

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@uefi.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

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 meta-cognition 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 (Bischof 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.

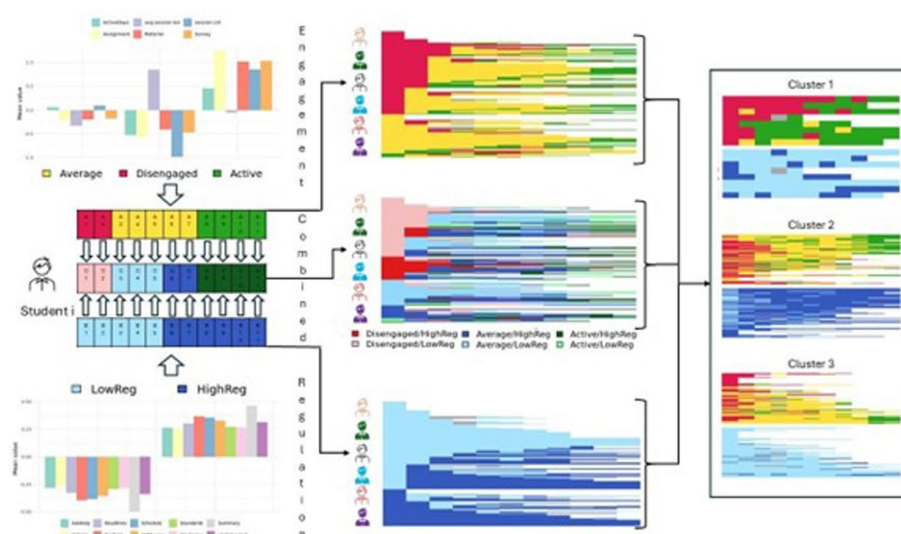


Fig. 1 The research process. On the left are the engagement and learning regulation states and an example of the sequence formation. In the middle are every student's index plots for engagement, regulation and combined sequences. On the right, similar sequences are clustered

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

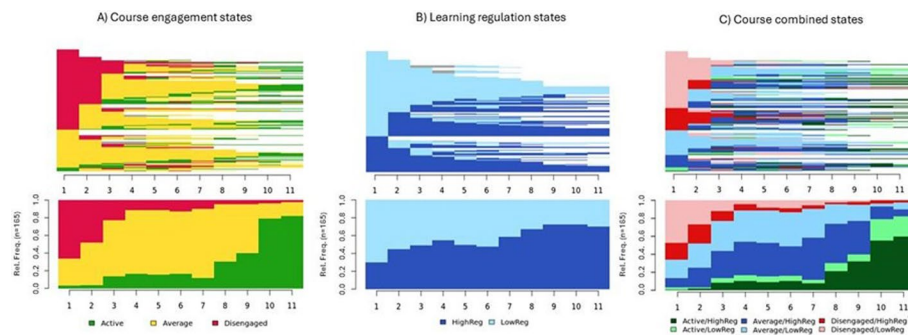


Fig. 2 Sequence index plot (top) and sequence distribution plot (bottom) for **A** course engagement states, **B** learning regulation states, and **C** course combined states

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

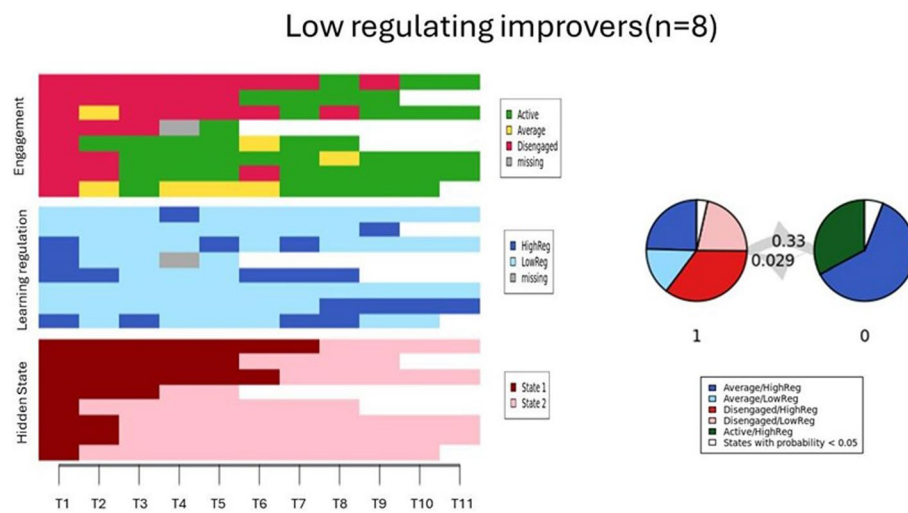


Fig. 3 Characteristics of Low regulating improvers. On the top left is an index plot of the engagement states; in the middle are learning regulation states; and at the bottom are hidden states. On the right-hand side is the transition diagram

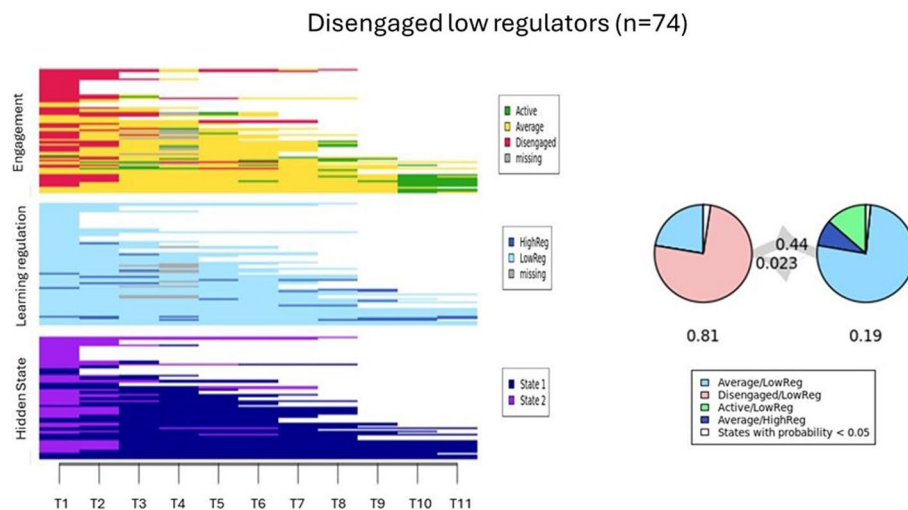


Fig. 4 Characteristics of Disengaged Low Regulators. On the top left is an index plot of the engagement states; in the middle are learning regulation states, and at the bottom are hidden states. On the right-hand side is the transition diagram

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

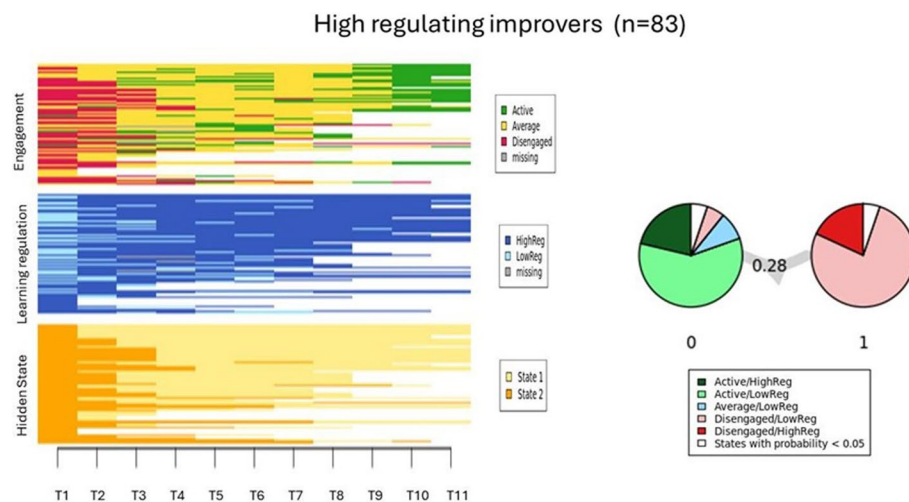


Fig. 5 Characteristics of High regulating improvers. On the top left is an index plot of the engagement states, in the middle learning regulation states and at the bottom hidden states. On the right-hand side is the transition diagram

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 & Saqr, 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>
- Bischl, B., Lang, M., Bossek, J., Horn, D., Richter, J., & Surmann, D. (2022). *Miscellaneous Helper Functions for B. Bischl*. <https://github.com/berndbischl/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 (*Pegomya 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. *ASCILITE*. <https://doi.org/10.14742/apubs.2013.1518>
- Cleary, T. J., Slem, 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>
- 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>
- 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 Staaldin, 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>
- 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>
- 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 CSCL. *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 OSLQ 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 CSDL 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/educ.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(S1), 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.