), e.g., due to low self-efficacy beliefs. Still, there may be students who are completely satisfied with the grade that is enough to pass the course. When designing interventions to support learning, it is essential to identify the students who are aiming for higher grades but do not have sufficient skills and distinguish these students from the ones who are satisfied with unpretentious learning outcomes. By aiming the intervention at the former group, the instructor's effort is expected to produce a better output (Dahling & Ruppel, 2016).

Clustering is an approach which can be used to distinguish the differences in students' learning profiles. Clustering has been applied to distinguish learner profiles in online management education. Das and Bhuwandeep (2022) used the Online Student Engagement (OSE) scale (Dixon, 2015) as a framework for clustering. The clustering was done to identify cohorts of students with similar learning preferences. The clusters differed based on the importance of attributes (e.g., the timing of online classes and teaching methods).

The power of LA has not been fully harnessed to effectively support SRL of business students' learning in project management. There are few LA studies conducted to foster SRL in the field of business studies. In software engineering project management, LA is used to assess and predict teamwork (Petkovic, 2016). James et al. (2020) use LA to explore social factors influencing learning performance with the help of social network analysis. Virtual business projects are supported with scaffolding customised with

automated reflection and feedback. These studies put their focus on the reflection phase of the SRL cycle. Thus, the understanding of goal-setting, self-efficacy beliefs, and how those affect learning remains unstudied in the field of business education.

Purpose and aims of the study

In light of these recent developments in LA, SRL, and adaptive learning strategies, there is a clear need for research that integrates these approaches in the context of business education. The current study aims to address this gap by examining how LA can help understand different learning tactics between students with low and high self-efficacy beliefs in a project management course. By understanding these differences, we can create targeted interventions capable of supporting low self-efficacy belief students to improve their learning tactics and learning processes to achieve their individual learning goals.

- RQ1 What kind of learning tactics do project management students use when studying online (RQ1.I), could the different learning tactics be used as an indicator towards achieving their goals (RQ1.II), and what are the differences in learning processes between different learning tactics (RQ1.III)?
- RQ2 How do learning tactics differ between low and high self-efficacy belief students (RQ2.I) and what are the differences in learning processes between low and high self-efficacy belief students (RQ2.II)?

By addressing these research questions, this study aims to contribute to the growing body of knowledge on SRL in online environments, particularly in the context of business education. The findings will not only enhance our understanding of the relationship between self-efficacy, goal-setting, and learning behaviors but also inform the development of more effective and personalized learning interventions in project management education.

Materials and methods

We apply a data-driven approach with log data captured from LMS to trace the actions students take studying online. This approach is widely used (Elmoazen et al., 2022; López-Pernas et al., 2021; Zhang & Paquette, 2023). Sequence analysis and process mining can be studied using statistical software, e.g., R. There are several libraries which have been used to research patterns, including TraMiner, seqHMM (López-Pernas & Saqr, 2021; Saqr & López-Pernas, 2021); pMineR, rENA (Uzir et al., 2020); and BupaR (López-Pernas et al., 2021; Uzir et al., 2020).

Study context

The current study was conducted at the Finnish University of Applied Sciences. The context was a project management course organised entirely online. There were three consecutive course implementations included in the research materials. Each implementation, lasting 14 weeks, was arranged during the academic year 2021–2022. The students participating in the course were first-year business students from the organising UAS and students from different UASes, who participated in the course during

the summer semester. Only the students ($n = 209$) who gave their informed consent are included in the research.

The extent of the course is 5 ECTS (European Credit Transfer System, equals c. 130 h of student's active participation). The course contains two themes: creative problem-solving and project management. These themes are divided into sub-themes following the phases of problem-solving (Lubart, 2001) and project planning (Meredith et al., 2017). In the problem-solving phases, students identify a problem, gather information, create ideas for the identified problem, and evaluate the created ideas to choose the best ones, whereas the phases for project planning include scope and work packages, schedule, budgeting, and compiling project plan documents. The project planning theme included an assignment focusing on SCRUM (Hron & Obwegeser, 2022; Takeuchi & Nonaka, 1986), aiming to provide tools for agile project execution. Students were expected to follow the order of these phases during their studying.

The course was organised as an asynchronous course in Moodle LMS. The course materials contained 17 video lectures. Trouble-shooting material was available for the students who struggled with the assignments, along with supplementary material for those who were eager to learn more. The students could use all the materials from the beginning of the course. Along with the video lectures, there was one assignment for each sub-theme, in total, ten assignments. In the first assignment, students chose the problem they drove to solve during the course in the following assignments. Moreover, to cultivate goal-setting (the first phase of SRL), each student was expected to set an individual learning goal for the course. Here a principle of formative assessment (William et al., 2009) was implemented even though it is not commonly used in business education (Ochuot & Modiba, 2018). Formative assessment was applied by using the number of accepted deliverables as criteria for the final grade of the course (five-tier numeric scale, 0 = failed, 5 = excellent). Within the instructions of the assignments, the criterion for accepted deliverables was defined, thus enabling the students to monitor (the second phase of SRL) their actions while preparing deliverables. If the criteria were not achieved, the instructor gave feedback to students on how to improve performance to achieve the level of an accepted deliverable. In the last assignment, the students were asked to recall their goals set in the first assignment and reflect (the third phase of SRL) on their learning and how they would improve their learning if they could start the same course again. These learning design elements can be seen as a tool to fulfil the need for interventions supporting all phases of SRL (Heikkinen et al., 2023).

Data sources

Three data sources were used from each student: (1) the student's numeric goal set (the overall grade for the course) in the deliverable of the first assignment, (2) LMS log data, and (3) the final grade. LMS log data included a timestamp for students' actions in LMS, user IDs, course module IDs, and event names. We used the log data as a basis for the analysis as it can unveil the learning processes (Siemens & Long, 2011) and it has been proven to be useful in the contexts of different LMSs (Viberg et al., 2018). Furthermore, time-stamped log data is a valuable tool for assessing the temporality of learning (Chen et al., 2018).

The final grades were used both as an indicator of course performance and as a benchmark to assess if a student had achieved their goal related to self-efficacy beliefs in terms of their intended goal set at the beginning of the course.

Data preparation

The first step of the analysis was cleaning the log data. The non-learning activities were removed from the LMS log data. These included teacher actions and learners’ actions not directly linked to active learning (e.g., profile views, viewing download centre, logging into the course). This was followed by the anonymization of the student IDs. The output of this step was re-coded into categories (Table 1) to illustrate the different contents and activities students use while studying in LMS. The coding scheme was grouping similar

Table 1 Recoding LMS event to learning actions

Recoded learning actions	LMS Event name
Assignment	A file has been uploaded
	An online text has been uploaded
	An online text uploaded (first assignment)
	Submission created (first assignment)
	Submission created
	Submission form viewed (first assignment)
	Submission form viewed
	Submission updated (first assignment)
	Submission updated
	The status of the submission has been updated (first assignment)
Course viewed	The status of the submission has been updated
	Course viewed
Deliverable	A submission has been submitted (first assignment)
	A submission has been submitted
Feedback	Feedback viewed
Forum	Discussion viewed
Interaction	Comment created
Learning resource	Course module viewed
	mod_hvp: attempt submitted
Monitoring	Badge listing viewed
	Course activity completion updated
	Course module instance list viewed
	Course user report viewed
	Grade overview report viewed
	Grade user report viewed
	Recent activity viewed
	The status of the submission has been viewed
	User report viewed
	Quiz attempt reviewed
Quiz	Quiz attempt started
	Quiz attempt submitted
	Quiz attempt summary viewed
	Quiz attempt viewed

Table 2 Descriptive statistics of students’ usage of learning content and actions

Learning action	n	Percent
Assignment	4262	7.60
Course viewed	10,718	19.12
Deliverable	1355	2.42
Feedback	3027	5.40
Interaction	32	0.06
Learning resource	20,343	36.30
Monitoring	13,104	23.38
Quiz	3091	5.52
Reading forum	115	0.21
	56,047	100.00

Table 3 Coding the goal achievement

Student cohort	The goal set vs. final grade
High self-efficacy beliefs	The Goal set in the first assignment $\leq$ The final grade
Low self-efficacy beliefs	The Goal set in the first assignment $>$ The final grade
Goal-not-set	The goal is not set (e.g., a student stated that “the grade is not important for me” or “I just want to pass the course”)

actions into categories (Molenaar & Chiu, 2014; Saqr et al., 2023) “assignment”, “course viewed”, “deliverable”, “feedback”, “interaction”, “learning resource”, “monitoring”, “quiz”, and “reading forum” to simplify the interpretation of analysis results. The cleaned and re-coded data included coded logs, timestamps, and student IDs, resulting in 56,047 rows of data. The descriptive statistics display the different activities students performed during studying (Table 2). This data was used in the following steps of analysis.

The students were allocated into three cohorts by comparing the student’s final grade with the grade student set as a goal for themselves in the first assignment. The coding scheme for this step is presented in Table 3. These cohorts are used to analyse the distinctive features between low and high self-efficacy belief students.

Data analysis

We analysed the data with sequence analysis to find how students apply the different features of LMS during learning. This approach has been used numerous times to identify online learning tactics and strategies (López-Pernas et al., 2021; Saqr et al., 2023).

Identification of learning tactics

To answer RQ1, to find what kind of learning tactics learners use and what is the distinction between low and high self-efficacy belief students, we applied sequence analysis. Sequence analysis is used to identify learning sessions with similar sequences of learning actions using the TraMineR R package (Gabadinho et al., 2011). The re-coded trace data was organised timewise to model and visualise the learning (Bogarín et al., 2018; Saqr et al., 2023). The re-organised data was grouped into sessions. Different learning actions are determined to belong to a certain session based on the interval between LMS

actions. Every LMS action within the interval of 30 min was considered to belong to the same session, which is the interval used in similar studies previously (Jovanović et al., 2017). We trimmed the excessively short sessions, with less than three actions performed and the excessively long learning sessions with numerous activities were identified as outliers and therefore were trimmed to 50 actions (Jovanović et al., 2017). Before trimming the learning sessions, the maximum session length was 266, with a mean length of 8.38 (SD = 12.13, median = 4). After trimming the overly long and short sessions, the mean session length decreased to 8.03 (SD = 9.75, median = 4). The total number of sessions was 6,688. We used these sessions to describe and interpret what students do during online studying.

To identify students' learning tactics (RQ1.I), each session resulted in a sequence object with timewise organised events. These sequence objects are representations of learning tactics and display how the learning patterns between students differ. We used agglomerative hierarchical clustering (AHC) and tried different clustering methods: optimal matching, longest common subsequence, Hamming distance and Euclidean distance and their applications (Gabadinho et al., 2011) with Ward's algorithm to cluster the sessions with similarities and find the optimal number of clusters (López-Pernas et al., 2021; Uzir et al., 2020). Spell-length-sensitive optimal matching (OMslen) method resulted in four interpretable clusters with the help of fit indices (Gabadinho et al., 2011). OMslen is a modification of the optimal matching algorithm that weights optimal matching elementary operations inversely with episode length (Halpin, 2009). We used three plots to interpret the results: sequence distribution, index, and implication plots (Saqr et al., 2023). A sequence distribution plot illustrates the proportion of different activities at each point in time, an index plot displays each sequence as a stacked bar plot, and an implication plot visualises the difference between the target cluster and other clusters in each step of the sequence. We used fit statistics (R squared and average silhouette width) to support the selection of the best number of clusters (Studer, 2013). We aimed to find distinct clusters that could be used in plotting to interpret the differences in online learning behaviour. The labels for the tactics are based on the context, thus following the procedure used in previous studies (Fincham et al., 2019; Matcha et al., 2020; Saqr et al., 2023).

To find out if there are differences between applying different tactics and achieving the goal a student has set (RQ1.II), we tested Spearman's correlation between different tactics, goal-setting, and final grade to see if these could be used as indicators towards achieving learning goals.

In order to find what are the differences in learning processes between different learning tactics (RQ1.III), we applied process mining methodology to visualise the consecutive order of actions students take during learning sessions. The clusters representing learning tactics were analysed using process mining to create process maps presenting differences between different learning tactics. Process maps illustrating students' actions during their learning sessions were created using process mining (Peeters et al., 2020; Uzir et al., 2020) with the BupaR package (Janssenswillen et al., 2019). The resulting visualisations present the probabilities of different consecutive actions students take during online studying (Chiu & Reimann, 2021), and it is used to explain the order of actions students perform online. This was done to map the relative flow (edge percentage)

between different actions (nodes) and the relative frequency (percentage within nodes) of the actions. In plotting, we used 0.02 as a threshold to suppress the noise in process maps with processmapR and DiagrammeR libraries. This action cuts off the transitions used least often. Thus, threshold limits the number of edges in graphs and makes it easier to interpret the plots.

Identification of differences between high and low self-efficacy belief students

To answer RQ2, we applied the same methodology as in the previous step. We continued by grouping the students into three cohorts: (a) students with high self-efficacy beliefs, (b) with low self-efficacy beliefs, and (c) who did not set learning goals for themselves, thus whose self-efficacy beliefs could not be determined (RQ2.I). This was done to find out if there are differences in applying different tactics between cohorts. We used process mining to create process maps of applying different learning activities, presenting differences between low and high self-efficacy belief students, and students who did not set their learning goals (RQ2.II).

Results

In this section, we present the results for each research question in subsections following the order of the research questions.

Identification of learning tactics

The sequence distribution plot (Fig. 1) displays the sequences (n = 6688) of LMS data. The sequences represent the actions students take during a learning session. The x-axis represents the order of actions (the first action in the first slice, the second in the second slice, etc.) and the y-axis represents the distribution of different actions of all sessions at the particular step of the sequence. The plot shows that most students

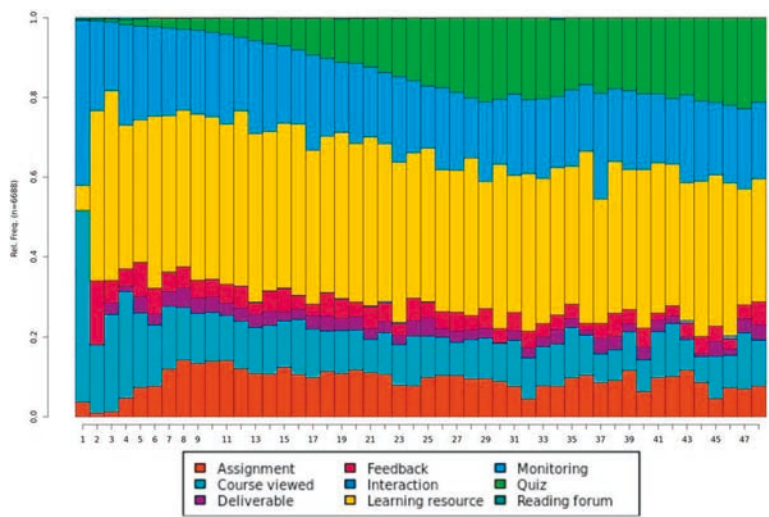


Fig. 1 Plot of overall sequence distribution of learner actions in the LMS

start (the first slice on the x-axis) their learning session by viewing the course or performing monitoring activities (the dark blue slice). Then, students continue using learning resources (the yellow slice). This is the dominating action type throughout sessions. Feedback is applied most often as the second step of the learning process. As the sessions unfold, the frequency of assignments and quizzes increases. This might indicate that students first check the status of their learning process and then check the feedback they have received before continuing to study. Once they have gotten familiar with the current theme, they check the instructions for the assignment and perform quizzes before submitting deliverables. The frequency of interaction and reading forum activities are very low. The thin lines present these on the sequences' 19th, 21st–23rd, 34th, 43rd, and 46th steps.

We used AHC to identify distinct learning tactics (RQ1.1). As a result of the OMslens method, we identified four distinct sequence clusters representing different learning tactics (Fig. 2):

- *Short-focused* ($n = 3746$, 56.0%): This is the shortest sequence cluster where students' learning tactic is focused mostly on viewing the course. Some effort is put into course materials before returning to viewing the course at the end of the learning session.
- *Quiz orientation* ($n = 1698$, 25.4%): A quiz is the most used activity in this sequence cluster. The relative frequency of quizzes increases as the sessions unfold. Prior steps before the quiz include using learning resources and monitoring and feedback activities. After the peak proportion of quiz actions, students return to check learning resources and monitoring activities.
- *Project focused* ($n = 689$, 10.3%): Within this sequence cluster, the relative frequency of the assignment actions is highest compared to the other clusters. In these sessions, students check the assignment's requirements, study the learning resources, and then submit the deliverables. However, the sessions do not end with submission. Instead,

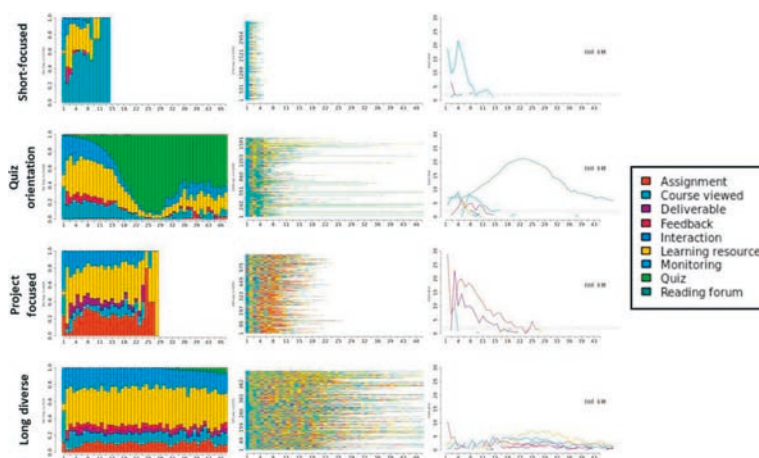


Fig. 2 Sequence distribution plot of learning tactics (left), sequence index plot (middle), and implication plot (right) of the distinct learning tactics

students check the assignment (perhaps to see the instructions of the consecutive assignment) once more before returning to check learning resources.

- *Long diverse* (n = 555, 8.3%): In the sessions belonging to this sequence cluster, students keep using the feedback throughout the learning sessions. Even though the relative frequency of feedback is not high, it is constantly used along with learning resources, monitoring activities and assignments.

We examined the correlation between goals, grades, and students' learning tactics (RQ1.II) (Fig. 3). There was a significant correlation ($p < 0.05$) between almost every factor, except short-focused tactic and goal $r(167) = 0.13$, $p > 0.05$. The highest correlation coefficients were between grade and project focused tactic $r(211) = 0.78$, $p < 0.05$; short-focused and quiz orientation tactics $r(211) = 0.73$, $p > 0.05$; short-focused and project focused tactics $r(211) = 0.71$, $p < 0.05$; and grade and quiz orientation tactic $r(211) = 0.71$, $p < 0.05$. The associations between different tactics had lower correlation coefficients: Long diverse and grade $r(211) = 0.62$, $p < 0.05$; quiz orientation and project focused $r(211) = 0.65$, $p < 0.05$, long diverse and quiz orientation $r(211) = 0.61$, $p < 0.05$; long diverse and short focused $r(211) = 0.44$, $p < 0.05$; goal and grade $r(167) = 0.43$, $p > 0.05$; long diverse and project focused $r(211) = 0.33$, $p < 0.05$; project focused and grade $r(211) = 0.29$, $p < 0.05$; quiz orientation and grade $r(211) = 0.24$, $p < 0.05$; long diverse and grade $r(211) = 0.24$, $p < 0.05$. These results show that all tactics are important towards higher grades, while project focused seems to be the most important learning tactic.

Identification of differences in learning processes

Sequential process maps of students' learning tactics (Fig. 4) illustrate the flow of learning activities and display the differences between students' learning tactics (RQ1.III). The percentages in the figure display the proportion of learning sessions which include the transition (edges) between different activities (nodes) included in the particular session type.

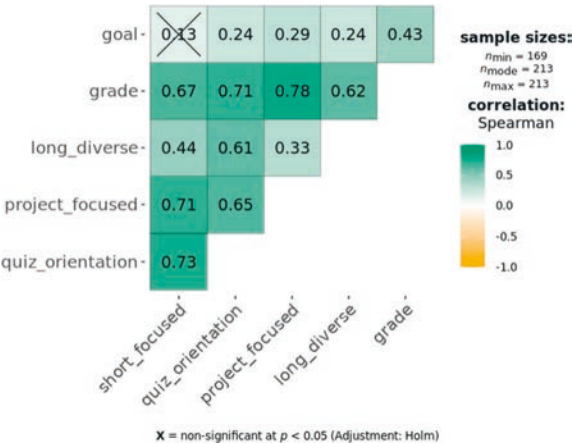


Fig. 3 Correlation matrix between goal, grade, and students' learning tactics

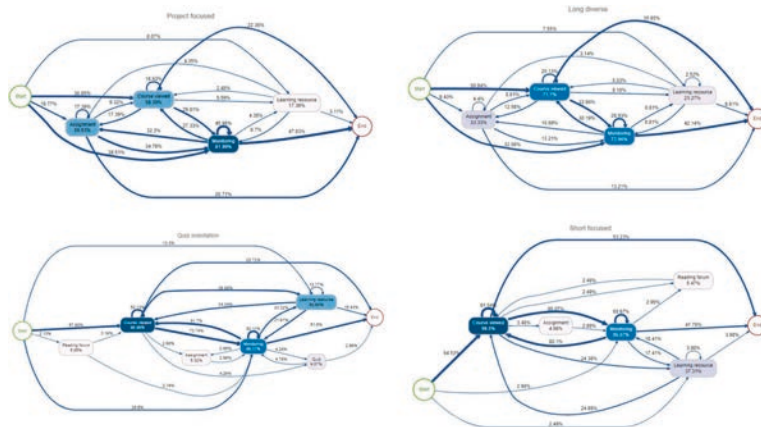


Fig. 4 Sequential process maps of students' learning tactics

- *Project focused* students are likely to start sessions with monitoring (38.5%), course-viewed (36.7%), or assignment (16.8%) actions. Then, students tend to move from monitoring to an assignment (32.3%) and back (34.8%), or from monitoring to course viewed (27.3%) and back (29.8%). This tactic rarely includes visits to learning resource actions (17.4%).
- *Long diverse* this tactic starts with course viewed (50.9%) or monitoring (32.1%). This is followed by the second step from course viewed to monitoring (34.0%) or from monitoring to course viewed (30.2%). An assignment (33.3%) is less frequently used when compared with project focused (59.6%), whereas learning resource actions have increased (from 17.4% to 23.3%).
- *Quiz orientation* starts with course viewed (58.8%), monitoring (27.8%), or learning resource (13.4%). From the course viewed the flow continues to either monitoring (62.0%) or learning resource (40.1%). Within this tactic, the learning resource usage has the highest value (60.6%). This process map shows a loop between course viewed, monitoring, and learning resource activities. An assignment (5.3%) and reading forum (5.9%) are visited rarely. This is the only process map, where the quiz (9.6%) is visited.
- *Short-focused* students start sessions with course viewed (94.5%). As the next step, students continue with monitoring (85.1%), which is followed most likely by returning to course viewed (80.1%) or learning resource (17.4%). Students using this tactic take more action to read forums (5.5%) than interacting with assignments (5.0%).

Differences between high and low self-efficacy belief students

The learning tactics distribution plot between student cohorts (Fig. 5) illustrates the differences in the usage of different learning tactics between student cohorts (RQ2.1):

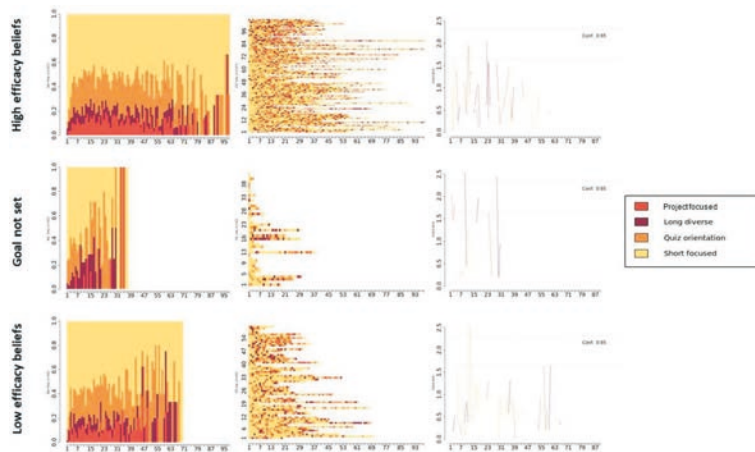


Fig. 5 Learning tactics distribution plot between student cohorts

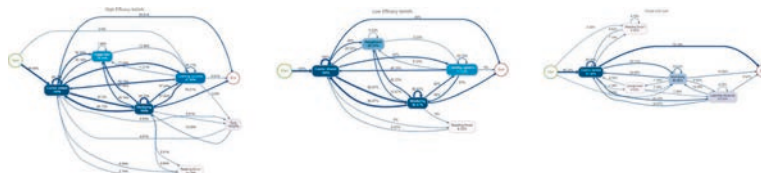


Fig. 6 Sequential process maps of student cohorts

- *High self-efficacy beliefs* ($n=107$, 51.2%): This is the largest cohort and the sequence of different tactics used is the longest. Students with high self-efficacy beliefs use the quiz orientation tactic more often than the rest of the cohorts. The short-focused is the most used tactic, while long diverse and project focused tactics are used during almost the entire span of the sequences.
- *Low self-efficacy beliefs* ($n=60$, 28.7%): The length of the overall distribution plot for this cohort is shorter than for the students with high self-efficacy beliefs. These students use fewer learning sessions during the course, and the proportion of the short-focused tactic is slightly more used than students with high self-efficacy beliefs.
- *Goal-not-set* ($n=42$, 20.1%): The most distinct feature of this cohort is the low frequency of the project focused tactic and higher frequency of the long diverse tactic. This suggests that these students wander through the learning space but do not put significant effort into assignment activities.

Differences in relative-case process maps between low and high self-efficacy beliefs students, and students who did not set their learning goals are illustrated in Fig. 6 (RQ2.II).

- *High self-efficacy beliefs* start their learning processes with course viewed (96.3%). This is followed by monitoring (97.2%), learning resource (70.1%), or assignment

(52.3%). There is a strong flow between each of these four actions. Students with high self-efficacy beliefs are also the most active cohort to read the forum (12.2%) and the only cohort that uses quizzes (16.9%).

- *Low self-efficacy beliefs* Start every learning session from the course viewed (100%), which is followed by monitoring (96.7%), learning resource (50%), or assignment (40%). From monitoring, students tend to return to the course viewed (96.7%). When compared to students with high self-efficacy beliefs, low self-efficacy beliefs students tend to use fewer assignments, learning resources, and reading forum actions and do not use quiz activities.
- *Goal-not-set* The students within this group start with course viewed (95.2%), which is the only step before ending the session (76.2%) or monitoring (35.7%). Within this cohort, the usage of monitoring, learning resource, and assignment actions are the lowest, whereas reading forum is higher than non-achievers.

Discussion

Goal-setting is a subprocess of SRL and can positively impact students' academic achievement (Handoko et al., 2019). Our study contributes to the current body of research by implementing sequence and process mining techniques to better understand the differences in online learning tactics between low and high self-efficacy beliefs students. We identified the learning tactics of students based on sequenced events from their learning sessions. We also used the goals students set at the beginning of the course to identify differences in learning behaviour and learning outcomes. Our results indicate a positive effect of goal-setting at the beginning of the course, expressed through better learning tactics employed, but also more sessions and time spent by students overall in the LMS.

Overall, the most used activity was accessing learning resources, followed by monitoring, and assignment access. Quizzes can be used as both a summative assessment method and a reflection tool; students fill them out to get a better image of their progress at various points in the course which can then be used to reflect on their learning behaviour and change it accordingly. We also notice that course viewing decreases as learning sessions unfold. This finding suggests that throughout the learning session, students start with more general learning behaviour, such as viewing the course and its learning resources. The next actions within the session show clear differences between the low and high self-efficacy student beliefs. The students with high self-efficacy beliefs move slowly towards more specific or reflective activities, such as quizzes, whereas students with low self-efficacy beliefs struggle to figure out what they should do next to progress with their learning process.

Our first research question was divided into three parts, all building upon the preceding part in a logical progression to give a comprehensive answer. For the first part, we used all the activities previously described to identify four distinct learning tactics used by students during the course. More than half of the learning sessions were short-focused, as students simply viewed the course or the available materials before ending the session. Moreover, such sessions are constant across student goal-setting categories, as they are the most common type of learning sessions for both students that set goals or not, and for low and high self-efficacy beliefs students. We argue such sessions can play

a role in the planning process of SRL (Winne & Hadwin, 1998), as students could access the course to clarify certain aspects of the course to plan future sessions. A more obvious explanation is that students simply get distracted or lose motivation after accessing the LMS; more research should follow up on this, especially considering how much overall time is spent during these sessions.

On the other side, we have the long-diverse tactic, where students are the most varied in their LMS usage. At first glance, we would expect such sessions to be the main learning activity, especially considering the constant usage of feedback, one of the main indicators in predicting academic performance in educational psychology (Hattie & Timperley, 2007). In the context of this course, the formative feedback received for the assignments gives students a perspective on the assessment criteria. After the evaluation, the student can assess their performance with feedback. Students with high self-efficacy beliefs use this as a tool to evaluate and, if needed, to enhance their performance, while other cohorts do not apply reflection at the same amount. However, long-diverse sessions have the second lowest correlation with the final grades and lowest with goal-setting. The quiz and project focused categories are the most straightforward; students use these sessions to focus on specific quizzes or assignments. Such sessions have the highest correlation with grades with project focused having the highest correlation with goals. Considering this difference between long-diverse and quiz or project focused sessions, we argue that having clearly set goals helps the students better focus during their learning sessions, such as focusing on specific assignments or quizzes, which is a more efficient and effective way of learning due to their testing nature. Such a difference could also portray the testing effect (Rowland, 2014), as completing an assignment or quiz would lead to better learning than rereading the materials associated with rote learning. While long-diverse represent a more varied learning behaviour, these could also hide confusion, students accessing everything with no clear purpose in mind. This could be the explanation for the students who did not set their learning goals: they lack the SRL skills, which would both help them to set precise goals for learning and to find ways how they could navigate themselves in the learning space. A final explanation stems from the stable form of the sessions. Long-diverse sessions have a steady structure throughout the course, while quiz and project focused sessions change drastically around the middle and end of sessions. Changing learning behaviour is important as it shows a clear SRL of students who need to adapt to the course features as a response to new material or feedback. Further research is needed to understand better why such differences occur.

Looking at the sequential process maps (Fig. 4) we see that viewing the course, accessing learning materials and monitoring are used the most across tactic categories. Monitoring is an essential subprocess of SRL (Pintrich, 2000; Winne & Hadwin, 1998), and such processes are missing from the only tactic category not correlating with goal-setting (i.e. short-focused); this is a clear indication of their importance. This is further corroborated by the fact that students that did not set any goals also had the lowest number of monitoring activities. As students have a clear goal in mind throughout the course, they know how to better monitor their learning behaviour, as they have a better image of what their final goal should be. This, in turn, offers clear criteria to which they can reflect and adapt the learning tactics used. The fact that a simple goal set at the beginning of the course can lead to such changes in the students' SRL behaviour brings further evidence

to the importance of supporting students in their SRL, changing their zone of proximal development, and allowing for further growth.

When we categorise students based on their self-efficacy beliefs, we see that students with high self-efficacy beliefs access the LMS the most, monitor their progress, access assignments and quizzes to test their progress, and apply feedback to improve their performance. This shows clear SRL behaviour, which further leads to successful learning and high academic performance. In contrast, students that did not set any goals have the most long-diverse sessions and read forums the most. This indicates an almost “confused” behaviour, as not having a clear goal in mind makes them not manage to employ effective learning methods and constantly check for others’ help. Moreover, the fact that they have almost no project focused sessions, with a sudden increase towards the end of the session seems to suggest blocked learning behaviour, which is a non-effective and harmful learning behaviour (Krasnoff & Overkott, 2022). The students with low self-efficacy beliefs have somewhat similar behaviour to students with high self-efficacy beliefs, however, they study less overall and make less use of assignments and forums. All these categories suggest a clear increase in learning and efficient learning tactics when students set a clear goal at the beginning of the course.

Relevance

Our study aligns with recent research in smart learning environments, particularly in the areas of SRL and LA. For instance, Gambo and Shakir (2021) conducted a systematic review on SRL in smart learning environments, highlighting the importance of motivation, cognitive, metacognitive, and self-efficacy components in the SRL process. Our findings on the differences between high and low self-efficacy students in their learning behaviors and goal-setting practices corroborate these components’ significance in SRL.

Furthermore, our findings on the importance of goal-setting and its impact on learning behaviors align with recent studies in smart learning environments. For example, Hsu et al. (2023) investigated how students allocate time during exam preparation and found that consistency in study time throughout the preparation period led to better performance. This supports our argument that clearly set goals help students focus better during their learning sessions, leading to more efficient and effective learning.

The differences we observed in learning behaviors between students with high and low self-efficacy beliefs are also relevant to recent research on adaptive learning strategies. Rincon-Flores et al. (2024) found that implementing an adaptive learning strategy with diverse didactic techniques led to better results in students’ learning and achievement of disciplinary and transversal sub-competencies. This suggests that tailoring learning experiences based on students’ self-efficacy and goal-setting behaviors could further enhance the effectiveness of online learning environments.

Our study’s use of LA to identify and analyze student learning tactics contributes to the growing body of research on using data-driven approaches to improve educational outcomes. For instance, Nakamura et al. (2024) used change point detection algorithms to analyze student engagement in classrooms, demonstrating the potential of such techniques to provide valuable insights for improving teaching and learning processes. Similarly, our sequence mining approach offers a way to automatically identify significant

patterns in student learning behaviors, which can inform instructional design and personalized interventions.

The implications of our findings extend beyond just understanding student behavior. They also suggest potential avenues for developing more effective learning support systems. For example, Takami et al. (2023) proposed a personality-based tailored explainable recommendation system for trustworthy smart learning. Our findings on the differences in learning tactics between high and low self-efficacy students could inform the development of similar adaptive systems that provide personalized support based on students' goal-setting behaviors and self-efficacy levels.

This study sheds more light on how business students use LA tactics (Heikkinen et al., 2022). The results give more evidence of the differences in self-efficacy beliefs (e.g., Knouse et al., 2014) affecting the student's ability to achieve their expected goals. The sequence mining approach gives another perspective to clustering students (Das & Bhuwandeep, 2022), thus, offering a better understanding of the learner perspective. Furthermore, the results show another way to analyse the project management study process besides assessing and predicting teamwork (Petkovic, 2016). The sequence mining and process mining approaches have been previously used in different educational settings (Elmoazen et al., 2022; López-Pernas et al., 2021; Zhang & Paquette, 2023), and this study verifies that these approaches are applicable also in business.

Our results show the importance of helping business students learning project management set clear goals and follow up on them, mainly through changes in online learning behaviour and SRL. We argue that such interventions should be encouraged and promoted more often due to their simplicity but large effects. If a student cannot set a proper goal, this can be seen as an indicator for a teacher to intervene. Such a student needs intervention at the very beginning of the course, which could greatly help students to first set goals for themselves. This should then be followed by a set of interventions aiming to help the students with low self-efficacy beliefs to improve their learning behaviour and match the patterns the students with high self-efficacy beliefs apply. As the meta-analysis by Heikkinen et al. (Heikkinen et al., 2023) discovered, less than half of the SRL interventions have positive significant effects, with very few targeting all phases of SRL. Our study aimed at filling this gap, as we also asked students to reflect on their work at the end of the course.

It is important to note, however, that while LA interventions, such as LA dashboards, can improve online engagement, they do not necessarily improve learning outcomes (Villalobos et al., 2024). This suggests that while feedback may help students improve their activity, it does not always help them understand what they should do to achieve better results. Therefore, as Tepgec et al. (2024) argue, feedback literacy skills should be supported to help students improve their ability to apply feedback and enhance their learning.

Our findings on the relationship between self-efficacy and learning behaviors are further supported by recent research. Shin (2024) found that self-efficacy can predict self-regulation and course outcomes, which aligns with our observations of high self-efficacy students demonstrating more effective SRL behaviors. Additionally, Khosravi et al. (2023) showed that effective use of metacognitive strategies can lead to a rise in self-efficacy, suggesting a potential reciprocal relationship between self-efficacy and SRL skills.

Theoretical and practical implications

This study has four theoretical implications. Firstly, our findings contribute to SRL theory by demonstrating how self-efficacy beliefs and goal-setting behaviors manifest in specific learning tactics within an online project management context. Secondly, this study showcases the potential of LA techniques, particularly sequence and process mining, in providing detailed insights into SRL processes, extending the methodological toolkit for SRL research. Thirdly, by examining SRL in the context of project management education, our study bridges these two fields, offering a new perspective on how SRL principles apply in professional skill development. Lastly, our identification and analysis of four distinct learning tactics contribute to the theoretical understanding of how students approach online learning, potentially informing future research on learning strategies.

Practical Implications focus on course design, targeted interventions, scaffold development, feedback mechanisms, LMS design, and student guidance. Instructors should incorporate explicit goal-setting activities at the beginning of online courses and provide opportunities for students to reflect on and revise these goals throughout the course. LA can be used to identify students with low self-efficacy beliefs early in the course, allowing for timely and targeted interventions to support their learning. Based on the identified learning tactics, educators can develop scaffolding strategies that guide students towards more effective learning behaviors, particularly for those with low self-efficacy beliefs. This can be done by implementing regular feedback mechanisms that not only address content knowledge but also promote self-reflection on learning strategies and goal progress. LMS developers should consider incorporating features that support goal-setting, self-monitoring, and adaptive learning pathways based on students' self-efficacy levels and chosen tactics. This requires developing resources to help students understand different learning tactics and how to adapt their approaches based on course demands and personal goals.

Limitations

While this study provides valuable insights into the learning tactics and SRL behaviors of business students in project management courses, it is important to acknowledge several limitations that may affect the generalizability and interpretation of our findings.

Firstly, our study focused on business students in a project management course at a single institution. While this allows for a deep understanding of this specific context, it may limit the generalizability of our findings to other disciplines, educational levels, or institutional settings. Future research should consider replicating this study across various disciplines and institutions to validate the broader applicability of our findings.

Secondly, the self-efficacy beliefs and goal-setting data were collected through self-reported measures. While widely used in educational research, self-reported data can be subject to social desirability bias and may not always accurately reflect students' true beliefs or behaviors. Future studies could incorporate additional objective measures or observational data to triangulate self-reported information.

Lastly, there were limited timeframe and intervention testing applied in this study. Our study analyzed student behavior over the course of a single academic term. A

longitudinal study spanning multiple terms or years could provide more insight into how learning tactics and SRL behaviors evolve over time, especially as students progress through their academic programs. Although our study identified differences in learning tactics between high and low self-efficacy students, we did not test specific interventions to improve self-efficacy or learning tactics. Future research could focus on designing and testing targeted interventions based on our findings.

Despite these limitations, our study provides valuable insights into the relationship between self-efficacy beliefs, goal-setting, and learning tactics in online project management education. By acknowledging these limitations, we hope to encourage future research that builds upon our findings and addresses these areas to further advance our understanding of SRL in online environments.

Conclusion and future work

The current study makes a novel contribution to the body of literature by portraying in detail the changes in business students' online project management learning behaviour based on assignment goal-setting at the beginning of the course, followed by a reflection at the end.

There are multiple possibilities for future research. One of the main "weaknesses" of LA research, in general, is the lack of context, such as not knowing what the students do or intend to do behind the monitor. The previously discussed short-focused learning tactics are the perfect example. These short sessions were the most common by far, even disregarding the goal-setting factor. While we did give multiple explanations, we believe research should focus more on such "contradicting" behaviour by possibly also including the context. It is expected that the students do not study with 100% efficiency, all learning sessions have a clear goal which is always achieved. However, we also believe the high number of such short-focused sessions mandates further study.

The studied course had quizzes and assignments the students had to complete. While they did seem to represent a more effective way of learning, we are not sure why. One argument we gave was them being associated with the testing effect, a proven method of effective learning. However, we are not sure exactly if there is an actual effect of such assignments and how it could influence online LMS learning. We believe future research should study this effect in particular, as it could lead to high academic performance.

Furthermore, there is a need for the development of interventions and research to study the impact of different interventions for goal-setting of students who cannot set learning goals, and interventions for students with low self-efficacy beliefs to help them regulate their learning processes towards achievement.

Lastly, the trace data, per se, is insufficient to explain the learning processes. The cognitive aspect of learning should be evaluated jointly with the traces. This could be done by researching survey data gathered from learners to find out how they perceive the state of SRL and the impact of interventions.

Acknowledgements

The first author was supported by the LAB-D program.

Author contributions

SH, MS, JM, CS, and MT contributed to the study conception and design. SH performed material preparation and data collection, and SH, MS, and TC performed analysis. SH and TC wrote the first draft of the manuscript. All authors commented on previous versions of the manuscript. SH wrote the revised version of the manuscript. All authors read and approved the final manuscript.

Funding

No funding details are to be reported.

Availability of data and materials

The authors do not have permission to share data.

Declarations**Competing interests**

The authors report there are no competing interests to declare.

Received: 23 June 2023 Accepted: 9 August 2024

Published online: 13 September 2024

References

- Alqahtani, H., & Panda, M. (2022). An education process mining framework: Unveiling meaningful information for understanding students' learning behavior and improving teaching quality. *Information*, 13(1), 29. <https://doi.org/10.3390/INFO13010029>
- Boekaerts, M., & Corno, L. (2005). Self-regulation in the classroom: A perspective on assessment and intervention. *Applied Psychology*, 54(2), 199–231. <https://doi.org/10.1111/j.1464-0597.2005.00205.x>
- Bogarín, A., Cerezo, R., & Romero, C. (2018). A survey on educational process mining. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*. <https://doi.org/10.1002/widm.1230>
- Bogarín, A., Romero, C., Cerezo, R., & Sánchez-Santillán, M. (2014). Clustering for improving Educational process mining. *ACM International Conference Proceeding Series*. <https://doi.org/10.1145/25675742567604>
- Chang, C. C., Liang, C., Chou, P. N., & Liao, Y. M. (2018). Using e-portfolio for learning goal setting to facilitate self-regulated learning of high school students. *Behaviour and Information Technology*, 37(12), 1237–1251. <https://doi.org/10.1080/0144929X.2018.1496275>
- Chen, B., Knight, S., & Wise, A. F. (2018). Critical issues in designing and implementing temporal analytics. *Journal of Learning Analytics*, 5(1), 1–9. <https://doi.org/10.18608/jla.2018.53.1>
- Chiu, M. M., & Reimann, P. (2021). Statistical and stochastic analysis of sequence data. *International Handbook of Computer-Supported Collaborative Learning*. https://doi.org/10.1007/978-3-030-65291-3_29
- Dahling, J. J., & Ruppel, C. L. (2016). Learning goal orientation buffers the effects of negative normative feedback on test self-efficacy and reattempt interest. *Learning and Individual Differences*. <https://doi.org/10.1016/j.lindif.2016.08.022>
- Das, P. (2022). Leveraging level of student engagement for online class design in management education. *Journal of Education for Business*, 97(1), 21–28. <https://doi.org/10.1080/08832323.2021.1884519>
- de Oliveira, C. F., Sobral, S. R., Ferreira, M. J., & Moreira, F. (2021). How does learning analytics contribute to prevent students' dropout in higher education: A systematic literature review. *Big Data and Cognitive Computing*, 5(4), 64. <https://doi.org/10.3390/BDCC5040064>
- De Oliveira Fassbinder, A. G., Fassbinder, M., Barbosa, E. F., & Magoulas, G. D. (2017). Massive open online courses in software engineering education. In *Proceedings: Frontiers in Education Conference, FIE, 2017-October* (pp. 1–9). <https://doi.org/10.1109/FIE.2017.8190588>
- Dixon, M. D. (2015). Measuring student engagement in the online course: The online student engagement scale (OSE). *Online Learning*, 19(4), n4.
- Doran, G. T. (1981). There's a S.M.A.R.T. way to write management's goals and objectives. *Management Review*, 70(11).
- Elmoazen, R., Saqr, M., Khalil, M., & Wasson, B. (2023). Learning analytics in virtual laboratories: A systematic literature review of empirical research. *Smart Learning Environments*, 10(1), 1–20. <https://doi.org/10.1186/S40561-023-00244-Y/FIGURES/4>
- Elmoazen, R., Saqr, M., Tedre, M., & Hirsto, L. (2022). How social interactions kindle productive online problem-based learning: An exploratory study of the temporal dynamics. In L. Hirsto, S. López-Pernas, M. Saqr, E. Sointu, T. Valtonen, & S. Väisänen (Eds.), *Proceedings of the 1st Finnish learning analytics and artificial intelligence in education conference (FLAIEC 2022)* (pp. 68–76). CEUR-WS.
- Fincham, E., Gašević, D., Jovanović, J., & Pardo, A. (2019). From study tactics to learning strategies: An analytical method for extracting interpretable representations. *IEEE Transactions on Learning Technologies*, 12(1), 59–72. <https://doi.org/10.1109/TLT.2018.2823317>
- Gabardin, A., Ritschard, G., Müller, N. S., & Studer, M. (2011). Analyzing and visualizing state sequences in R with TraMineR. *Journal of Statistical Software*. <https://doi.org/10.18637/jss.v040.i04>
- Gallagher, T., Slof, B., van der Schaaf, M., Toyoda, R., Tehreem, Y., Garcia Fracaro, S., & Kester, L. (2024). Reference frames for learning analytics dashboards: The progress and social reference frame and occupational self-efficacy. *Journal of Computer Assisted Learning*, 40(2), 742–760. <https://doi.org/10.1111/JCAL.12912>
- Gambo, Y., & Shakir, M. Z. (2021). Review on self-regulated learning in smart learning environment. *Smart Learning Environments*, 8(1), 1–14. <https://doi.org/10.1186/S40561-021-00157-8/TABLES/3>
- Hadwin, A. F., Nesbit, J. C., Jamieson-Noel, D., Code, J., & Winne, P. H. (2007). Examining trace data to explore self-regulated learning. *Metacognition and Learning*. <https://doi.org/10.1007/s11409-007-9016-7>
- Halpin, B. (2009). Optimal matching analysis and life-course data: The importance of duration. *Sociological Methods and Research*. <https://doi.org/10.1177/0049124110363590>
- Handoko, E., Gronseth, S. L., McNeil, S. G., Bonk, C. J., & Robin, B. R. (2019). Goal setting and MOOC completion. *The International Review of Research in Open and Distributed Learning*. <https://doi.org/10.19173/irrodl.v20i4.4270>

- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112. <https://doi.org/10.3102/003465430298487>
- Heikkinen, S., López-Pernas, S., Malmberg, J., Tedre, M., & Saqr, M. (2022). How do business students self-regulate their project management learning? A sequence mining study. In L. Hirsto, S. López-Pernas, M. Saqr, E. Sointu, T. Valttonen, & S. Väisänen (Eds.), *Proceedings of the 1st finnish learning analytics and artificial intelligence in education conference (FLAIEC 2022)* (pp. 51–59). CEUR-WS.
- Heikkinen, S., Saqr, M., Malmberg, J., & Tedre, M. (2023). Supporting self-regulated learning with learning analytics interventions: a systematic literature review. *Education and Information Technologies*, 28, 3059–3088. <https://doi.org/10.1007/s10639-022-11281-4>
- Hirsto, L., Valttonen, T., Saqr, M., Hallberg, S., Sointu, E., Kankaanpää, J., & Väisänen, S. (2022). Pupils' experiences of utilizing learning analytics to support self-regulated. *Society for Information Technology & Teacher Education International Conference*, 2022(1), 1879–1885.
- Hron, M., & Obwegeser, N. (2022). Why and how is Scrum being adapted in practice: A systematic review. *Journal of Systems and Software*, 183, 111110. <https://doi.org/10.1016/j.jss.2021.111110>
- Hsu, C. Y., Horikoshi, I., Li, H., Majumdar, R., & Ogata, H. (2023). Supporting "time awareness" in self-regulated learning: How do students allocate time during exam preparation? *Smart Learning Environments*, 10(1), 1–15. <https://doi.org/10.1186/s40561-023-00243-z/TABLES/3>
- James, N., Humez, A., & Laufenberg, P. (2020). Using technology to structure and scaffold real world experiential learning in distance education. *TechTrends*, 64(4), 636–645. <https://doi.org/10.1007/S11528-020-00515-2>
- Janssenswillen, G., Depaire, B., Swennen, M., Jans, M., & Vanhoof, K. (2019). bupaR: Enabling reproducible business process analysis. *Knowledge-Based Systems*, 163, 927–930. <https://doi.org/10.1016/j.knosys.2018.10.018>
- Jovanovic, J., Dawson, S., Joksimovic, S., & Siemens, G. (2020). Supporting actionable intelligence: Reframing the analysis of observed study strategies. *ACM International Conference Proceeding Series*. <https://doi.org/10.1145/3375462.3375474>
- Jovanović, J., Gašević, D., Dawson, S., Pardo, A., & Mirriahi, N. (2017). Learning analytics to unveil learning strategies in a flipped classroom. *Internet and Higher Education*. <https://doi.org/10.1016/j.iheduc.2017.02.001>
- Juhaňák, L., Zounek, J., & Rohlíková, L. (2019). Using process mining to analyze students' quiz-taking behavior patterns in a learning management system. *Computers in Human Behavior*, 92, 496–506. <https://doi.org/10.1016/j.chb.2017.12.015>
- Khosravi, R., Dastgoshadeh, A., & Jalilzadeh, K. (2023). Writing metacognitive strategy-based instruction through flipped classroom: An investigation of writing performance, anxiety, and self-efficacy. *Smart Learning Environments*, 10(1), 1–26. <https://doi.org/10.1186/s40561-023-00264-8/TABLES/5>
- Knouse, L. E., Feldman, G., & Blevins, E. J. (2014). Executive functioning difficulties as predictors of academic performance: Examining the role of grade goals. *Learning and Individual Differences*. <https://doi.org/10.1016/j.lindif.2014.07.001>
- Krasnoff, J., & Overkott, C. (2022). Why do people overestimate the effectiveness of blocked learning? *Psychonomic Bulletin and Review*, 1, 1–11. <https://doi.org/10.3758/S13423-022-02225-7/TABLES/1>
- Liu, M., Cai, Y., Han, S., & Shao, P. (2023). Understanding middle school students' self-efficacy and performance in a technology-enriched problem-based learning program: A learning analytics approach. *Journal of Educational Technology Systems*, 51(4), 513–543. <https://doi.org/10.1177/00472395231174034>
- López-Pernas, S., & Saqr, M. (2021). Bringing synchrony and clarity to complex multi-channel data : A learning analytics study in programming education. *IEEE Access*, 9, 166531–166541. <https://doi.org/10.1109/ACCESS.2021.3134844>
- López-Pernas, S., Saqr, M., & Viberg, O. (2021). Putting it all together: Combining learning analytics methods and data sources to understand students' approaches to learning programming. *Sustainability (Switzerland)*, 13(9), 4825. <https://doi.org/10.3390/su13094825>
- Lubart, T. I. (2001). Models of the creative process: Past, present and future. *Creativity Research Journal*, 13(3–4), 295–308. https://doi.org/10.1207/S15326934CRJ1334_07
- Malmberg, J., Saqr, M., Järvenoja, H., Haataja, E., Pijera-Díaz, H. J., & Järvelä, S. (2022). Modeling the complex interplay between monitoring events for regulated learning with psychological networks. *The Multimodal Learning Analytics Handbook*. https://doi.org/10.1007/978-3-031-08076-0_4
- Marcelino-Sádaba, S., & Perez-Ezcurdia, A. (2020). Competence training for project management: holistic analysis framework. In *Handbook of research on project management strategies and tools for organizational success* (pp. 196–222).
- Matcha, W., Gašević, D., Jovanović, J., Uzir, N. A., Oliver, C. W., Murray, A., & Gasevic, D. (2020). *Analytics of learning strategies: Associations with Academic Performance and Feedback*. 151–160. <https://doi.org/10.1145/3375462.3375534>
- McCardle, L., & Hadwin, A. F. (2015). Using multiple, contextualized data sources to measure learners' perceptions of their self-regulated learning. *Metacognition and Learning*. <https://doi.org/10.1007/s11409-014-9132-0>
- Meredith, J. R., Shafer, S. M., & Mantel, S. J., Jr. (2017). *Project management: A strategic managerial approach*. Wiley.
- Molenaar, I., & Chiu, M. M. (2014). Dissecting sequences of regulation and cognition: Statistical discourse analysis of primary school children's collaborative learning. *Metacognition and Learning*. <https://doi.org/10.1007/s11409-013-9105-8>
- Nakamura, K., Ishihara, M., Horikoshi, I., & Ogata, H. (2024). Uncovering insights from big data: Change point detection of classroom engagement. *Smart Learning Environments*, 11(1), 1–19. <https://doi.org/10.1186/s40561-024-00317-6/TABLES/3>
- Namoun, A., & Alshantqiti, A. (2020). Predicting student performance using data mining and learning analytics techniques: A systematic literature review. *Applied Sciences*, 11(1), 237. <https://doi.org/10.3390/AP11010237>
- Ochuot, H. A., & Modiba, M. (2018). Formative assessment as critical pedagogy: A case of business studies. *Interchange*. <https://doi.org/10.1007/s10780-018-9341-6>
- Pajares, F. (1996). Self-efficacy beliefs in academic settings. *Review of Educational Research*. <https://doi.org/10.3102/00346543066004543>
- Panadero, E., Jonsson, A., & Botella, J. (2017). Effects of self-assessment on self-regulated learning and self-efficacy: Four meta-analyses. In *Educational research review* (Vol. 22). <https://doi.org/10.1016/j.edurev.2017.08.004>

- Peeters, W., Saqr, M., & Viberg, O. (2020). Applying learning analytics to map students' self-regulated learning tactics in an academic writing course. In *Proceedings of the 28th international conference on computers in education, August*.
- Petkovic, D. (2016). Using learning analytics to assess capstone project teams. *Computer*, 49(1), 80–83. <https://doi.org/10.1109/MC.2016.3>
- 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. 452–502). Academic Press.
- Rincon-Flores, E. G., Castano, L., Guerrero Solis, S. L., Olmos Lopez, O., Rodríguez Hernández, C. F., Castillo Lara, L. A., & Aldape Valdés, L. P. (2024). Improving the learning-teaching process through adaptive learning strategy. *Smart Learning Environments*, 11(1), 1–27. <https://doi.org/10.1186/S40561-024-00314-9/FIGURES/9>
- Rowland, C. A. (2014). The effect of testing versus restudy on retention: A meta-analytic review of the testing effect. *Psychological Bulletin*, 140(6), 1432–1463. <https://doi.org/10.1037/A0037559>
- Saqr, M., & Alamo, A. (2019). The role of social network analysis as a learning analytics tool in online problem based learning. *BMC Medical Education*, 19(1), 1–11. <https://doi.org/10.1186/s12909-019-1599-6>
- Saqr, M., & López-Pernas, S. (2021). The longitudinal trajectories of online engagement over a full program. *Computers and Education*, 175(September), 104325. <https://doi.org/10.1016/j.compedu.2021.104325>
- Saqr, M., López-Pernas, S., Jovanović, J., & Gašević, D. (2023). Intense, turbulent, or wallowing in the mire: A longitudinal study of cross-course online tactics, strategies, and trajectories. *Internet and Higher Education*. <https://doi.org/10.1016/j.ihe.2022.100902>
- Schloemer, P., & Brenan, K. (2006). From students to learners: Developing self-regulated learning. *Journal of Education for Business*, 82(2), 81–87. <https://doi.org/10.3200/JOEB.82.2.81-87>
- Shin, D. D. (2024). Curiosity promotes self-regulated learning and achievement in online courses for students with varying self-efficacy levels. *Educational Psychology*. <https://doi.org/10.1080/01443410.2024.2372302>
- Siadat, M., Gašević, D., & Hatala, M. (2016). Associations between technological scaffolding and micro-level processes of self-regulated learning: A workplace study. *Computers in Human Behavior*, 55, 1007–1019. <https://doi.org/10.1016/j.chb.2015.10.035>
- Siemens, G., & Long, P. (2011). Penetrating the fog: Analytics in learning and education. *EDUCAUSE Review*, 46(5).
- Stoyanov, S., & Kirschner, P. A. (2023). Text analytics for uncovering untapped ideas at the intersection of learning design and learning analytics: Critical interpretative synthesis. *Journal of Computer Assisted Learning*. <https://doi.org/10.1111/jcal.12775>
- Studer, M. (2013). *WeightedCluster Library Manual: A practical guide to creating typologies of trajectories in the social sciences with R*. (LIVES Working Papers, 24). <https://doi.org/10.12682/lives.2296-1658>
- Takami, K., Flanagan, B., Dai, Y., & Ogata, H. (2023). Personality-based tailored explainable recommendation for trustworthy smart learning system in the age of artificial intelligence. *Smart Learning Environments*, 10(1), 1–19. <https://doi.org/10.1186/S40561-023-00282-6/FIGURES/10>
- Takeuchi, H., & Nonaka, I. (1986). The new new product development game. *Harvard Business Review*, 137–146.
- Teppeck, M., Heil, J., & Ifenthaler, D. (2024). Feedback literacy matters: unlocking the potential of learning analytics-based feedback. *Assessment & Evaluation in Higher Education*. <https://doi.org/10.1080/02602938.2024.2367587>
- Uzir, N. A., Gašević, D., Jovanović, J., Matcha, W., Lim, L. A., & Fudge, A. (2020). 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. <https://doi.org/10.1145/3375462.3375493>
- Viberg, O., Hatakka, M., Bälter, O., & Mavroudi, A. (2018). The current landscape of learning analytics in higher education. *Computers in Human Behavior*, 89, 98–110. <https://doi.org/10.1016/j.chb.2018.07.027>
- Villalobos, E., Pérez-Sanagustín, M., & Broisin, J. (2024). From Learning Actions to Dynamics: Characterizing Students' Individual Temporal Behavior with Sequence Analysis. 3–17. https://doi.org/10.1007/978-3-031-64302-6_1
- Wigfield, A., & Eccles, J. S. (2000). Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology*. <https://doi.org/10.1006/ceps.1999.1015>
- William, D., Black, P., & William, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation, and Accountability*, 21(1), 5–31. <https://doi.org/10.1007/s11092-008-9068-5>
- Winne, P. H., & Hadwin, A. F. (1998). Studying as self-regulated engagement in learning. In D. Hacker, J. Dunlosky, & A. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 277–304). Erlbaum.
- Wong, J., Baars, M., He, M., de Koning, B. B., & Paas, F. (2021). Facilitating goal setting and planning to enhance online self-regulation of learning. *Computers in Human Behavior*. <https://doi.org/10.1016/j.chb.2021.106913>
- Zhang, Y., & Paquette, L. (2023). Sequential pattern mining in educational data: the application context, potential, strengths, and limitations. In A. Peña-Ayala (Ed.), *Educational data science: Essentials, approaches, and tendencies: proactive education based on empirical big data evidence* (1st ed., pp. 219–254). Springer. <https://doi.org/10.1007/978-981-99-0026-8>
- Zimmerman, B., & Schunk, D. (2011). *Handbook of self-regulation of learning and performance*. Routledge, Taylor & Francis Group.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

ARTICLE III

© Springer 2025. Reprinted, with permission from:

Heikkinen, S., Saqr, M., Malmberg, J., & Tedre, M. (2025). A longitudinal study of interplay between student engagement and self-regulation. *Int J Educ Technol High Educ* 22, 21.

<https://doi.org/10.1186/s41239-025-00523-3>

RESEARCH ARTICLE

Open Access

A longitudinal study of interplay between student engagement and self-regulation

Sami Heikkinen^{1,2*} , Mohammed Saqr¹ , Jonna Malmberg³  and Matti Tedre¹ 

*Correspondence:
samp@uef.fi

¹ School of Computing, The University of Eastern Finland, Yliopistonkatu 2, 80100 Joensuu, Finland

² LAB University of Applied Sciences, Mikkulankatu 19, 15210 Lahti, Finland

³ Department of Educational Sciences and Teacher Education, The University of Oulu, P.O. Box 2000, 90014 Oulu, Finland

Abstract

This study investigates the dynamic interplay between student engagement and self-regulated learning (SRL) in an online project management course. Using learning management system trace data combined with longitudinal self-reported SRL measures, we analysed data from 165 first-year business students through multi-channel sequencing and clustering algorithms. Results revealed three distinct patterns of learning behaviour with significant implications for intervention design: Low-regulating improvers (4.85%) who showed potential for growth with targeted support, Disengaged low regulators (44.85%) requiring comprehensive intervention in both engagement and self-regulation skills, and High-regulating improvers (50.30%) who demonstrated successful adaptation to online learning. Our innovative approach of analysing student progress individually rather than using fixed time points enabled more precise identification of support needs in asynchronous learning settings. The findings demonstrate that both engagement and self-regulation skills can develop over time with appropriate support, suggesting opportunities for adaptive intervention throughout the course. This research provides actionable insights for developing targeted support strategies, including early warning systems for at-risk students and personalized scaffolding approaches based on students' behavioural patterns. These findings advance learning analytics by providing a framework for real-time identification of student needs and evidence-based intervention design in online education. The results particularly emphasize the importance of continuous, pattern-based support for developing self-regulation skills in online environments.

Keywords: Learning analytics, Self-regulated learning, Student engagement, Online education, Multi-channel clustering, Sequence analysis, Hidden Markov models, Asynchronous learning

Introduction

In the rapidly evolving landscape of online education, understanding the dynamics of student engagement and self-regulated learning (SRL) has become increasingly important (Dent & Koenka, 2016). As online learning environments become more popular, educators and researchers face the challenge of supporting diverse learners in settings where traditional face-to-face interactions are limited or absent. This

study aims to shed light on the interplay between student engagement and self-regulation in online learning contexts, utilising advanced analytical techniques to uncover learning trajectories.

Our study integrates several complementary theoretical frameworks to understand online learning dynamics. At its foundation is Zimmerman's cyclical model of self-regulated learning, which conceptualizes learning as an active, iterative process of planning, performance, and self-reflection (Zimmerman & Schunk, 2011). We view these processes through the lens of complex dynamic systems theory (Schmitz, 2006), which suggests that learning patterns emerge from the interaction of multiple components over time. This theoretical integration supports our novel multi-channel analytical approach and helps explain why students may follow different trajectories in their development of engagement and self-regulation skills.

The shift towards online learning modalities has accelerated in recent years, and research indicates that educational institutions worldwide have increasingly integrated online courses into their curricula to maintain continuity in education during lockdowns (Ali, 2020; Lapitan et al., 2021). This transition has brought to the forefront the importance of self-regulated learning skills, as students navigate more autonomous learning environments. Research indicates that self-regulation significantly influences learning outcomes, with studies showing that higher levels of self-regulation correlate with improved academic performance in online settings (Almoslamani, 2022; Hensley et al., 2022). Simultaneously, maintaining student engagement in these digital spaces has become a critical concern for educators and instructional designers (Churchill et al., 2013; Li et al., 2024).

While previous research has examined engagement and self-regulation as separate or static constructs, there is a growing recognition of the need to understand their dynamic relationship over time. Understanding the interplay between engagement and self-regulation at a granular level makes it possible to support learners during their online studies. When we can identify learners who struggle with either engagement or self-regulation, we can scaffold their learning processes. Learning analytics (LA) offers powerful tools for capturing and analysing the digital footprints left by students as they interact with online learning platforms. However, the challenge lies in integrating these behavioural traces with students' self-reported experiences of self-regulation to create a more holistic picture of the learning process.

This study employs a multi-channel clustering approach to analyse longitudinal student engagement and self-regulation data in an online project management course. By leveraging various R packages and clustering algorithms, we classify students based on their engagement patterns and self-regulation abilities over time. The analysis utilises data from multiple sources, including learning management system (LMS) logs and repeated survey responses, to uncover hidden states and transitions in student behaviours.

The following section will introduce the study's core concepts: SRL, engagement, LA, sequence analysis, and multi-channel sequence analysis; discuss the field's recent development; and finally, explain the study's motivation.

Background

Our study integrates several complementary theoretical frameworks to understand online learning dynamics. At its foundation is Zimmerman's cyclical model of self-regulated learning, which conceptualizes learning as an active, iterative process of planning, performance, and self-reflection (Zimmerman & Schunk, 2011). This interacts with cognitive load theory (Sweller, 2020), which helps explain how students' cognitive resources are allocated between managing the learning environment and content mastery—a particularly crucial consideration in online settings where students must simultaneously navigate technical interfaces while engaging with course material. Self-determination theory (Ryan & Deci, 2020) further enriches our framework by explaining how autonomy, competence, and relatedness drive engagement and sustained self-regulation in online environments. These theoretical perspectives are complemented by student engagement theory, which posits three dimensions: behavioural, emotional, and cognitive (Martin & Borup, 2022). We view these processes through the lens of complex dynamic systems theory (Schmitz, 2006), which suggests that learning patterns emerge from the interaction of multiple components over time. This theoretical integration supports our novel multi-channel analytical approach and helps explain why students may follow different trajectories in their development of engagement and self-regulation skills.

Self-regulation

SRL is a complex, multi-faceted process through which learners actively control their thoughts, behaviours, and emotions to achieve their learning goals (Zimmerman & Schunk, 2011). It represents the intentional planning, monitoring, and regulation of cognitive, behavioural, and affective processes before, during, and after learning experiences (Panadero, 2017). This cyclical process involves three main phases: forethought (planning), performance (execution and monitoring), and self-reflection (evaluation) (Zimmerman & Schunk, 2011).

In the forethought phase, learners analyse tasks, set goals, and select appropriate strategies. During the performance phase, they implement chosen strategies while continuously monitoring their progress and understanding. In the self-reflection phase, learners evaluate their performance against their goals and adjust their approaches for future learning situations (Winne & Hadwin, 2008).

The effectiveness of SRL depends on the interplay between cognitive, metacognitive, and motivational components (Schunk, 2005). Cognitive components include strategies for processing and manipulating information; metacognitive components involve planning, monitoring, and evaluating learning processes, while motivational components encompass beliefs about one's capabilities, task value, and learning goals (Schunk & Greene, 2017).

SRL involves setting goals, monitoring progress, and regulating the learning process (Järvelä et al., 2019; Pintrich, 2000). Different SRL models focus on different parts of learning, such as motivation (Boekaerts & Corno, 2005), metacognitive aspects (Efklides, 2011), effective cognitive strategies (Winne & Hadwin, 1998), and how the regulation can be shared (Hadwin et al., 2011). Some educational models are also

regarded as comprehensive, prioritising motivation and metacognition as fundamental components of learning (Pintrich, 2000; Zimmerman, 2000). There are diverse methods for assessing this key aspect of learning. One such method for evaluating SRL is through the use of surveys (Barnard et al., 2009; Bremer & Kaufmann, 1927; Broadbent et al., 2023; Jansen et al., 2017; Lin et al., 2023), which can vary in the learning domains they address. These tools can be employed to investigate students' perceptions of their current learning status.

Self-reported SRL data can be combined with trace data captured from LMS. In previous studies, students with strong SRL skills demonstrate higher engagement (Cleary et al., 2021), achieve better grades, and effectively manage their study environments (Estévez et al., 2021). A study environment such as an online course designed to help students enhance their SRL skills can improve their engagement (Wolters et al., 2023) and grades (Hilpert et al., 2023). The engagement is one of the elements of this current study.

Engagement

Engagement is a three-dimensional construct incorporating behaviour, emotion, and cognition (Martin & Borup, 2022; Redmond et al., 2018). Behavioural engagement involves, e.g., study efforts such as being present, involvement, and commitment to learning. Emotional engagement explains a student's different emotions towards studying and other learners. Cognitive engagement indicates students' efforts to start and proceed with learning assignments (Borup et al., 2020). In the current study, the behavioural dimension is crucial as it can be measured using data collected from an LMS. This data includes variables such as completed assignments, frequency of logins, number and frequency of posts, time spent creating posts, time spent online, and the number of accessed learning resources (Fincham et al., 2019b; Henrie et al., 2015). Saqr et al., (2023a, 2023b, 2023c) have given an example of the methodology that can be used to measure online engagement.

Previous studies about engagement show that prior engagement can predict imminent engagement (Gottfried et al., 2009; Wylie & Hodgen, 2012), but engagement is not continuously increasing (Wigfield et al., 2007). Research has identified several categories of variables that influence engagement development. At the individual level, psychological factors such as locus of control and self-esteem, along with emotional states, personality characteristics, and prior learning experiences and grade history shape how students engage with learning. These individual characteristics interact with social factors, including parental expectations, communication patterns, shared values across contexts, and socio-economic status (You & Sharkey, 2009). Additionally, the educational context itself plays a crucial role, particularly through extracurricular activities and various learning environment characteristics (Wang & Degol, 2014).

These diverse influencing factors highlight that engagement is a complex phenomenon that cannot be evaluated in isolation. Therefore, there is a need to inspect engagement changes with more advanced approaches. One option is to investigate it using data triangulation.

Learning analytics

LA is defined as "the measurement, collection, analysis, and reporting of data about learners and their contexts for the purposes of understanding and optimising learning and the environments in which it occurs" (Siemens, 2013). LA is used to analyse and report learner performance and to give insights into the learning processes that can be improved (Lodge et al., 2019; Winne, 2022). Trace data captured from an LMS is input for calculating, analysing, and tracking learners' processes. Instructors typically use the results of the analysis to execute interventions supporting learning activities (Ifenthaler & Yau, 2020). LA interventions aim to support changes learners can make to achieve better outcomes.

While early LA methods relied on cross-sectional data and had limited predictive capabilities, contemporary dynamic models can be applied to more complex learning contexts. A plethora of studies have analysed learning in the context of SRL and used LA to support SRL interventions (Heikkinen et al., 2023). However, implementing effective interventions remains challenging and requires careful design to support learning successfully (Rienties et al., 2017; Sedrakyan et al., 2020; Wise, 2014). A typically used approach for LA interventions is the learning analytics dashboard (LAD; Jivet et al., 2018), that are used to offer actionable insights that empower learners to adjust their behaviours for improved outcomes (Susnjak et al., 2022). A study by Matcha et al., (2020a, 2020b, 2020c) explicitly focused on LAD interventions aimed at supporting learning. Their systematic review highlighted the potential of LADs in supporting various aspects of SRL but also noted challenges in implementation and effectiveness across different contexts. According to Larrabee Sønderlund et al. (2019), most interventions aim to identify at-risk students, which refers to students who are highly likely not to pass the course or if passing with poor results (Shafiq et al., 2022). Highly topical AI has also been used to support SRL by providing personalized feedback and resources with promising results (Jin et al., 2023).

Some supportive processes have primarily focused on enhancing engagement. Suraworachet et al. (2023) investigated the combined effects of human and analytics feedback on student engagement in reflective writing tasks, revealing that such feedback mechanisms can significantly enhance students' engagement levels and performance. This is further supported by the work of Ouyang et al., (2023a, 2023b), which employed social network analysis to examine engagement dimensions in online learning, demonstrating how analytics can illuminate patterns of student interaction and participation. Finally, the findings of Zhang et al. (2023) emphasize that self-regulated learning is crucial for achieving learning goals, suggesting that educational interventions should enhance students' self-regulation skills to promote engagement and academic success.

Since SRL is a complex phenomenon including numerous viewpoints, Järvelä and Hadwin (2024) suggest that multimodal data can be used to analyse SRL. Their framework proposes that disruptive trigger events are markers or catalysts for strategic metacognition or other regulatory approaches. Triggers can be external, such as the elements of learning design or interventions aiming to invite the learner to adapt their learning to meet the needs of the learning process. When learners recognise and respond to these triggers, they may initiate changes in their regulatory processes, leading to adapted learning strategies and behaviours. The effectiveness of triggers depends on

learners' ability to acknowledge them and implement appropriate regulatory responses. This cyclical process of trigger recognition, response, and regulatory adaptation has been observed in various learning contexts, such as collaborative learning environments where multimodal data has been used to inform pedagogical strategies (Xu et al., 2023).

Chatbots are one form of intervention that has gained popularity in recent years. Most studies exploring the usage of chatbots have put effort into promoting productive SRL processes (Sáiz-Manzanares et al., 2023). Still, the support for SRL, such as setting learning goals, planning learning, or reflecting on studies, remains limited (Guan et al., 2024). Building on Järvelä and Hadwin's (2024) framework of triggers, our study implements regular self-evaluation surveys as intentional triggers for SRL processes. By prompting students to assess their SRL skills at multiple points during the online course, these surveys serve two purposes: they provide researchers with data about skill progression over time, and more importantly, they act as metacognitive triggers that encourage students to reflect on and adjust their learning strategies. This structured self-reflection can help students become more aware of their learning processes and make necessary adjustments, potentially leading to enhanced SRL skills and improved learning outcomes.

Self-reported SRL and engagement have been analysed together by (Wang et al., 2024), who used self-reporting at the end of the research experiment. Our approach improves on this setting by applying the multi-channel approach to assess engagement with LA data accompanied by self-reported SRL data. The granularity is improved with temporal data that examines the development of engagement and SRL multiple times during the course.

Sequence analysis

Sequence analysis was developed to address the need for pattern mining (Abbott, 1983). Although it has been used in various disciplines since its initial development in education, it has been adopted somewhat recently (e.g., Saqr et al., 2023a, 2023b, 2023c). Typically, sequence mining is used to identify patterns or sub-sequences (Fournier-Viger et al., 2017). Sequential mining is beneficial in the context of learning because of the processual nature of learning. This process unfolds in time, as states, in a transformative way (Schmitz, 2006). With sequence mining, repeating patterns and sequences can be identified. This is useful when we want to identify similarities between learners. Taking sequence mining to the next level allows the examination of trajectories of events, which helps to assess the sequences in a broader context (Saqr et al., 2023a, 2023b, 2023c).

In recent years, sequence analysis has been used more often in LA. It is typical to use the trace data captured from LMS as a data source for sequence mining study (Jovanović et al., 2017; Matcha et al., 2020a; Saqr et al., 2023c). Trace data forms sessions, the basic unit of analysis in sequence analysis. Most often, sessions are determined based on the uninterrupted sequence of the learner's actions (Jovanović et al., 2017; Matcha et al., 2020a). While most studies focus on intra-course sequences, some studies expand the timespan over the entire study program (Saqr & López-Pernas, 2021b, 2022; Saqr et al., 2023a, 2023b, 2023c).

When sequences are identified, the next step is typically to cluster similar sequences. Different methods are used for this purpose, but Agglomerative Hierarchical Clustering (AHC) has been the most used method (Saqr et al., 2024). Clustering requires the

determination of distance measures to find dissimilarities between clusters. For this purpose, there are different distance measures used, such as Euclidean (Matcha et al., 2020a; Saqr et al., 2023a, 2023b, 2023c), Longest Common Subsequence (López-Pernas et al., 2023), Longest Common Prefix (Saqr & López-Pernas, 2022), and Optimal Matching (Jovanović et al., 2017) used in the previous studies. Along with these dissimilarity measures, recent studies have started to use methods developed to identify more detailed nuances. Markovian Models is this type of approach gaining more popularity (Matcha et al., 2020a; Saqr & López-Pernas, 2022). Other approaches are used to extract meaningful metrics and indicators for distinguishing the nature of learners' sequential patterns (López-Pernas & Saqr, 2024; Poquet et al., 2023; Saqr et al., 2023a, 2023b, 2023c; Tan & Samavedham, 2022).

Recent sequence mining studies have used different categorical data types as data sources. These types of data include behavioural states (Raab & Struffolino, 2023), learning activities (Matcha et al., 2020b), moves in educational games (López-Pernas et al., 2023), or different roles taken during collaborative learning (Saqr & López-Pernas, 2022). One common approach is to define learning actions based on interactions with LMS (López-Pernas et al., 2021; Matcha et al., 2019; Saqr et al., 2023a, 2023b, 2023c; Villalobos, 2022). Another option is to map interactions to higher-level meta-cognitive actions (e.g., SRL strategies) to connect the practical actions to existing theoretical concepts (Dever et al., 2023; Poquet et al., 2023). It is also possible to continue the analysis by sequencing the sequences and thus to analyse the learning, e.g., at the course level (Fincham et al., 2019a, 2019b; Saqr et al., 2023a, 2023b, 2023c; Villalobos, 2022).

The data logs can be used to generate indicators representing students' engagement, such as activity level or study regularity. These indicators can have varying levels of temporal granularity, ranging from daily or weekly data analysis to actions performed during specific learning sessions at particular times (Jovanović et al., 2021). Indicators have also been computed for the COPES model (Winne & Hadwin, 1998) focusing on task definition, goal-setting, enactment, and adaptation (Cristea et al., 2024). This approach aims to leverage the indicators at a theoretical level to trace SRL-related constructs.

Another approach to indicators is the metrics proposed by Ritschard (2023) who divide metrics into three categories measuring sequences: (a) basic indicators (counts of states), (b) diversity (the heterogeneity of visited states), and (c) complexity (unpredictability of state arrangement). Basic indicators are like metrics proposed by (Jovanović et al., 2021) because these measurements are based on counting the number of interactions. Diversity metrics focus on the distribution of states in the sequence, e.g., course strategies students use during a higher education program (Saqr et al., 2023a, 2023b, 2023c). Complexity metrics refer to the instability or unpredictability of state arrangement in the sequence, e.g., students' procrastination behaviour in an online course (Tan & Samavedham, 2022).

The goals of the recent sequence mining studies have varied. Sequence analysis is used to evaluate the relationship between identified patterns of sequences and students' performance (Kinnebrew & Biswas, 2012; Matcha et al., 2020b; Saqr et al., 2023a, 2023b, 2023c), to compare the patterns of course delivery (Matcha et al., 2020a), and examine how sequences are associated with dropout (Saqr & López-Pernas, 2021b) or self-efficacy beliefs (Heikkinen et al., 2024). López-Pernas and Saqr (2024) have used clustering

techniques to detect students' engagement patterns and found three clusters of behaviour: 'Active', 'Average' and 'Disengaged'. They used the concept of integrative potential to group students based on this classification. This metric assesses achieving a positive state, such as 'Active' engagement. Tan and Samavedham (2022) used the precarity index to analyse students' procrastination in an online course. They grouped students based on the time spent watching videos and used the index and other metrics in a regression model. They found that students' behaviour complexity explained about 14% of the variation in their performance. Saqr et al., (2023a, 2023b, 2023c) performed a longitudinal study to analyse students' learning strategies variation across a learning program. Researchers found common context-specific actions across courses and examined student behaviour at different levels.

The integration of multiple theoretical perspectives necessitates a methodological approach that can capture both the temporal nature of learning and the interaction between different learning processes. Cognitive load theory (Sweller, 2020) suggests that students in online environments must balance their cognitive resources between managing the learning platform and engaging with content, while self-determination theory (Ryan & Deci, 2020) emphasizes how autonomy and competence influence sustained engagement and self-regulation. These theoretical perspectives help explain why students might show different patterns of engagement and regulation over time. When combined with SRL's cyclical model and complex dynamic systems theory, these frameworks suggest that learning analytics should consider multiple simultaneous processes. Multi-channel sequence analysis provides a powerful tool for operationalizing these theoretical constructs, allowing us to examine how engagement and self-regulation co-evolve while maintaining the theoretical integrity of each construct. This approach aligns with recent theoretical developments in learning analytics that emphasize the need to consider multiple data streams when studying complex learning phenomena (Järvelä & Hadwin, 2024; Wise, 2014).

Multi-channel sequence analysis

Multi-channel analysis is gaining ground in LA. Multi-channel sequence analysis requires at least two temporal datasets, which can be analysed in parallel. Previous studies have used multi-channel data to create the channels necessary for multi-channel analysis. Along with clickstream or log data (Saqr & López-Pernas, 2021a), scholars have used assessment data (Bacci & Bertaccini, 2022; Lopez-Pernas & Saqr, 2021; Saqr et al., 2023a, 2023b, 2023c), video data (Törmänen et al., 2022, 2023), and electrodermal activity (Törmänen et al., 2022, 2023). The data used in multi-channel sequence analysis is typically mutated into states to simplify the interpretation of the results. The states used in previous studies have explained the relationship between emotional valence with arousal (Törmänen et al., 2022), tactics used while studying online with automated assessments (Lopez-Pernas & Saqr, 2021), and engagement with achievement (Saqr et al., 2023a, 2023b, 2023c). The highest number of channels, totalling five, was used by Ouyang et al. (2023a, 2023b), who used five channels: social interactions, cognitive dimension, collaboration regulation, students' online behaviours, and the socio-emotional dimension. It is uncommon to use many channels as the number of clusters will increase multiplied by the number of states in each state. This will make it harder

to interpret the results meaningfully and create practical implications based on these findings.

Clustering is the standard step that the studies mentioned above have performed. Hidden Markov models (HMM) are used to identify hidden states in the data. In HMM, there are two types of states: observed and hidden. Hidden states are also known as latent states (Lopez-Pernas & Saqr, 2021). Sequence data are observed states, which can be understood as probabilistic functions of hidden latent states (Lopez-Pernas & Saqr, 2021). Hidden states cannot be observed directly, but it becomes possible to observe them through the sequence of observed states (Helske & Helske, 2019). In previous studies, HMM has been used to analyse unified strategies and the transition probabilities of students moving from a cluster to another cluster during studies (Lopez-Pernas & Saqr, 2021). Lopez-Pernas and Saqr (2021) found three hidden states of learning strategies combining the use of the LMS and an automated assessment tool. When the hidden states are analysed timewise, distinct trajectories can be found. In recent studies, this approach has been used to assess the fluctuation of socio-emotional interaction episodes and regulation of learning (Törmänen et al., 2023) and collaborative group patterns (Ouyang et al., 2023a, 2023b).

Motivation of the study

SRL is proven to be an essential element of learning, leading to better learning outcomes. However, there remains a significant gap in our understanding of how SRL interacts with student engagement over time in online courses. While previous studies have examined these constructs separately or as static measures, the dynamic interplay between engagement and SRL throughout a course has not been thoroughly investigated.

Our multi-channel approach represents a theoretical advancement in how we understand the relationship between engagement and self-regulation. While previous research has typically treated these constructs as separate or examined them through single data streams, our approach acknowledges their inherent interconnectedness as suggested by social cognitive theory (Bandura, 1986) and reinforced by recent theoretical developments (Panadero et al., 2019). By simultaneously tracking behavioural engagement through trace data and self-regulation through repeated self-reports, we can observe how improvements in one area might catalyse changes in the other, a dynamic highlighted in recent studies of creative and collaborative learning contexts (Greene et al., 2019; Malmberg et al., 2017). This bi-directional relationship is significant in online learning environments, where students must actively manage both their engagement with course materials and their regulatory processes. Our theoretical framework suggests that these processes are not merely parallel but interacting, with changes in engagement potentially triggering adjustments in self-regulation strategies and vice versa. This dynamic interaction helps explain why some students show improvement patterns while others remain consistently disengaged.

This study uses a multi-channel sequence analysis approach to address this knowledge gap, combining LMS log data with longitudinal self-reported SRL measures. This innovative method allows us to examine the co-evolution of engagement and SRL throughout an online course. We aim to identify distinct patterns and trajectories of student

behaviour, considering both engagement and SRL in parallel. This information provides a more nuanced understanding of how learners navigate online courses over time.

By uncovering these patterns, we can inform the development of more targeted and effective support strategies for online learners. This research builds upon previous work in learning analytics and SRL, extending it by (a) using a longitudinal approach to capture the dynamic nature of engagement and SRL, (b) combining behavioural trace data with self-reported measures for a more comprehensive view of the learning process, and (c) employing advanced analytical techniques to identify complex patterns and transitions in student behaviour.

The research questions of this study are:

RQ1: How do patterns of engagement and self-regulation co-evolve throughout an online project management course, and what temporal relationships exist between these constructs?

RQ2: What distinct clusters of student groupings emerge when analysing their engagement and self-regulation trajectories together, and what characteristics define these clusters?

These questions address critical gaps in our understanding of online learning. RQ1 moves beyond static measurements to capture the dynamic nature of learning processes, essential for developing responsive support systems in online education. RQ2 aims to identify distinct patterns of learning behaviour that can inform the design of targeted interventions, addressing the growing need for personalized support in online environments. Together, these questions contribute to our understanding of how students develop and maintain effective learning strategies in digital spaces, a crucial concern as higher education increasingly embraces online and hybrid modalities.

Methods

Context

The context of the study was an online project management course. The course was organised in a Finnish higher education institution (HEI), and the participants were first-year business students. The course logistics can be described asynchronous as a student could study the course at their own pace between preset start and end dates. All materials and assignments were available on the course platform once the course platform opened for students. The course was organised following the principles of formative evaluation. Therefore, students set personal learning goals at the beginning of the course. The numerical goal (scale: 0–5) determines the number of assignments students submit during the course; for the lowest passing grade (1), a student had to complete three assignments, whereas, for the highest grade (5), the required number of assignments was ten. For each assignment, predetermined requirements were set and available for students throughout the course to enable self-assessment. Once the requirements of the submitted assignments are met, the student achieves the grade they have set for themselves. Due to these evaluation principles, there is a natural variance in students' activity: some students aim for lower grades; therefore, they must do less in the course platform than those who aim for higher grades. This variance has to be considered when computing engagement indicators. The HEI has given a research permit for this study, and each student ($n=165$) gave informed consent for their participation. The research

permit did not allow the researchers to collect contextual data on the students. Therefore, it is not reported in this study.

Data collection and preparation

The data for this study were gathered from two sources: (a) trace data captured from a course organised in Moodle LMS and (b) survey data from the same course. These parallel datasets represent two perspectives on learning. LMS data (ID, timestamp, and the context of every student's action) illustrates the actions students take while studying online, while survey data shows the students' perceptions of their learning process. The survey data was captured at the beginning of the course when students set their goals for the course and every time a student submitted an assignment. As a result, a student aiming for grade 1 answered the survey 4 times (once in the beginning and three times when submitting the three assignments required for grade 1), and the student aiming for grade 5 answered the survey 11 times (once in the beginning and ten times when submitting the assignments). We used assignment submission timestamps as natural segmentation points for the LMS data analysis, as these represent significant transition points in students' learning processes. Research has shown that assignment submissions represent natural regulatory episodes in learning cycles, where students transition between different phases of self-regulated learning Järvelä & Hadwin, 2013; Winne & Hadwin, 2008). This segmentation approach aligns with the cyclical nature of SRL, where each segment represents a complete cycle: planning phase for the assignment, performance phase during task completion, and reflection phase upon submission. The same approach has been used in our previous study (Heikkinen et al., 2025). This progressive data collection approach resulted in a total of 1,164 survey responses across all 165 participants, with the number of responses per student varying based on their grade target and course completion level. Each survey response included ratings for all ten SRL variables on an 11-point Likert scale (1 = strongly disagree to 11 = strongly agree). We used the timestamps of the submissions as a cutting point to segment the trace data into corresponding parts of the course. This helps us to analyse the data granularly. The segments are this way represent the cyclical SRL process: each segment starts with the planning phase of the assignment, unfolds with the performing phase, and finally ends with the submission and reflecting phase, which should be used as a feed-forward for the next assignment and SRL cycle.

We use the LMS log data (62,070 log entries) to represent the student's engagement during the course and survey data to represent the SRL channel. LMS logs provided detailed interaction data capturing student engagement with the online learning environment. Each log entry contained four key elements: a precise timestamp of the interaction, an anonymised student identifier to maintain privacy while enabling individual-level analysis, the context of the interaction (such as course materials), and the specific event or action taken by the student (such as viewing content). See Appendix 1 for more details. The complete data set is available in OSF [OSF reference]. We used LMS log data to compute indicators representing engagement (Jovanović et al., 2021). The indicators included *ActiveDays* (the count of days a student had enrolled on LMS), *Assignment* (the count of student's interactions with assignment-related activities in LMS), *AvgSessionLen* (mean time of the length of student's LMS sessions), *Material* (the

count of a student’s interactions with the learning resources), *SessionCnt* (the count of sessions a student had taken in LMS), and *Survey* (the count of student’s interactions with the SRL survey related activities: e.g., answering or revisiting; in LMS).

For our survey design, we opted for single-item measures rather than traditional multi-item constructs. While lengthy surveys provide benefits for development and validity evaluation (Barnard et al., 2009; Broadbent et al., 2023), they create significant response burden during repeated measures in a course (Fisher et al., 2016). For longitudinal studies tracking competency development, short surveys are crucial (Beymer et al., 2022). Single-item surveys offer a feasible solution (Allen et al., 2022) for real-time reporting, as people struggle to accurately report development retrospectively (Solhan et al., 2009). However, it is important to note that this approach has potential validity limitations.

Our survey included 11 statements on SRL-related factors selected from validated instruments (Barnard et al., 2009; Broadbent et al., 2023; Jansen et al., 2017) to match the course context. This approach, previously used by Saqr and López-Pernas (2024), allows for capturing numerous observations even with small sample sizes. While this raises validity questions, research suggests single-item construct questionnaires can be as reliable as multi-item versions (Allen et al., 2022). Small sample sizes can be adequate when focusing on within-person variance (Maas & Hox, 2005; Martin et al., 2015; Raudenbush & Bryk, 2002).

The statements, their source per statement and the variables derived from them are displayed in Table 1. The statement variables were *AskHelp*, *Criteria*, *Deadlines*, *Prethink*, *Schedule*, *SelfAssess*, *Standards*, *Strategies*, *Summary*, *Time*, and *Location*. We performed a correlation test to check if there were correlations between survey variables. *Time* and *Location* variables had a very high correlation. Thus, we combined these variables by calculating a mean for these two variables and named the newly created variable *Undistracted*. This new variable was then used to replace variables

Table 1 The variables, original statements used to derive the variables, and the source of original statements

Variable	Statement	Source
Askhelp	"I am persistent in getting help from the instructor through e-mail."	OSLQ (Barnard et al., 2009)
Criteria	"I think about how I might improve my work by evaluating it against the marking criteria provided by the teacher."	SRL-O (Broadbent et al., 2023)
Deadlines	"I set realistic deadlines for learning."	SRL-O (Broadbent et al., 2023)
Prethink	"I think about what I really need to learn before I begin a task in this online course."	SOL-Q (Jansen et al., 2017)
Schedule	"I plan out my schedule each week, so I have the appropriate amount of time available for online study."	SRL-O (Broadbent et al., 2023)
SelfAssess	"I usually self-assess my performance once I finish."	SRL-O (Broadbent et al., 2023)
Standards	"I set standards for my assignments in online courses."	OSLQ (Barnard et al., 2009)
Strategies	"I am aware of what strategies I use when I study for this online course."	SOL-Q (Jansen et al., 2017)
Summary	"I summarise my learning in online courses to examine my understanding of what I have learned."	OSLQ (Barnard et al., 2009)
Undistracted	"I choose a time with few distractions for studying for my online courses." and "I choose the location where I study to avoid too much distraction."	OSLQ (Barnard et al., 2009) OSLQ (Barnard et al., 2009)

Time and *Location* in the survey data set. These ten variables represent the student's perception of learning regulation.

RQ1: The engagement and learning regulation states

To answer the first RQ, we first normalised both data sets to ensure comparability across different scales and course iterations using the *BBmisc* package (Bischl et al., 2022). This will set the mean value for each variable to zero, making comparing different variables using different scales easier. We then used the normalised data sets to detect states for both data sets. Figure 1, illustrates the process. The LMS data was used to detect an engagement state, and the survey data was used to detect the learning regulation state.

To identify distinct patterns in both engagement and self-regulation data, we employed Gaussian Parsimonious Clustering Models using the *MoEClust* package (Murphy & Murphy, 2020). This approach was chosen because it can effectively capture complex multivariate patterns in continuous data while accounting for different cluster shapes, orientations, and variance structures. For both engagement and self-regulation data, we compared multiple models with varying covariance structures (EII through VVV) using the Bayesian Information Criterion (BIC) to determine the optimal number of clusters and model complexity. This method provides robust model selection criteria and allows for thorough evaluation of cluster quality through measures such as entropy and posterior probabilities.

For the engagement indicators, the VVV model with three clusters was selected as optimal based on BIC, showing high entropy (0.879) and strong average posterior probabilities (>0.90) indicating good cluster separation. Similarly, for the self-regulation measures, a VVV model with two clusters provided the best fit according to BIC, also demonstrating high entropy (0.908) and reliable classification probabilities.

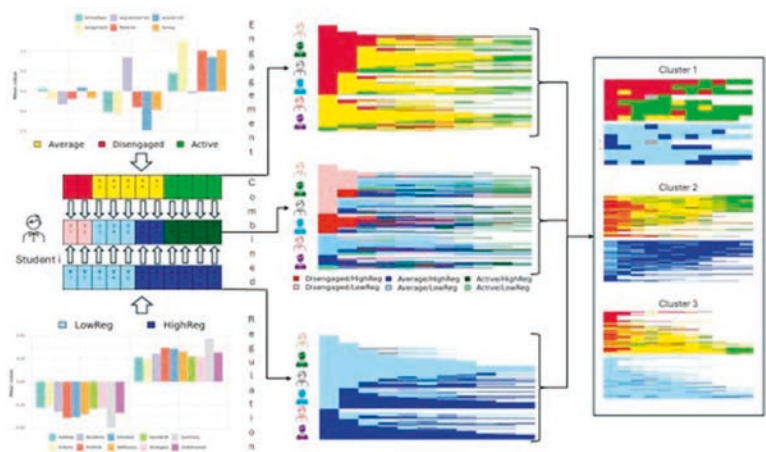


Fig. 1 The research process. On the left are the engagement and learning regulation states and an example of the sequence formation. In the middle are every student's index plots for engagement, regulation and combined sequences. On the right, similar sequences are clustered

These models effectively captured the underlying structure of both data channels while maintaining interpretability and statistical robustness.

From here on, the two sets of clusters are referred to as states. The LMS clusters are referred to as course engagement states with levels of *Active*, *Average*, and *Disengaged* (the same states are used in López-Pernas & Saqr, 2024), while survey clusters are referred to as learning regulation states, with levels of *LowReg* representing low regulation of learning and *HighReg* representing high regulation of learning.

After the states were determined, we created sequences to represent the evolution of learning during the course. This was done by dividing the course into 11 segments (Fig. 1, left centre). The segmentation was done with the help of survey submissions. We wanted to combine the two states to represent the parallel learning development on survey data and LMS. Thus, we used the timestamps of the survey submissions to cut the LMS data into segments, representing the learning development within the course. We used an interval between two survey submissions for each student from their trace data to calculate the LMS variables for each student for each segment individually. This approach enables the longitudinal division of the LMS course data following the individual progression of each student, still making it possible to make the between-student comparison. Sequences of engagement and self-regulation states were defined and visualised using the *TraMineR* package (Gabadinho et al., 2008). For quality assessment of the clustering solutions, we used the *WeightedCluster* package (Studer, 2013) to evaluate cluster stability and separation.

RQ2: Clustering similar trajectories

To answer the second RQ, we performed clustering, to identify patterns with similarities. Dissimilarities between sequences were computed using the Longest Common Subsequence (LCS) method through *TraMineR* (Gabadinho et al., 2008), to find patterns with longest identical sequences. Hierarchical clustering with Ward's method was then applied to these dissimilarities. We visualised sequences for both states independently (Fig. 1, top and bottom middle) and then combined the states (Fig. 1, centre) from both channels to be analysed. When combined, each state included information on both regulation and engagement. Thus, there were six states: *Active/HighReg*, *Active/LowReg*, *Average/HighReg*, *Average/LowReg*, *Disengaged/LowReg*, and *Disengaged/HighReg*. To illustrate the sequences, we used sequence index plots and sequence distribution plots.

We continued the analysis by clustering the multi-channel data (Fig. 1, right). Gaussian Parsimonious Clustering were fitted using the *MoE_clust* function from the *MoEClust* package (Murphy & Murphy, 2020). The optimal number of clusters was determined using the BIC. The optimal number of multi-channel clusters was found to be three. Hidden states within the multi-channel data were identified using the *seqHMM* package (Helske & Helske, 2019). This approach allowed for modelling latent states and their transitions over time. *MoEClust* (Murphy & Murphy, 2020) is powerful here, as it helps to choose a better model start by being more sensitive to where pattern starts. We built mixture hidden Markov models (MHMM) with two hidden states for each cluster. We fitted the model a hundred times to find the globally optimal solution. To check the reliability of the clustering, we checked that we reached the same solution with different starting values. Furthermore, we checked that the results were interpretable.

Finally, a transition network analysis was performed to visualise the transitions between different states in the sequences, using the *qgraph* package (Epskamp et al., 2012) for visualisation. This comprehensive methodological approach allowed for a nuanced examination of students' engagement and self-regulation patterns over time in the online learning environment.

Results

RQ1: Engagement, learning regulation, and combined states

The clustering analysis revealed distinct profiles of student engagement and self-regulation. According to BIC, the optimal model for LMS engagement indicators was a 3-component VVV model. The three clusters were labelled as *Average* (56.3%), *Disengaged* (23.4%), and *Active* (20.4%). The model showed high entropy (0.879) and average posterior probabilities (0.949, 0.905, 0.974 for the three clusters, respectively), indicating good cluster separation. The *Average* state represents learners whose engagement indicators (ActiveDays, Assignment, AvgSessionLen, Material, SessionCnt, and Survey) were clustered around the normalized mean values of these LMS trace data variables. These normalized values were calculated by standardizing each engagement indicator across all students, setting the mean to zero and standard deviation to one. The *disengaged* state had below-average values on the number of active days, interactions with assignment-related activities, learning resources, and surveys. Students in this state showed a distinctive pattern in their LMS usage: they logged in infrequently (low number of sessions) but when they did, they stayed logged in for extended periods (high mean session length), suggesting inefficient or irregular study patterns. In contrast, the *active* state represents learners who demonstrated consistent engagement through above-average values in active days, interactions with assignment-related activities, learning resources, and survey activities. These students exhibited a more systematic study pattern: they accessed the LMS frequently but had shorter, more focused sessions (below-mean session length). This pattern suggests they engaged in regular, purposeful learning activities rather than sporadic, lengthy sessions.

There are two states for learning regulation. For survey responses, a 2-component VVV model was optimal. The two clusters were labelled as *LowReg* (48.2%) and *HighReg* (51.8%). This model also showed high entropy (0.908) and average posterior probabilities (0.978, 0.972), suggesting a clear cluster distinction. *LowReg* represents learners whose survey responses were below the normalised mean across all ten SRL variables (AskHelp, Criteria, Deadlines, Prethink, Schedule, SelfAssess, Standards, Strategies, Summary, and Undistracted). Each variable was measured on a Likert scale, with responses collected at the beginning of the course and before each assignment submission. Similarly, *HighReg* represents learners whose survey responses were consistently above the normalised mean on all these SRL variables.

When the two channels are combined, we get the third state, which is the combined state. For the combined state the labels are *Active/HighReg*, *Active/LowReg*, *Average/HighReg*, *Average/LowReg*, *Disengaged/LowReg*, and *Disengaged/HighReg*. The sequence analysis provided insights into the transitions between different states. In Fig. 2, we can see the sequences of learners. Each slice in the sequence plot shows an idiographic sequence (Ito et al., 2024) from the same individual across time to devise person-specific

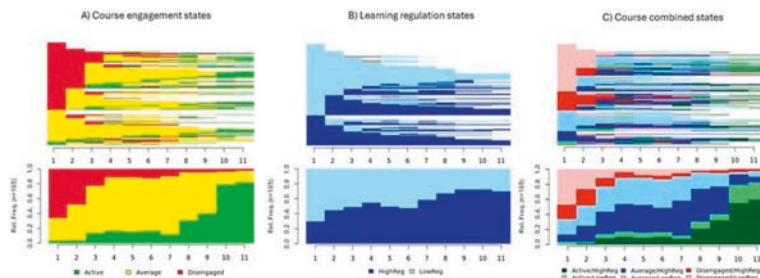


Fig. 2 Sequence index plot (top) and sequence distribution plot (bottom) for **A** course engagement states, **B** learning regulation states, and **C** course combined states

precise information about the very individual. The coloured blocks represent the state: in (A) the engagement state and in (B) the learning regulation state. When these states are combined, (C) represents the course combined state. The numbers at the bottom of the Fig. 2 represent the course segment. Thus, in the sequence index plot (upper plot), the continuum from left to right represents the sequence of the states for each student during the course. In the sequence distribution plot (lower plot), the y-axis represents the proportion of students in a particular state during the course segment.

The sequence analysis revealed distinct patterns of engagement and self-regulation over time. The state distribution plots show that while some students maintained consistent engagement and self-regulation levels, others exhibited more variable patterns.

RQ2: Clusters of similar trajectories

Next, we performed clustering to identify similar patterns in learners' sequences. Hierarchical clustering of the multi-channel sequences suggested a 3-cluster solution based on the quality measures. The resulting clusters showed distinct patterns of engagement and self-regulation. The hidden Markov models further elucidated the underlying patterns in the data.

Within the low-regulating improvers cluster (Fig. 3, $n=8$, 4.85%), the students perceive themselves as low-learning regulators (middle left). At the beginning of the course, they are disengaged and become active later on (upper left). For some learners, this happens already at the second segment of the course, while some of them change state to active at the midpoint of the course. For hidden states (lower left), state 1 suggests the learner is disengaged, and vice versa; state 2 represents Average or engaged learners. From the transition plot on the figure's right side, we can see that the initial state is state 1 with a probability of 100% (below the pie chart), whereas 0% starts at state 2. The transition probabilities between states (the edges between states on the right) are much higher from state 1 to state 2 (33%) than from state 1 to state 2 (2.9%).

Disengaged low regulators (Fig. 4, $n=74$, 44.85%) belong to a cluster where almost every student perceives themselves as having low skills in learning regulation. At the beginning of the course, the engagement channel is mostly disengaged or average. These represent state 2 for this cluster. Some students transit to state 1, which improves engagement from disengaged to average or active engagement. Initial

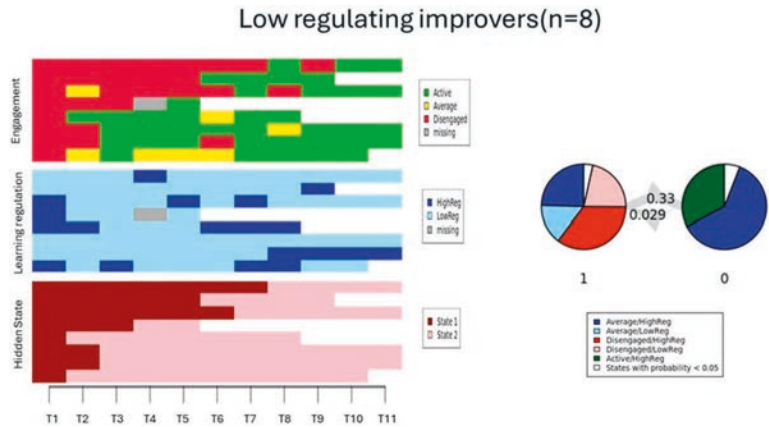


Fig. 3 Characteristics of Low regulating improvers. On the top left is an index plot of the engagement states; in the middle are learning regulation states; and at the bottom are hidden states. On the right-hand side is the transition diagram

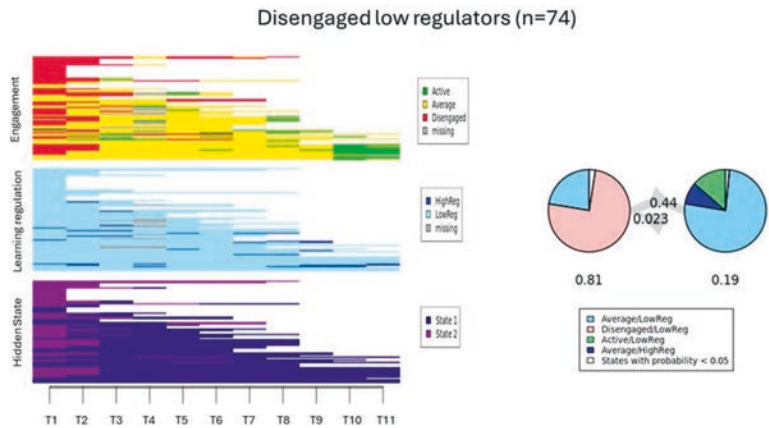


Fig. 4 Characteristics of Disengaged Low Regulators. On the top left is an index plot of the engagement states; in the middle are learning regulation states, and at the bottom are hidden states. On the right-hand side is the transition diagram

probabilities are 81%/19%, and the transition is more likely (44%) from state 1 to state 2 than from state 2 to state 1 (2.3%).

The last cluster (Fig. 5, $n = 83$, 50.30%) represents learners with high learning regulation skills. At the beginning of the course, the first channel shows that they are disengaged, but their engagement status changes from disengaged to average and then towards the end of the course to active. The learning regulation channel shows a similar improvement from low regulation to high regulation during the course. State 1 represents students who are disengaged with low regulation and then move to State 2, which can be described as students with average or active engagement and high

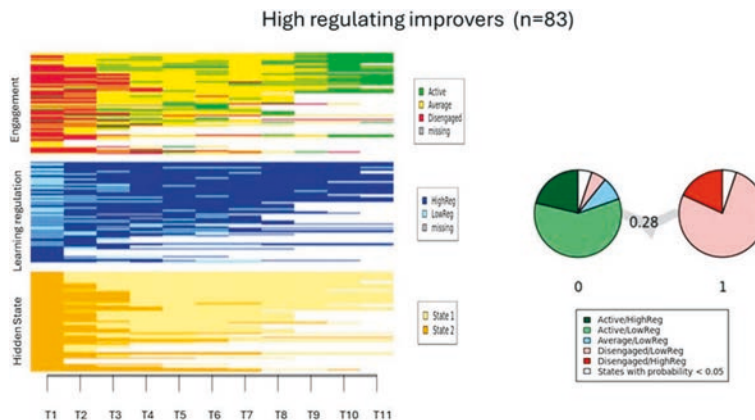


Fig. 5 Characteristics of High regulating improvers. On the top left is an index plot of the engagement states, in the middle learning regulation states and at the bottom hidden states. On the right-hand side is the transition diagram

learning regulation. All students in this cluster start with state 2 (initial probability is 100%), and the transitions occur only from state 2 to state 1 (28%).

Discussion

This study employed a multi-channel clustering approach to analyse longitudinal student engagement and self-regulation data in an online project management course. By combining trace data from an LMS with self-reported survey data on SRL, we were able to identify distinct patterns of student behaviour and their evolution over time. Our findings provide valuable insights into the complex interplay between engagement, self-regulation, and learning outcomes in online educational environments.

Low-regulating improvers (4.85%, $n=8$) demonstrated an interesting pattern of behavioural adaptation despite consistently low self-reported regulation skills. These students began in a disengaged state but showed remarkable improvement in their engagement levels over time, as evidenced by their transition probabilities (33% probability of moving from state 1 to state 2). Their improvement appears to be driven by behavioural adaptations rather than enhanced regulatory skills—they learned to engage more effectively with the course platform even while continuing to report low confidence in their self-regulation abilities. This suggests that these students may have developed practical strategies for course engagement (such as regular login patterns and consistent interaction with course materials) without necessarily developing the metacognitive awareness of their learning strategies. This pattern indicates that practical engagement behaviours can improve even when students don't perceive themselves as strong self-regulators. The importance of emphasizing the regularity of course activity would be beneficial for the students in order to help them improve their engagement within course.

Disengaged low regulators (44.85%, $n=74$), the second largest cluster, showed the most concerning pattern. These students exhibited both low engagement and low self-regulation throughout the course, with a high probability (81%) of remaining in their

initial disengaged state. Their sequence patterns reveal brief periods of average engagement interspersed with disengagement, suggesting unsuccessful attempts to establish consistent study routines. The transition probability from state 2 to state 1 was very low (2.3%), indicating that once these students became disengaged, they struggled to re-engage with the course. This pattern suggests that early intervention is crucial for this group—once they fall into disengaged patterns, they are unlikely to recover without external support.

High-regulating improvers (50.30%, $n=83$), the largest cluster, showed the most positive trajectory. While these students initially appeared disengaged, they demonstrated a consistent pattern of improvement in both engagement and self-regulation. Their sequence patterns reveal a gradual progression from disengaged to average and finally to active engagement states, accompanied by increasing self-regulation scores. Interestingly, all students in this cluster started in state 2 (100% initial probability), suggesting they had the foundational skills necessary for improvement. The 28% transition probability from state 2 to state 1 indicates that even these successful students experienced temporary setbacks, but they were able to recover and maintain their overall positive trajectory. This resilience in the face of temporary setbacks appears to be a key distinguishing feature of this group. To help all students, the interventions should focus on prompts helping students to acknowledge their current state of SRL and identify their needs on these skills. Once the metacognitive needs are identified, the interventions can be aimed to help students in the skills they lack sufficient toolkit to help their SRL.

The transition patterns across clusters reveal important insights about learning trajectories in online environments. The high-regulating improvers' ability to maintain positive trajectories despite occasional setbacks suggests that strong self-regulation skills serve as a buffer against disengagement. In contrast, the disengaged low regulators' difficulty in transitioning out of disengaged states highlights how initial struggles with self-regulation can create self-reinforcing cycles of disengagement. The low-regulating improvers demonstrate that behavioural engagement can improve even without strong self-reported regulation skills, suggesting multiple potential pathways for student success in online learning environments.

Our findings align seamlessly with previous research, reinforcing the importance of self-regulation in online learning environments (Broadbent et al., 2023; Järvelä et al., 2019). Our results both support and extend previous findings on student engagement. While we observed that prior engagement can indeed predict future engagement, aligning with studies by Gottfried et al. (2009) and Wylie and Hodgen (2012), our analysis revealed a more nuanced picture. We found that engagement is not a continuously increasing phenomenon, confirming Wigfield et al. (2007) observations. Our study goes further by identifying distinct patterns of engagement evolution, including "improvers" who show increasing engagement over time and "disengaged low regulators" who fluctuate between disengaged and average states. This highlights the dynamic nature of engagement and the potential for positive change even among initially disengaged students.

While previous longitudinal studies, such as Saqr et al., (2023a, 2023b, 2023c), have examined engagement with academic achievement; our study broadens this perspective by simultaneously considering engagement and self-regulation. This dual focus allows us

to uncover complex interactions between these factors that may not be apparent when studying engagement alone. For instance, we found that students with similar engagement levels may differ significantly in their self-regulation skills, leading to divergent learning outcomes. This multifaceted approach provides educators with a more comprehensive understanding of student learning processes and potential areas for targeted support. The observed improvement in engagement and self-regulation for some students over the course duration supports the notion that these skills can be developed over time (Wolters et al., 2023). This aligns with the conceptual framework proposed by Järvelä and Hadwin (2024), suggesting that certain events or experiences during a course can trigger changes in regulatory processes.

Our study expands the field of sequence mining in education in several ways. While previous studies have used sequence analysis to examine relationships between learning patterns and performance (Kinnebrew & Biswas, 2012; Matcha et al., 2020b; Saqr et al., 2023a, 2023b, 2023c), compare course delivery patterns (Matcha et al., 2020a), or investigate associations with dropout (Saqr & López-Pernas, 2021b), our study integrates engagement and self-regulation sequences to provide a more holistic view of the learning process. Building on the work of López-Pernas and Saqr (2024), who identified three clusters of engagement behaviour; our study reveals how these engagement patterns interact with self-regulation strategies over time. This combined analysis of engagement and self-regulation sequences offers a more nuanced understanding of student learning trajectories and potential intervention points.

Our results contribute to the growing literature on LA and multi-channel sequence analysis in educational research (Saqr et al., 2023a, 2023b, 2023c). Combining LMS trace data with self-reported SRL data provides a more holistic view of the learning process than studies relying solely on self-reported measures or behavioural data alone. Our study significantly improves upon previous approaches to analysing SRL and engagement. Unlike studies such as Wang et al. (2024) that relied on self-reported data collected at the end of the research experiment, our approach combines LA data for engagement assessment with self-reported SRL data collected multiple times throughout the course. This multi-channel, temporal approach provides a more granular and dynamic view of how engagement and SRL co-evolve during learning. By capturing these constructs at multiple time points, we can track individual trajectories and identify critical periods where interventions might be most effective.

Our analysis revealed several unexpected findings that challenge common assumptions about online learning. First, the existence of a small but distinct group of 'Low-regulating improvers' (4.85%) demonstrates that students can enhance their engagement patterns without corresponding improvements in self-reported regulation skills. This challenges traditional models that posit a strong interdependence between engagement and self-regulation (Panadero, 2017; Winne & Hadwin, 2008). This suggests that behavioural adaptation can occur independently of metacognitive development, aligning with recent work on the dissociation between behavioural and cognitive aspects of learning (Törmänen et al., 2023). Second, even the most successful students in the 'High-regulating improvers' cluster began in a disengaged state and showed significant probability (28%) of transitioning back to lower engagement states, supporting Wigfield et al.'s (2007) observation that engagement is not continuously

increasing. This challenges the assumption that high self-regulation skills guarantee consistent engagement (Zimmerman & Schunk, 2011). Finally, the stark bimodal distribution of students between 'Disengaged low regulators' (44.85%) and 'High-regulating improvers' (50.30%) suggests a potential Matthew effect in online learning (Stanovich, 2009), where students tend to either succeed or struggle with few middle-ground cases. This polarization effect has been observed in recent online learning studies (López-Pernas & Saq, 2024) but not at such a pronounced level. These findings highlight the complexity of online learning trajectories and suggest that traditional assumptions about the relationship between engagement and self-regulation may need to be reconsidered (Järvelä & Hadwin, 2024).

A notable methodological contribution of our study is the data segmentation based on the number of assignments completed by each student. This approach allows for a more equitable comparison of students with different achievement levels and course progression rates. From the sequence plots, we can see that some students have achieved lower grades (shorter sequences) but still had high engagement or SRL. If this analysis had been done on a course level, these students would have got lower indicators than those who continuously put effort into their learning. By anchoring our analysis to individual student progress rather than arbitrary time points, we can more accurately capture the relationship between engagement, self-regulation, and learning outcomes across diverse student populations. This method is particularly valuable in asynchronous online courses where students may progress at different rates, providing a more nuanced understanding of learning trajectories than traditional time-based analyses.

Identifying distinct student clusters has several implications for educational practice and instructional design. First, knowing which cluster a student belongs to can help educators provide more personalised support. For instance, disengaged low regulators might benefit from early interventions focused on developing self-regulation skills and increasing engagement. In order to achieve this, systematic analysis of log data in parallel with surveys should be performed to assess the level of engagement and SRL. This requires the analysis of students engagement analysing LMS data (e.g., when a student has logged in most recently) and assessing their learning skills (e.g., with a survey in the beginning of the course). When either one is absent, interventions should be performed as early as possible to offer timely support. Second, the results can be used to adapt course design. The observed transitions between states suggest that course design could be adapted to better support students at different stages. For example, additional scaffolding could be provided early in the course when many students show low engagement and self-regulation. Engagement information can be computed automatically with the log data. When low engagement is identified, support should be performed either automatically or by the instructor contacting a student with low engagement. Third, constant support for self-regulation. The improvement seen in the improver clusters underscores the importance of explicitly supporting the development of self-regulation skills throughout the course. One intervention is not necessarily enough. Instead, the intervention strategy should be systematic and constant throughout the course. Finally, the granular information can enable the creation of early warning systems.

The clustering approach used in this study could be incorporated into early warning systems to identify at-risk students based on their engagement and self-regulation patterns. In order to execute these practical implications, policy makers should allocate the necessary resources to LA in order to these functions to be able to provide timely insights helping to identify students in need.

Limitations and future research

While this study provides valuable insights, several limitations should be acknowledged. First, the study was conducted in one institution's online project management course. Future research should explore the generalizability of these findings across different subjects, institutions, and cultural contexts. The learning patterns might be different in different contexts. The comparison of the different contexts might help us to understand if the identified patterns are typical only for business students in Finnish cultural context. Second, while combining self-reported data with trace data strengthens our approach, self-reported measures may be subject to biases. Future studies could incorporate additional data sources, such as think-aloud protocols or eye-tracking data, to further validate SRL measurements. Third, the causal relationships should be considered. While we identified patterns and associations, our study design does not allow for causal inferences. Experimental studies could help establish causal relationships between interventions, self-regulation, engagement, and learning outcomes. Fourth, the single item measures used in this study are not a validated approach and needs further research to assess the validity of single-item measures against the longer surveys. Fifth, this data did not collect contextual data of the students. To assess the relevancy of the results in different contexts, this kind of information would be important. This limits the generalisability of the findings of the current research. In future studies, this should be acknowledged to set the results of the study in right context and help to identify the relation of context with the identified patterns. Finally, the focus should be put on long-term outcomes. Our study focused on patterns within a single course. Longitudinal studies spanning multiple courses or entire degree programs could provide insights into the long-term development of engagement and self-regulation skills.

Future research could also explore the following areas. As Järvelä and Hadwin (2024) have suggested that the triggers are important factors in SRL. The specific triggers that lead to transitions between states, particularly for students who show improvement over time, should be studied to find the optimal triggers leading to a desired response. This should be done by examining the effectiveness of targeted interventions for identified clusters. This would help to overcome the challenge of offering one-size-fits-all interventions for every student, even though we already know the students are heterogenic. This study used two data channels. Yet, other factors, such as motivation, prior knowledge, or social interactions, shape engagement and self-regulation patterns that should be examined with multi-channel studies. Relating the current findings with students' self-efficacy beliefs would also be essential. This will help estimate if we aim to support students who perceive they would need it to meet their goals.

Conclusion

This study has provided valuable insights into the complex dynamics of student engagement and self-regulated learning in online educational environments. By employing a multi-channel clustering approach to analyse longitudinal data, we have uncovered distinct patterns and trajectories that offer a nuanced understanding of how students study in online courses over time.

Our findings reveal three primary clusters of students based on their engagement and self-regulation patterns: Low-regulating improvers, Disengaged low regulators, and High-regulating improvers. These clusters highlight the diversity of learning trajectories and underscore the need for targeted support strategies in online education. Identifying these patterns allows for a more sophisticated approach to student support, moving beyond one-size-fits-all interventions to more personalised and timely assistance.

In conclusion, this research marks a significant advancement in our understanding of online learning dynamics through its innovative application of multi-channel sequence analysis. By uncovering the complex interplay between engagement and self-regulation, we have revealed how different student groups navigate online learning environments in distinctly different ways, challenging traditional one-size-fits-all approaches to online education. The identification of three distinct student clusters—Low-regulating improvers, Disengaged low regulators, and High-regulating improvers—provides concrete evidence that online learning trajectories are both diverse and malleable, suggesting opportunities for targeted intervention and support. These findings have immediate practical implications for the design of online learning environments, the development of early warning systems, and the implementation of personalized support strategies. As higher education continues its digital transformation, with online and blended learning becoming increasingly prevalent, our methodological framework and empirical findings provide a robust foundation for creating more adaptive, responsive, and inclusive learning environments. This work not only advances our theoretical understanding of how students learn online but also offers practical pathways for institutions to better support the diverse needs of their student populations, ultimately contributing to more equitable and effective online education practices.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s41239-025-00523-3>.

Supplementary Material 1.

Acknowledgements

The first author would like to thank the LAB-D program for supporting the study. The authors wish to acknowledge CSC – IT Center for Science, Finland, for computational resources.

Author contributions

All authors contributed to the study's conception and design. The first author performed material preparation and data collection, and the first and the second author performed analysis. The first author wrote the first draft of the manuscript, and all authors commented on previous versions.

Funding

The paper is co-funded by the Academy of Finland (Suomen Akatemia) Research Council for Natural Sciences and Engineering for the project "Towards precision education: Idiographic learning analytics (TOPEILA)", Decision Number 350560 which was received by M Saqr.

Availability of data and materials

The datasets generated and analysed during the current study are available in the Open Science Framework repository, <https://osf.io/kcjm2/>.

Declarations

Competing interests

The authors declare that they have no competing interests.

Received: 18 November 2024 Accepted: 11 March 2025

Published online: 08 April 2025

References

- Abbott, A. (1983). Sequences of social events: Concepts and methods for the analysis of order in social processes. *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 16(4), 129–147. <https://doi.org/10.1080/01615440.1983.10594107>
- Ali, W. (2020). Online and remote learning in higher education institutes: A necessity in light of COVID-19 pandemic. *Higher Education Studies*, 10(3), 16. <https://doi.org/10.5539/hes.v10n3p16>
- Allen, M. S., Iliescu, D., & Greiff, S. (2022). Single item measures in psychological science. *European Journal of Psychological Assessment*, 38(1), 1–5. <https://doi.org/10.1027/1015-5759/a000699>
- Almoslamani, Y. (2022). The relationship between self-regulation learning and online learning adoption. *Cypriot Journal of Educational Sciences*, 17(6), 2117–2126. <https://doi.org/10.18844/cjes.v17i6.7550>
- Bacci, S., & Bertaccini, B. (2022). A mixture hidden Markov model to mine Students' University curricula. *Data*, 7(2), 25. <https://doi.org/10.3390/data7020025>
- Bandura, A. (1986). *Social foundations of thought and action: a social cognitive theory* (Vol. 16). Prentice-Hall Inc.
- Barnard, L., Lan, W. Y., To, Y. M., Paton, V. O., & Lai, S.-L. (2009). Measuring self-regulation in online and blended learning environments. *The Internet and Higher Education*, 12(1), 1–6. <https://doi.org/10.1016/j.iheduc.2008.10.005>
- Beymer, P. N., Ferland, M., & Flake, J. K. (2022). Validity evidence for a short scale of college students' perceptions of cost. *Current Psychology*, 41(11), 7937–7956. <https://doi.org/10.1007/s12144-020-01218-w>
- Bischi, B., Lang, M., Bossek, J., Horn, D., Richter, J., & Surmann, D. (2022). *Miscellaneous Helper Functions for B. Bischi*. <https://github.com/berndbischi/BBmisc/issues>
- Boekaerts, M., & Corno, L. (2005). Self-regulation in the classroom: A perspective on assessment and intervention. *Applied Psychology*, 54(2), 199–231. <https://doi.org/10.1111/j.1464-0597.2005.00205.x>
- Borup, J., Graham, C. R., West, R. E., Archambault, L., & Spring, K. J. (2020). Academic communities of engagement: An expansive lens for examining support structures in blended and online learning. *Educational Technology Research and Development*, 68(2), 807–832. <https://doi.org/10.1007/s11423-020-09744-x>
- Bremer, H., & Kaufmann, O. (1927). Die Bekämpfung der Rübenfliege (*Pegomyia hyoscyami* Pz.) mit Fluornatrium und Kieselfluornatrium. *Anzeiger Für Schädlingskunde*, 3(2), 13–15. <https://doi.org/10.1007/BF02338175>
- Broadbent, J., Panadero, E., Lodge, J. M., & Fuller-Tyszkiewicz, M. (2023). The self-regulation for learning online (SRL-O) questionnaire. *Metacognition and Learning*, 18(1), 135–163. <https://doi.org/10.1007/s11409-022-09319-6>
- Churchill, D., King, M., Webster, B., & Fox, B. (2013). Integrating learning design, Interactivity, and technology. *ASCLITE*. <https://doi.org/10.14742/apubs.2013.1518>
- Cleary, T. J., Slomp, J., & Pawlo, E. R. (2021). Linking student self-regulated learning profiles to achievement and engagement in mathematics. *Psychology in the Schools*, 58(3), 443–457. <https://doi.org/10.1002/pits.22456>
- Cristea, T., Snijders, C., Matzat, U., & Kleingeld, A. (2024). Unobtrusive measurement of self-regulated learning: A clickstream-based multi-dimensional scale. *Education and Information Technologies*, 29(11), 13465–13494. <https://doi.org/10.1007/s10639-023-12372-6>
- Dent, A. L., & Koenka, A. C. (2016). The relation between self-regulated learning and academic achievement across childhood and adolescence: A meta-analysis. *Educational Psychology Review*, 28(3), 425–474. <https://doi.org/10.1007/s10648-015-9320-8>
- Dever, D. A., Sonnenfeld, N. A., Wiedbusch, M. D., Schmorow, S. G., Amon, M. J., & Azevedo, R. (2023). A complex systems approach to analyzing pedagogical agents' scaffolding of self-regulated learning within an intelligent tutoring system. *Metacognition and Learning*, 18(3), 659–691. <https://doi.org/10.1007/s11409-023-09346-x>
- Efkliades, A. (2011). Interactions of metacognition with motivation and affect in self-regulated learning: The MASRL model. *Educational Psychologist*, 46(1), 6–25. <https://doi.org/10.1080/00461520.2011.538645>
- Epskamp, S., Cramer, A. O. J., Waldorp, L. J., Schmittmann, V. D., & Borsboom, D. (2012). qgraph: Network visualizations of relationships in psychometric data. *Journal of Statistical Software*, 48(4), 1–18. <https://doi.org/10.18637/jss.v048.i04>
- Estévez, I., Rodríguez-Llorente, C., Piñeiro, I., González-Suárez, R., & Valle, A. (2021). School engagement, academic achievement, and self-regulated learning. *Sustainability*, 13(6), 3011. <https://doi.org/10.3390/su13063011>
- Fincham, E., Whitelock-Wainwright, A., Kovanović, V., Joksimović, S., van Staalduinen, J.-P., & Gašević, D. (2019). Counting clicks is not enough. *Proceedings of the 9th International Conference on Learning Analytics & Knowledge*, 501–510. <https://doi.org/10.1145/3303772.3303775>
- Fincham, E., Gasevic, D., Jovanovic, J., & Pardo, A. (2019a). From study tactics to learning strategies: An analytical method for extracting interpretable representations. *IEEE Transactions on Learning Technologies*, 12(1), 59–72. <https://doi.org/10.1109/TLT.2018.2823317>
- Fisher, G. G., Matthews, R. A., & Gibbons, A. M. (2016). Developing and investigating the use of single-item measures in organizational research. *Journal of Occupational Health Psychology*, 21(1), 3–23. <https://doi.org/10.1037/a0039139>

- Fournier-Viger, P., Chun, J., Lin, W., Kiran, R. U., Koh, Y. S., & Thomas, R. (2017). A survey of sequential pattern mining. *Data Science and Pattern Recognition*, 1(1), 54.
- Gabadinho, A., Ritschard, G., Studer, M., & Müller, N. S. (2008). Mining sequence data in R with the TraMineR package: A user's guide. *Department of Econometrics and Laboratory of Demography, University of Geneva, Switzerland*. <http://Mephisto.Unige.Ch/Traminer.1>
- Gottfried, A. E., Marcoulides, G. A., Gottfried, A. W., & Oliver, P. H. (2009). A latent curve model of parental motivational practices and developmental decline in math and science academic intrinsic motivation. *Journal of Educational Psychology*, 101(3), 729–739. <https://doi.org/10.1037/a0015084>
- Greene, J. A., Freed, R., & Sawyer, R. K. (2019). Fostering creative performance in art and design education via self-regulated learning. *Instructional Science*. <https://doi.org/10.1007/s11251-018-9479-8>
- Guan, R., Raković, M., Chen, G., & Gašević, D. (2024). How educational chatbots support self-regulated learning? A systematic review of the literature. *Education and Information Technologies*. <https://doi.org/10.1007/s10639-024-12881-y>
- Hadwin, A. F., Järvelä, S., & Miller, M. (2011). 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* (pp. 65–84). Routledge.
- Heikkinen, S., Cristea, T., Saqr, M., Malmberg, J., Kleingeld, A., Snijders, C., Matzat, U., & Tedre, M. (2024). Sequence analysis and process mining perspectives to goal setting: What distinguishes business students with high and low self-efficacy beliefs? *Smart Learning Environments*, 11(1), 40. <https://doi.org/10.1186/s40561-024-00327-4>
- Heikkinen, S., Saqr, M., Malmberg, J., & Tedre, M. (2023). Supporting self-regulated learning with learning analytics interventions—a systematic literature review. *Education and Information Technologies*, 28(3), 3059–3088. <https://doi.org/10.1007/s10639-022-11281-4>
- Heikkinen, S., Saqr, M., Malmberg, J., & Tedre, M. (2025). The interplay of engagement and learning regulation in online learning. In *Proceedings of TEEM 2024. Technological Ecosystems for Enhancing Multiculturality, 23–25 October 2024, Alicante, Spain. Lecture Notes in Educational Technology*. Springer, Singapore.
- Helske, S., & Helske, J. (2019). Mixture hidden markov models for sequence data: The seqHMM package in R. *Journal of Statistical Software*. <https://doi.org/10.18637/jss.v088.i03>
- Henrie, C. R., Halverson, L. R., & Graham, C. R. (2015). Measuring student engagement in technology-mediated learning: A review. *Computers and Education*, 90, 36–53. <https://doi.org/10.1016/j.compedu.2015.09.005>
- Hensley, L. C., Iaconelli, R., & Wolters, C. A. (2022). “This weird time we’re in”: How a sudden change to remote education impacted college students’ self-regulated learning. *Journal of Research on Technology in Education*. <https://doi.org/10.1080/15391523.2021.1916414>
- Hilpert, J. C., Greene, J. A., & Bernacki, M. (2023). Leveraging complexity frameworks to refine theories of engagement: Advancing self-regulated learning in the age of artificial intelligence. *British Journal of Educational Technology*, 54(5), 1204–1221. <https://doi.org/10.1111/bjet.13340>
- Ilfenthaler, D., & Yau, J.Y.-K. (2020). Utilising learning analytics to support study success in higher education: A systematic review. *Educational Technology Research and Development*, 68(4), 1961–1990. <https://doi.org/10.1007/s11423-020-09788-z>
- Ito, H., López-Pemas, S., & Saqr, M. (2024). A scoping review of idiographic research in education: Too little, but not too late. *IEEE International Conference on Advanced Learning Technologies (ICALT)*, 2024, 10–12. <https://doi.org/10.1109/ICALT61570.2024.00010>
- Jansen, R. S., van Leeuwen, A., Janssen, J., Kester, L., & Kalz, M. (2017). Validation of the self-regulated online learning questionnaire. *Journal of Computing in Higher Education*, 29(1), 6–27. <https://doi.org/10.1007/s12528-016-9125-x>
- Järvelä, S., & Hadwin, A. F. (2013). New frontiers: Regulating learning in CSDL. *Educational Psychologist*. <https://doi.org/10.1080/00461520.2012.748006>
- Järvelä, S., & Hadwin, A. (2024). Triggers for self-regulated learning: A conceptual framework for advancing multimodal research about SRL. *Learning and Individual Differences*, 115, 102526. <https://doi.org/10.1016/j.lindif.2024.102526>
- Järvelä, S., Järvenoja, H., & Malmberg, J. (2019). Capturing the dynamic and cyclical nature of regulation: Methodological progress in understanding socially shared regulation in learning. *International Journal of Computer-Supported Collaborative Learning*, 14(4), 425–441. <https://doi.org/10.1007/s11412-019-09313-2>
- Jin, S.-H., Im, K., Yoo, M., Roll, I., & Seo, K. (2023). Supporting students’ self-regulated learning in online learning using artificial intelligence applications. *International Journal of Educational Technology in Higher Education*, 20(1). <https://doi.org/10.1186/s41239-023-00406-5>
- Jivet, I., Scheffel, M., Specht, M., & Drachsler, H. (2018). License to evaluate. *Proceedings of the 8th International Conference on Learning Analytics and Knowledge*. 31–40. <https://doi.org/10.1145/3170358.3170421>
- Jovanović, J., Gašević, D., Dawson, S., Pardo, A., & Mirriahi, N. (2017). Learning analytics to unveil learning strategies in a flipped classroom. *The Internet and Higher Education*, 33, 74–85. <https://doi.org/10.1016/j.iheduc.2017.02.001>
- Jovanović, J., Saqr, M., Joksimović, S., & Gašević, D. (2021). Students matter the most in learning analytics: The effects of internal and instructional conditions in predicting academic success. *Computers and Education*, 172, 104251. <https://doi.org/10.1016/j.compedu.2021.104251>
- Kinnebrew, J. S., & Biswas, G. (2012). Identifying learning behaviors by contextualizing differential sequence mining with action features and performance evolution. *Proceedings of the 5th International Conference on Educational Data Mining, EDM 2012*.
- Lapitan, L. D. S., Tiangco, C. E., Sumalinog, D. A. G., Sabarillo, N. S., & Diaz, J. M. (2021). An effective blended online teaching and learning strategy during the COVID-19 pandemic. *Education for Chemical Engineers*, 35, 116–131. <https://doi.org/10.1016/j.ece.2021.01.012>
- Larrabee Sønderlund, A., Hughes, E., & Smith, J. (2019). The efficacy of learning analytics interventions in higher education: A systematic review. *British Journal of Educational Technology*, 50(5), 2594–2618. <https://doi.org/10.1111/bjet.12720>
- Li, Y., Chen, D., & Deng, X. (2024). The impact of digital educational games on student’s motivation for learning: The mediating effect of learning engagement and the moderating effect of the digital environment. *PLoS ONE*, 19(1), e0294350. <https://doi.org/10.1371/journal.pone.0294350>

- Lin, F., Lim, L., Ho, Y. Y., Tan, B. Z., & Lim, W. Y. (2023). Validating and applying an adapted OSLO to examine adult learners' online self-regulation. *Higher Education Research and Development*, 42(7), 1684–1699. <https://doi.org/10.1080/07294360.2023.2183938>
- Lodge, J. M., Panadero, E., Broadbent, J., & De Barba, P. G. (2019). Supporting self-regulated learning with learning analytics. In J. Lodge, J. C. Horvath, & L. Corrin (Eds.), *Learning analytics in the classroom* (pp. 45–55). Routledge.
- Lopez-Pernas, S., & Saqr, M. (2021). Bringing synchrony and clarity to complex multi-channel data: A learning analytics study in programming education. *IEEE Access*, 9, 166531–166541. <https://doi.org/10.1109/ACCESS.2021.3134844>
- López-Pernas, S., & Saqr, M. (2024). How the dynamics of engagement explain the momentum of achievement and the inertia of disengagement: A complex systems theory approach. *Computers in Human Behavior*, 153, 108126. <https://doi.org/10.1016/j.chb.2023.108126>
- López-Pernas, S., Saqr, M., Gordillo, A., & Barra, E. (2023). A learning analytics perspective on educational escape rooms. *Interactive Learning Environments*, 31(10), 6509–6525. <https://doi.org/10.1080/10494820.2022.2041045>
- López-Pernas, S., Saqr, M., & Viberg, O. (2021). Putting it all together: combining learning analytics methods and data sources to understand Students' Approaches to learning programming. *Sustainability*, 13(9), 4825. <https://doi.org/10.3390/su13094825>
- Maas, C. J. M., & Hox, J. J. (2005). Sufficient sample sizes for multilevel modeling. *Methodology*, 1(3), 86–92. <https://doi.org/10.1027/1614-2241.1.3.86>
- Malmberg, J., Järvelä, S., & Järvenoja, H. (2017). Capturing temporal and sequential patterns of self-, co-, and socially shared regulation in the context of collaborative learning. *Contemporary Educational Psychology*. <https://doi.org/10.1016/j.cedpsych.2017.01.009>
- Martin, A. J., Papworth, B., Ginns, P., Malmberg, L.-E., Collie, R. J., & Calvo, R. A. (2015). Real-time motivation and engagement during a month at school: Every moment of every day for every student matters. *Learning and Individual Differences*, 38, 26–35. <https://doi.org/10.1016/j.lindif.2015.01.014>
- Martin, F., & Borup, J. (2022). Online learner engagement: Conceptual definitions, research themes, and supportive practices. *Educational Psychologist*, 57(3), 162–177. <https://doi.org/10.1080/00461520.2022.2089147>
- Matcha, W., Gašević, D., Ahmad Uzir, N., Jovanović, J., Pardo, A., Maldonado-Mahauad, J., & Pérez-Sanagustín, M. (2019). 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)*: Vol. 11722 LNCS: 525–540. https://doi.org/10.1007/978-3-030-29736-7_39
- Matcha, W., Gašević, D., Jovanović, J., Uzir, N. A., Oliver, C. W., Murray, A., & Gasevic, D. (2020). Analytics of learning strategies. *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge*. 151–160. <https://doi.org/10.1145/3375462.3375534>
- Matcha, W., Gašević, D., Ahmad Uzir, N., Jovanović, J., Pardo, A., Lim, L., Maldonado-Mahauad, J., Gentili, S., Pérez-Sanagustín, M., & Tsai, Y.-S. (2020a). Analytics of learning strategies: role of course design and delivery modality. *Journal of Learning Analytics*, 7(2), 45–71. <https://doi.org/10.18608/jla.2020.72.3>
- Matcha, W., Uzir, N. A., Gasevic, D., & Pardo, A. (2020c). A systematic review of empirical studies on learning analytics dashboards: A self-regulated learning perspective. *IEEE Transactions on Learning Technologies*, 13(2), 226–245. <https://doi.org/10.1109/TLT.2019.2916802>
- Murphy, K., & Murphy, T. B. (2020). Gaussian parsimonious clustering models with covariates and a noise component. *Advances in Data Analysis and Classification*, 14(2), 293–325. <https://doi.org/10.1007/s11634-019-00373-8>
- Ouyang, F., Wu, M., Zheng, L., Zhang, L., & Jiao, P. (2023a). Integration of artificial intelligence performance prediction and learning analytics to improve student learning in online engineering course. *International Journal of Educational Technology in Higher Education*, 20(1), 4. <https://doi.org/10.1186/s41239-022-00372-4>
- Ouyang, F., Xu, W., & Cukurova, M. (2023b). An artificial intelligence-driven learning analytics method to examine the collaborative problem-solving process from the complex adaptive systems perspective. *International Journal of Computer-Supported Collaborative Learning*, 18(1), 39–66. <https://doi.org/10.1007/s11412-023-09387-z>
- Panadero, E. (2017). A review of self-regulated learning: Six models and four directions for research. *Frontiers in Psychology*. <https://doi.org/10.3389/fpsyg.2017.00422>
- Panadero, E., Broadbent, J., Boud, D., & Lodge, J. M. (2019). Using formative assessment to influence self- and co-regulated learning: the role of evaluative judgement. *European Journal of Psychology of Education*. <https://doi.org/10.1007/s10212-018-0407-8>
- 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. 452–502). Academic Press.
- Poquet, O., Jovanovic, J., & Pardo, A. (2023). Student profiles of change in a university course: A complex dynamical systems perspective. *LAK23: 13th International Learning Analytics and Knowledge Conference*. 197–207. <https://doi.org/10.1145/3576050.3576077>
- Raab, M., & Struffolino, E. (2023). *Sequence analysis*. SAGE Publications.
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models applications and data analysis methods* (Vol. 1). Sage.
- Redmond, P., Heffernan, A., Abawi, L., Brown, A., & Henderson, R. (2018). An online engagement framework for higher education. *Online Learning*, 22(1), 183–204. <https://doi.org/10.24059/olj.v22i1.1175>
- Rienties, B., Cross, S., & Zdrahal, Z. (2017). Implementing a learning analytics intervention and evaluation framework: What works? In B. K. Daniel (Ed.), *Big data and learning analytics in higher education* (pp. 147–166). Springer.
- Ritschard, G. (2023). Measuring the nature of individual sequences. *Sociological Methods and Research*, 52(4), 2016–2049. <https://doi.org/10.1177/00491241211036156>
- Ryan, R. M., & Deci, E. L. (2020). Intrinsic and extrinsic motivation from a self-determination theory perspective: Definitions, theory, practices, and future directions. *Contemporary Educational Psychology*. <https://doi.org/10.1016/j.cedpsych.2020.101860>
- Sáiz-Manzanares, M. C., Marticorena-Sánchez, R., Martín-Antón, L. J., González Díez, I., & Almeida, L. (2023). Perceived satisfaction of university students with the use of chatbots as a tool for self-regulated learning. *Heliyon*, 9(1), e12843. <https://doi.org/10.1016/j.heliyon.2023.e12843>

- Saqr, M., & López-Pernas, S. (2021a). Idiographic learning analytics: A single student (N=1) approach using psychological networks. *CEUR Workshop Proceedings*. 2868.
- Saqr, M., & López-Pernas, S. (2021b). The longitudinal trajectories of online engagement over a full program. *Computers and Education*, 175(September), 104325. <https://doi.org/10.1016/j.compedu.2021.104325>
- Saqr, M., & López-Pernas, S. (2022). How CSCL roles emerge, persist, transition, and evolve over time: A four-year longitudinal study. *Computers and Education*, 189, 104581. <https://doi.org/10.1016/j.compedu.2022.104581>
- Saqr, M., & López-Pernas, S. (2024). Mapping the self in self-regulation using complex dynamic systems approach. *British Journal of Educational Technology*, 55(4), 1376–1397. <https://doi.org/10.1111/bjet.13452>
- Saqr, M., López-Pernas, S., Helske, S., Durand, M., Murphy, K., Studer, M., & Ritschard, G. (2024). Sequence analysis in education: Principles, technique, and tutorial with r. In M. Saqr & S. López-Pernas (Eds.), *Learning analytics methods and tutorials: A practical guide using R*. Springer.
- Saqr, M., López-Pernas, S., Helske, S., & Hrastinski, S. (2023a). The longitudinal association between engagement and achievement varies by time, students' profiles, and achievement state: A full program study. *Computers and Education*, 199, 104787. <https://doi.org/10.1016/J.COMPEDU.2023.104787>
- Saqr, M., López-Pernas, S., Jovanović, J., & Gašević, D. (2023b). Intense, turbulent, or wallowing in the mire: A longitudinal study of cross-course online tactics, strategies, and trajectories. *The Internet and Higher Education*, 57, 100902. <https://doi.org/10.1016/j.iheduc.2022.100902>
- Saqr, M., Matcha, W., Ahmad Uzir, N., Jovanovic, J., Gašević, D., & López-Pernas, S. (2023c). Transferring effective learning strategies across learning contexts matters: A study in problem-based learning. *Australasian Journal of Educational Technology*, 2023(3), 35–57. <https://doi.org/10.14742/ajet.8303>
- Schmitz, B. (2006). Advantages of studying processes in educational research. *Learning and Instruction*, 16(5), 433–449. <https://doi.org/10.1016/j.learninstruc.2006.09.004>
- Schunk, D. H. (2005). Self-regulated learning: the educational legacy of Paul R. Pintrich. *Educational Psychologist*, 40(2), 85–94. https://doi.org/10.1207/s15326985Sep4002_3
- Schunk, D. H., & Greene, J. A. (2017). Handbook of self-regulation of learning and performance. In D. H. Schunk & J. A. Greene (Eds.), *Handbook of self-regulation of learning and performance*. Routledge.
- Sedrakyan, G., Malmberg, J., Verbert, K., Järvelä, S., & Kirschner, P. A. (2020). Linking learning behavior analytics and learning science concepts: Designing a learning analytics dashboard for feedback to support learning regulation. *Computers in Human Behavior*, 107(2014), 105512. <https://doi.org/10.1016/j.chb.2018.05.004>
- Shafiq, D. A., Marjani, M., Habeeb, R. A. A., & Asirvatham, D. (2022). Student retention using educational data mining and predictive analytics: A systematic literature review. *IEEE Access*, 10, 72480–72503. <https://doi.org/10.1109/ACCESS.2022.3188767>
- Siemens, G. (2013). Learning analytics. *American Behavioral Scientist*, 57(10), 1380–1400. <https://doi.org/10.1177/0002764213498851>
- Solhan, M. B., Trull, T. J., Jahng, S., & Wood, P. K. (2009). Clinical assessment of affective instability: Comparing EMA indices, questionnaire reports, and retrospective recall. *Psychological Assessment*, 21(3), 425–436. <https://doi.org/10.1037/a0016869>
- Stanovich, K. E. (2009). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Journal of Education*. <https://doi.org/10.1177/0022057409189001-204>
- Studer, M. (2013). WeightedCluster Library Manual: A practical guide to creating typologies of trajectories in the social sciences with R. (LIVES Working Papers, 24). <https://doi.org/10.12682/lives.2296-1658>
- Suraworachet, W., Zhou, Q., & Cukurova, M. (2023). Impact of combining human and analytics feedback on students' engagement with, and performance in, reflective writing tasks. *International Journal of Educational Technology in Higher Education*, 20(1), 1. <https://doi.org/10.1186/s41239-022-00368-0>
- Susnjak, T., Ramaswami, G. S., & Mathrani, A. (2022). Learning analytics dashboard: A tool for providing actionable insights to learners. *International Journal of Educational Technology in Higher Education*, 19(1), 12. <https://doi.org/10.1186/s41239-021-00313-7>
- Sweller, J. (2020). Cognitive load theory and educational technology. *Educational Technology Research and Development*. <https://doi.org/10.1007/s11423-019-09701-3>
- Tan, T. K., & Samavedham, L. (2022). The learning process matter: A sequence analysis perspective of examining procrastination using learning management system. *Computers and Education Open*, 3, 100112. <https://doi.org/10.1016/j.caeo.2022.100112>
- Törmänen, T., Järvenoja, H., Saqr, M., Malmberg, J., & Järvelä, S. (2022). A person-centered approach to study Students' socio-emotional interaction profiles and regulation of collaborative learning. *Frontiers in Education*. <https://doi.org/10.3389/feduc.2022.866612>
- Törmänen, T., Järvenoja, H., Saqr, M., Malmberg, J., & Järvelä, S. (2023). Affective states and regulation of learning during <scp>socio-emotional</scp> interactions in secondary school collaborative groups. *British Journal of Educational Psychology*, 93(51), 48–70. <https://doi.org/10.1111/bjep.12525>
- Villalobos, E. (2022). Supporting Self-regulated Learning in BL: Exploring Learners' Tactics and Strategies. *Educating for a New Future: Making Sense of Technology-Enhanced Learning Adoption: 17th European Conference on Technology Enhanced Learning, EC-TEL 2022, Toulouse, France, September 12–16, 2022, Proceedings*.
- Wang, M., & Degol, J. (2014). Staying engaged: Knowledge and research needs in student engagement. *Child Development Perspectives*, 8(3), 137–143. <https://doi.org/10.1111/cdep.12073>
- Wang, W., Lee, H., Lin, C., Li, P., Huang, Y., & Wu, T. (2024). Enhancing students' learning outcomes in self-regulated virtual reality learning environment with learning aid mechanisms. *British Journal of Educational Technology*. <https://doi.org/10.1111/bjet.13512>
- Wigfield, A., Eccles, J. S., Schiefele, U., Roeser, R. W., & Davis-Kean, P. (2007). *Development of achievement motivation*. Wiley. <https://doi.org/10.1002/9780470147658.chpsy0315>
- Winne, P. H. (2022). Modeling self-regulated learning as learners doing learning science: How trace data and learning analytics help develop skills for self-regulated learning. *Metacognition and Learning*, 17(3), 773–791. <https://doi.org/10.1007/s11409-022-09305-y>

- Winne, P. H., & Hadwin, A. F. (1998). Studying as self-regulated engagement in learning. In D. Hacker, J. Dunlosky, & A. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 277–304). Erlbaum.
- Winne, P. H., & Hadwin, A. F. (2008). The weave of motivation and self-regulated learning. In D. H. Schunk & B. J. Zimmerman (Eds.), *Motivation and self-regulated learning: Theory, research, and applications* (pp. 297–314). Lawrence Erlbaum Associates Publishers.
- Wise, A. F. (2014). Designing pedagogical interventions to support student use of learning analytics. *Proceedings of the Fourth International Conference on Learning Analytics and Knowledge*. <https://doi.org/10.1145/2567574>
- Wolters, C. A., Iaconelli, R., Peri, J., Hensley, L. C., & Kim, M. (2023). Improving self-regulated learning and academic engagement: Evaluating a college learning to learn course. *Learning and Individual Differences*, 103, 102282. <https://doi.org/10.1016/j.lindif.2023.102282>
- Wylie, C., & Hodgen, E. (2012). Trajectories and patterns of student engagement: Evidence from a longitudinal study. *Handbook of Research on Student Engagement* (pp. 585–599). Springer. https://doi.org/10.1007/978-1-4614-2018-7_28
- Xu, W., Wu, Y., & Ouyang, F. (2023). Multimodal learning analytics of collaborative patterns during pair programming in higher education. *International Journal of Educational Technology in Higher Education*, 20(1), 8. <https://doi.org/10.1186/s41239-022-00377-z>
- You, S., & Sharkey, J. (2009). Testing a developmental–ecological model of student engagement: A multilevel latent growth curve analysis. *Educational Psychology*, 29(6), 659–684. <https://doi.org/10.1080/01443410903206815>
- Zhang, X., Zhang, B., & Zhang, F. (2023). Student-centered case-based teaching and online–offline case discussion in postgraduate courses of computer science. *International Journal of Educational Technology in Higher Education*, 20(1), 6. <https://doi.org/10.1186/s41239-022-00374-2>
- Zimmerman, B. J. (2000). Attaining self-regulation. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13–39). Elsevier. <https://doi.org/10.1016/B978-012109890-2/50031-7>
- Zimmerman, B., & Schunk, D. (2011). *Handbook of self-regulation of learning and performance*. Routledge Taylor & Francis Group.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

SAMI HEIKKINEN

This dissertation explores how self-regulated learning, student engagement, and learning analytics interact in online education. Through mixed-methods analysis of business student cohorts, the thesis identifies distinct engagement patterns and self-regulation trajectories and explores the impact of learning analytics interventions on learning. The findings provide insights for designing adaptive online learning environments that support diverse student needs.



UNIVERSITY OF
EASTERN FINLAND

uef.fi

**PUBLICATIONS OF
THE UNIVERSITY OF EASTERN FINLAND**
Dissertations in Science, Forestry and Technology

ISBN 978-952-61-5624-8
ISSN 2954-131X