Computers & Education 199 (2023) 104787

Contents lists available at ScienceDirect

Computers & Education

journal homepage: www.elsevier.com/locate/compedu

The longitudinal association between engagement and achievement varies by time, students' profiles, and achievement state: A full program study

Mohammed Saqr [a,*], Sonsoles López-Pernas [a], Satu Helske [b], Stefan Hrastinski [c]

[a] *University of Eastern Finland, School of Computing, Yliopistokatu 2, FI-80100, Joensuu, Finland*
[b] *University of Turku, INVEST Research Flagship Center & Department of Social Research, Turun Yliopisto, FI-20014, Turku, Finland*
[c] *KTH Royal Institute of Technology, Division of Digital Learning, Lindstedtsvägen 30, SE-100 44, Stockholm, Sweden*

ARTICLE INFO

Keywords:
Learning analytics
Multi-channel sequence analysis
Longitudinal engagement
Academic achievement
Mixture hidden Markov models
Student profiles
Person-centered methods

ABSTRACT

There is a paucity of longitudinal studies in online learning across courses or throughout programs. Our study intends to add to this emerging body of research by analyzing the longitudinal trajectories of interaction between student engagement and achievement over a full four-year program. We use learning analytics and life-course methods to study how achievement and engagement are intertwined and how such relationship evolves over a full program for 106 students. Our findings have indicated that the association between engagement and achievement varies between students and progresses differently between such groups over time. Our results showed that online engagement at any single time-point is not a consistent indicator for high achievement. It takes more than a single point of time to reliably forecast high achievement throughout the program. Longitudinal high grades, or longitudinal high levels of engagement (either separately or combined) were indicators of a stable academic trajectory in which students remained engaged —at least on average— and had a higher level of achievement. On the other hand, disengagement at any time point was consistently associated with lower achievement among low-engaged students. Improving to a higher level of engagement was associated with —at least— acceptable achievement levels and rare dropouts. Lack of improvement or "catching up" may be a more ominous sign that should be proactively addressed.

1. Introduction

The association between learners' engagement and academic achievement is largely recognized in the literature (King, 2015; Lei, Cui, & Zhou, 2018). Yet, existing research has mostly explored the association with little regard to temporality and evolution, e.g., at a single time-point in a task (Frishkoff, Collins-Thompson, Hodges, & Scott, 2016; Lyubovnikova, Napiersky, & Vlachopoulos, 2015) or a single course (Gašević, Jovanović, Pardo, & Dawson, 2017; Kovanović, Gašević, Joksimović, Hatala, & Adesope, 2015). Thus, a gap exists in our knowledge regarding the longitudinal association between engagement and achievement across time, i.e., how such association unfolds, changes, or remains stable. Given the increasing recognition of the heterogeneity of students' behavior, i.e., the presence of subpopulations with different behavioral patterns, we use methods that help find out for *whom* the evolution happens, that

* Corresponding author. University of Eastern Finland School of Computing, Joensuu Campus, Yliopistokatu 2, FI-80100, Joensuu, P.O. Box 111, FI-80101 Joensuu, Finland, Finland.
E-mail address: mohammed.saqr@uef.fi (M. Saqr).

https://doi.org/10.1016/j.compedu.2023.104787
Received 18 July 2022; Received in revised form 24 February 2023; Accepted 18 March 2023
Available online 29 March 2023
0360-1315/

is, for which students' subpopulation a change in engagement (improvement or decline) is associated with a change in achievement (increase or decrease). More importantly, we address *when* such changes happen, i.e., at the beginning, during, in certain courses, or towards the end or if it lags for some time. Our study aims to bridge the aforementioned gaps regarding the longitudinal association and evolution of engagement and achievement using a dataset of a full healthcare program (four years). We take advantage of the latest advances in learning analytics, statistics, and multichannel sequence analysis to achieve the study aims.

The contributions of our paper are manifold. First, we show that engagement and achievement evolve differently across students' subgroups. For instance, high achievers are relatively stable despite occasional decline in engagement whereas decline or failure to catch up is rather ominous in average and low achievers. Second, we map the trends of evolutions, i.e., the trajectories of engagement and achievement per student subgroup to find out when and for whom changes happen and their temporal trends. Third, we offer a detailed methodology for the study of an association across time using multichannel sequence mining which is a valuable methodological contribution.

The paper is organized as follows: in the background section, we review the concept of student engagement, its evolution, and its longitudinal unfolding. We follow by a review of academic achievement and how it changes over time. Subsequently, we review the association between engagement and achievement and how this relationship evolves. We follow by explaining why studying student subgroups is important, followed by stating the motivation of our study and the respective research questions. Since the methods of the study are novel to the educational community, we describe them in detail and explain how they help us study the phenomena under investigation. In the Methods section, we contextualize the study, describe the data operationalized, and provide a step-by-step explanation of the data analysis performed. The following section describes the results of the analysis, followed by a thorough discussion.

2. Background

2.1. Student engagement

Engagement is a multidimensional construct that integrates three dimensions: behavior, emotion, and cognition (Martin & Borup, 2022; Redmond, Abawi, et al., 2018). These dimensions are linked and intertwined (Borup, Graham, Leanna Archambault, & Spring, 2020; Martin & Borup, 2022). Behavioral engagement includes following school rules, efforts in studying, and involvement in academic activities. Examples of such engagement are attendance, participation, and commitment to schoolwork. Emotional engagement comprises how students feel about the school, other learners, and their emotional experience in learning which could be positive emotions (e.g., pride, interest, or happiness) or negative emotions (e.g., disinterest, sadness, or anxiety). Lastly, cognitive engagement represents students' thoughtful investment and going beyond the required in learning and can encompass students' efforts to tackle learning tasks, understand critical concepts, and self-regulate (Borup et al., 2020). All three dimensions of engagement have been found to affect achievement at all educational levels (King, 2015; Lei et al., 2018). These dimensions and their components are interrelated. For example, feeling positively about learning catalyzes being behaviorally engaged in course work and boosts cognitive engagement (Azevedo, 2015; Martin & Borup, 2022). Whereas initial conceptualizations and theoretical frameworks have been developed for face-to-face engagement, the concepts and types of engagement have been described in both online and blended learning environments (Martin & Borup, 2022; Redmond, Abawi, et al., 2018).

Although studying online engagement has received considerable attention over the years, it has been studied in different ways, meaning that we lack a sound and generally established theoretical understanding of how online engagement can be measured (Martin & Borup, 2022; Redmond, Abawi, et al., 2018). However, most existing research attempts to identify student engagement by analyzing the observable indicators of online activities (Henrie, Halverson, & Graham, 2015; Redmond, Abawi, et al., 2018). Examples of such indicators are the number of assignments completed, login frequency, number and frequency of postings, time spent creating posts, time spent online, number of learning resources accessed (Fincham et al., 2019; Henrie et al., 2015). Such operationalizations of engagement are commonly used in the online learning and learning analytics literature (Azevedo, 2015; Henrie et al., 2015). Although trace data from online activities have offered an unobtrusive method to capture proxy indicators of levels of engagement, they remain limited to behavioral and cognitive engagement.

2.2. Evolution of engagement

A considerable volume of research in face-to-face education has been devoted to examining the evolution of student engagement across school years (e.g., Archambault & Dupéré, 2017; Froiland & Oros, 2014; Li & Lerner, 2011; Zhen, De Liu et al., 2020). Such research has established that engagement changes over time as a result of many factors, such as motivation, teachers, peers, learning tasks and social context (e.g., Wang & Degol, 2014; You & Jill Sharkey, 2009). However, research is so far inconclusive: some scholars have reported that current engagement in a course predicts future engagement in the next courses (Gottfried, Marcoulides, Gottfried, & Oliver, 2009; Wylie & Hodgen, 2012); whereas other scholars have reported that engagement declines over time (Wigfield, Eccles, Ulrich, Roeser, & Davis-Kean, 2007).

Unlike studies in face-to-face education, most of the existing research in online learning has focused on individual courses, hence a gap of research on longitudinal online engagement across courses or throughout a program exists. What is more, most studies addressing longitudinal engagement rely on students' or teachers' perceptions using self-reported questionnaires that are subject to recall bias and attrition challenges (Azevedo, 2015). An exception is the work of Barthakur et al. (2021) and Saqr and López-Pernas (2021), who clustered students' online activities and reported the presence of program-level longitudinal patterns of engagement that varied across subgroups. Both studies reported a predominantly engaged cluster where students remained engaged most of their program, a disengaged cluster, and an intermediate cluster where students showed low levels of engagement across the whole study time. As education increasingly relies on online learning environments, especially in higher education in which independent online study is key, it is imperative to study student engagement with the online learning materials available to them (Martin & Borup, 2022). Furthermore, such online data has proven useful in capturing levels of engagement (e.g., Henrie et al., 2015; Joksimović et al., 2018).

2.3. Academic achievement

Academic achievement is an overarching construct that could be defined differently based on contextual variables such as the program, study field, and educational level. A general conceptualization can be viewed as the performance outcome that indicates the level of accomplishment of the objectives of a learning activity, task, or course. Most educational institutions base academic achievement on students' graded performance over a series of academic years among other methods (Kinai, Ndambuki, & Peter, 2019). The search for variables that correlate with or support academic achievement has garnered a vast corpus of research which spans several domains (Schneider & Preckel, 2017). Two main categories of variables have been generally explored: 1) demographic and contextual variables (e.g., gender, socioeconomics, and background), and 2) psychological, behavioral or dispositional factors (e.g., motivation, learning approach and personality) (Richardson, Abraham, & Bond, 2012; Schneider & Preckel, 2017).

Research exploring the relationship between demographic and contextual variables has traditionally reported that learners from underprivileged families are more prone to low academic achievement and possibly fall behind in their education (Jimerson, Egeland, & Teo, 1999). However, recent and comprehensive meta-analyses concluded that such demographic factors have —at best— a weak correlation with achievement (Richardson et al., 2012), which was the smallest effect sizes of all the categories of variables examined by Schneider and Preckel (2017) in a synthesis of 38 meta-analyses and 3330 effect sizes. In fact, the authors argued that some of these effect sizes are possibly a result of a spurious correlation. Taken together, the evidence from the meta-analyses emphasizes the notion that changes in academic achievement are concurrent with changes related to student behavior, e.g., engagement, self-efficacy, and goal-directedness.

2.4. Engagement and achievement: association and evolution

A positive association between engagement and achievement —as previously described— has been repeatedly reported in the literature across several contexts, types of engagements and subpopulations (Lei et al., 2018). A well-cited and widely used theoretical model that defines and explains such a relationship is the model of Christenson, Wylie, and Reschly (2012), which conceptualized such an association as the product of interaction between three main factors: contextual factors, learning process (i.e., engagement), and learning outcome (or achievement). While we use the model by Chirstenson and colleagues as a framing for our study, the model's "contextual factors" are time invariant and immutable, i.e., they cannot be modified and do not change with time and, therefore, will not be accounted for in our longitudinal model.

Few studies have researched the longitudinal relationship between engagement and academic achievement. One example is the work by Hughes, Luo, Kwok, and Loyd (2008), who measured effortful engagement (evaluated through a questionnaire), student-teacher relationship quality, and achievement in mathematics and reading on a yearly basis for three years in an elementary school. Their findings revealed that achievement, efforts in engagement, and student-teacher relationship form a dynamic system that affects early school years, whereas early intervention could positively affect children's school achievement trajectory. Using a large sample of students in the United States, Froiland and Oros (2014) followed students longitudinally across school years (5th to 8th) and found that students who have intrinsic motivation and are engaged (reported by teachers' rating) early in the 5th grade are more likely to have better reading achievement in the 8th grade. Moreover, the longitudinal study by Wang and Eccles (2013) examined the association between how learners perceive achievement motivation, school environment, and school engagement (cognitive engagement, behavioral and emotional). They found that students' perceptions of different aspects of the school environment had a significant contribution in all three types of engagement as well as in achievement motivation.

2.5. Modeling longitudinal engagement

This study takes advantage of the latest advances in modeling the typology of the longitudinal life course events and learning analytics. In particular, we use multi-channel sequence analysis, which allows modeling the unfolding of both engagement (first channel) and achievement (second channel) over time. In doing so, multi-channel sequence analysis offers an innovative method that fits the purpose of our study. In the present section, we offer a brief overview of sequence mining as a method and multi-channel sequence analysis.

A sequence can be defined as a chronologically ordered group of states (e.g., engagement or achievement level in our case). Sequence mining has been conceptualized to mine such time-ordered data and take advantage of their sequential state (Agrawal & Srikant, 1995). The method offers a rich array to visualize, cluster and statistically analyze sequential data. Therefore, several studies in education have adopted such methods to study the successions of students' activities, as well as self-regulation tactics or strategies (Matcha et al., 2020). What is more, sequence mining offers several tools and methods that allow the study of typical recurrent states commonly referred to as trajectories (Agrawal & Srikant, 1995; Gauthier, Widmer, Bucher, & Notredame, 2010; Helske, Helske, & Eerola, 2018). A trajectory is a time-ordered sequence of states that are similar to each other and distinct from other trajectories (or roughly a cluster of sequential data).

The abundance of multidimensional sequences as several variables evolving in parallel (e.g., engagement and achievement in our case) has led to the emergence of multi-channel sequence analysis, which extends the traditional sequence mining methods to account for several streams of sequences (Gauthier et al., 2010; Helske et al., 2018). Multi-channel sequence analysis offers a method for studying such combined sequences. For instance, multimodal learning analytics could be used to study the sequence of heart rate changes, along with the sequence of skin changes. One could align both variables as channels and use multi-channel sequence analysis to gain insights about their sequential relationship.

Clustering or mining the trajectories in sequential data is a common method in learning science research and learning analytics (Jovanovic, Dawson, Joksimovic, & George, 2020; Matcha et al., 2020). However, clustering multi-channel sequences requires special methods (Helske et al., 2018; Zhang, Lee, & Lee, 2018). Therefore, researchers have applied Hidden Markov Models (HMM) to cluster multi-channel sequence data based on their established role in clustering sequences and other data streams (Helske et al., 2018; Helske & Helske, 2019).

2.6. Heterogeneity of students' behavior

Asikainen and Gijbels's (2017) ascribed the lack of conclusive evidence of the evolution of student strategies to the fact that some studies have investigated a whole group, i.e., aggregated results over the entire sample without considering the possible existence of subgroups. That is, the presence of heterogeneous subgroups with contradictory trends (a group increasing and another decreasing) would give rise to a false conclusion of a stable trend. The authors noted that the studies that used a person-centered approach (that explores students' subgroups) —similar to our approach here— had more consistent findings where "different subgroups of students develop differently and the development is individual in nature" (Fryer, 2017; Postareff, Lindblom-Ylänne, & Anna, 2014, 2015). Such a variation by subgroup is often referred to as heterogeneity. Heterogeneity has received vast empirical evidence as well as theoretical backing across social and behavioral sciences. (Bryan, Tipton, & Yeager, 2021; Hickendorff, Edelsbrunner, McMullen, Schneider, & Kelly, 2018; Rosato & Judith, 2012).

Oftentimes, students can hardly conform to a common "average" resulting in variations between findings. Therefore, researchers have addressed such heterogeneity in psychological phenomena with appropriate methods that look for hidden or latent groups that have distinct patterns of evolution, e.g., person-centered methods (Bryan et al., 2021).

Such person-centered methods hold the hope for helping —at least partially— address some of the challenges of research replicability. As pointed out by Bryan et al. (2021), the replicability crisis can be overcome with a heterogeneity revolution. Recently, it has been demonstrated that the patterns of engagement evolution may vary in different students' subgroups, i.e., engagement could be stable in a subgroup of students and declining in another subgroup (Archambault & Dupéré, 2017; Saqr & López-Pernas, 2021; You & Jill Sharkey, 2009; Zhen, Liu, et al., 2020). When and for whom such temporal changes occur is an area that has recently attracted the attention of researchers. Thereupon, a main objective of this paper is to analyze the heterogeneity within temporal trends and answer the question "for whom". In other words, we examine the variability of evolutionary patterns among students' subgroups.

2.7. Motivation of the study

Whereas efforts have been devoted to studying the relationship between academic achievement and students' online behavior using different variables, e.g., regularity of online behavior as evidence of engagement (Saqr, Fors, & Tedre, 2017), number of learning activities or students' contributions in online collaborative work (Fincham et al., 2019). A dominant pattern of such work is that it oftentimes focuses on a single course or learning activity (Martin, Sun, & Westine, 2020). Therefore, little is known about how student achievement and engagement evolve together throughout time and if such evolution is consistent among all students or have distinct patterns in different subpopulations. If engagement decreases, do we expect achievement to decrease? Does it change for every student, or are there sub-populations of students in which the relationships evolve differently? And, if so, which groups of students are prone to drop in engagement? Within these groups, which subgroups of students can return to their levels of engagement and achievement? When can we consider engagement a good indication of achievement, i.e., at the beginning of the program or later? Does the association between engagement and achievement remain stable? Or does it vary across time or across groups?

Our study aims to answer such questions using a dataset of four years of students' online activities using an innovative longitudinal method that combines multiple channels of data (described in the next section). Our methods model the longitudinal evolution while looking at the heterogeneity and variability among subgroups. To that end, our research questions are as follows.

- How are engagement and achievement associated across a full program?
- How do student engagement and achievement evolve across a full program across different students' subgroups?

3. Methods

3.1. Context and program description

This study is based on a Problem-based learning (PBL) healthcare program which integrates clinical and basic sciences in most courses as early as the first course. All courses in the study are designed based on an integrated approach, which means they cover several subjects under a shared theme. Examples include the biochemistry, biology, anatomy, and histology of the kidney will be integrated together into the Renal System Source. In the same token, the Nervous System course contains the biology, physiology, and anatomy of the brain and other systems. The courses also have a similar assessment strategy which includes continuous assessment, practical exams, and written exams (see section 3.5) for details. The program is a PBL program, and therefore, the courses share the PBL approach. Courses differ in subject matter, and there is a slight difference in course duration as the course duration ranges from 6 to 8 weeks. Teachers may differ, but there is a large overlap between courses where the same subject teacher could be teaching the subject in multiple courses, e.g., anatomy may be given by the same teacher across several courses. The courses are based on PBL, in which students are supposed to engage in small group discussions around a weekly *problem*. The *problem* is an ill-structured clinical scenario that follows the week's intended learning objectives to trigger discussions (Wood, 2003). The intended learning objectives of all the *problems* should cover most of the intended learning objectives of the courses. The PBL implementation in the program extends weekly face-to-face discussions with an online forum as a platform for PBL interactions (blended PBL). A typical PBL starts with a group of students (5–10) meeting with a teacher on the first day of the week (face-to-face) where they read the problem, clarify the objectives, and co-construct learning issues. The students then continue the discussions online with the same group structure and the same teacher throughout the week. On the last day of the week, students convene to discuss conclusions, feedback, reflect on their performance, and group work (Wood, 2003). The courses include lectures, seminars, laboratory sessions, as well as clinical skills. These activities are aligned with the problem discussions.

Courses are supported by an online learning management system (LMS). The online LMS provided a platform for 1) distributing course materials (lectures, videos, PowerPoint presentations of lectures, books or links to external learning resources), 2) a platform to support PBL discussions as an extension for the face-to-face sessions, 3) a gateway for course updates, news, questions or answers regarding the course, 4) an event calendar that shows what is scheduled and what is to come, 5) a gradebook that students use to know their grade and performance, and 6) a survey tool for course evaluation. Therefore, engaged students are expected to follow online course materials, interact in the online PBL discussions, check course updates through the main page, and communicate with teachers or peers. Please see section 3.3 for a description of the collected data and how they were operationalized. The courses in the program are sequential and therefore are referred to as blocks. There are other courses that are year-long (e.g., clinical skills). These courses were excluded since they are mostly practical, do not have an online component, and are assessed mainly based on performance of practical skills.

3.2. Data collection

Logs were extracted from the University's Moodle LMS for all the subjects and for all the students who attended the years: 2015, 2016, 2017 and 2018. Only Moodle logs representing events related to learning were considered, i.e., teacher events and non-learning related activities e.g., viewing personal profile pages, resetting the password, or enrolling in the course. Rare clicks (e.g., clicks on inactive modules) were also excluded as they were very infrequent.

3.3. Engagement indicators and operationalization

Our choice and operationalization of engagement indicators followed the literature on engagement measurement (e.g., Sinatra, Heddy, & Lombardi, 2015; Azevedo, 2015; Henrie et al., 2015), the theoretical frameworks, (e.g., Redmond, Heffernan, Abawi, Brown, & Henderson, 2018; Martin & Borup, 2022), as well as studies that studies engagement in individual courses (Jovanović, Gašević, Dawson, Pardo, & Mirriahi, 2017, 2020; Kovanović et al., 2016, 2019; Matcha et al., 2019), or longitudinal studies (Barthakur et al., 2021). Three types of indicators for each course were collected: 1) indicators that represent the intensity of engagement with the learning materials and modules (e.g., lectures, forums, announcements), 2) indicators that reflect the regularity and consistency in engagement with the course materials, and 3) indicators of time-on-task that reflect the time spent working with online materials (Joksimović et al., 2018; Lei et al., 2018). While several indicators were computed, we selected the indicators that represent the context (PBL) and blended type of learning, reflect intensity of engagement according to literature as well as avoid redundancy, e.g., regularity indicators were computed for all activities, nonetheless, they were highly correlated and offered little added value. Two dimensions of engagement can be inferred from the online data: 1) behavioral engagement which is captured by the frequency of activities, time spent on task and access to learning materials, and 2) cognitive engagement, which is captured by indicators that reflect intense investment in learning (considerable time-on-task), regularity and involvement in activities that necessitate that students engage in challenging ill-structured problems that require analysis, synthesis of evidence, argumentation, connecting complex concepts, elaboration and explanation as indicators of cognitive engagement (Henrie et al., 2015; Rotgans & Schmidt, 2011). We do not explicitly operationalize emotional enragement since online trace data can only enable the attainment of indicators of behavioral and cognitive engagement (Azevedo, 2015; Henrie et al., 2015). Nonetheless, as Halverson and Graham (2019) stated, behavioral engagement is an outward manifestation of cognitive and emotional engagement "Researchers may infer internal processes from external behaviors, and while those behaviors are not trivial, they still can be recognized as the outward displays of the mental and emotional energies that fuel learning" (Halverson & Graham, 2019, p. 153).

3.3.1. Frequency of activities

Frequency was calculated as the total count or total cumulative number of a certain action (e.g., reading forums) over the entire duration of the course. The frequency was computed using the R programming language with janitor and tidyverse packages (Firke, 2020; R Core Team, 2021; Wickham et al., 2019).

- **Course browsing frequency (FCB):** The frequency (total cumulative count over the whole course duration) of the times a student browsed the course front page, which contains the course information, course updates, and announcements. The front page of the course serves as the main interface for other course materials.
- **Forum consuming frequency (FFCs):** The frequency (total cumulative count over the whole course duration) of the times a student read, looked up or interacted with posts contributed by peers. These posts constitute a valuable resource for learning as they contain the course objectives, shared learning resources, arguments about the problem in discussion, and possible solutions.
- **Forum contributing frequency (FFCb):** The frequency (total cumulative count over the whole course duration) of the times a student posted or updated (i.e., contributed) to the PBL online discussion forums, i.e., started a thread, or uploaded or replied to a post. Contributions to the PBL forum required students to compose arguments, synthesize evidence, connect complex concepts, and curate resources that he/she studies, or write reflections on the group work and therefore, reflect elements of cognitive engagement.
- **Lecture viewing frequency (FLV):** The frequency (total cumulative count over the whole course duration) of the times a student accessed a piece of learning material (e.g., a presentation, a video, a report, etc.). This metric reflects students' interest in the course and its learning resources.

3.3.2. Online time

- **Session count (SC):** The frequency (total cumulative count over the whole course duration) of the times a student had an online learning session. Sessions are non-interrupted sequences of learning activities where the difference between two successive learning activities is below a given limit (Gašević et al., 2017). A time difference of 20 min of no learning activity was selected as a threshold corresponding to the 85th percentile. Since there is no guidance on a threshold, it usually follows the contextual conditions. We opted for a relatively longer time as students in our program had to read several contributions by other colleagues or compose forum posts, either of which can take time.
- **Total duration (TD):** The total time (in seconds) spent accessing learning resources, contributing to forums or reading others' contributions. The TD was computed as the difference between the start and the end of a learning session (excluding sessions and events not related to learning).
- **Active days (AD):** The count of days that the students had at least a single log record of online learning activity in the LMS (e.g., reading a forum, downloading a learning resource, or posting to a forum). High levels of time-on-task and investment in learning are evidence of cognitive engagement according to (Henrie et al., 2015).

3.3.3. Regularity

- **Regularity (R):** How regular a student is in performing learning activities, computed based on Shannon's entropy using the method by Jovanović, Mirriahi, Gašević, Dawson, and Pardo (2019). The measure reflects students' consistency and investment in learning. Regular students are self-regulating students who attend to their online learning on a regular basis, and thus, high levels of regularity are indicative of cognitive engagement.

Modeling a longitudinal process entails dealing with large variability within the LMS indicators over four years. Therefore, to account for the differences between courses, we used a technique known as binning, which is commonly used for modeling students' data over longer periods of time (Dziuban, Moskal, Cavanagh, & Watts, 2012; Sander Pete and Information Services, 2016). All indicators were converted to deciles of equal width (Alves, Morais, & Miranda, 2017; Dewar, Hope, Jaap, & Cameron, 2021). There is growing evidence that such a process of discretizing LMS indicators offers comparable predictive power (Jovanović, Saqr, Joksimović, & Gašević, 2021) to "raw" indicators, and in some cases even better predictive models (Jishan, Rashu, Haque, & Rahman, 2015). As such, changing the level of granularity or converting to discrete levels (mostly deciles) has been commonly performed as a data-preprocessing in longitudinal studies when deemed essential for the analysis (Dewar et al., 2021; Jishan et al., 2015).

3.4. Clustering

We used Latent Class Analysis (LCA) to group students based on their engagement state in each course (Porcu & Giambona, 2017) based on their LMS activity indicators (frequencies of activities, learning time invested in learning, consistency). LCA is a statistical technique that allows the clustering of students into categories of unobserved latent variables (engagement in our case) based on observed features (LMS indicators). Several studies have established the role of LCA as a clustering method in educational contexts (Hickendorff et al., 2018). To determine the best number of clusters we used the lowest Bayes information criterion (BIC) and the Akaike information criterion (AIC) (Hickendorff et al., 2018; Weller, Bowen, & Faubert, 2020). We tested a number of clusters between 1 and 10.

To examine the separation of clusters, we performed a Kruskal–Wallis non-parametric one-way analysis of variance (ANOVA) (Ostertagova, Ostertag, & Kováč, 2014) to compare the resulting engagement states in terms of their levels of activity for each of the LMS indicators. The ANOVA results' effect size was measured using epsilon-squared (Tomczak & Tomczak, 2014). Pairwise Post-hoc comparisons were computed using the Dunn's test with p-value Holm's adjustment for multiple testing (Holm, 1979). The resulting clusters were also plotted for each course using boxplots.

3.5. *Measuring student achievement*

Achievement in this study was measured by respective course final grades. The final grades in each course are the total grades of assessments performed before final exams (20% of the grade), and the final examination grade (80%). The exams are composed of multiple-choice questions (MCQs) and Modified short Essay Questions (MEQs). The exams assess the knowledge acquired during the course. The questions are revised and checked for quality before they are included in the final exam. After the students take the test, psychometric analysis is performed and problematic questions (e.g., extremely difficult, or extremely easy or confusing) from the grades by the grading committee and grades are adjusted to maximize fairness and objectivity.

We have done extensive research on how to classify students according to their achievement for the current study. The decision was informed by the literature on standard setting which has extensively addressed the issue systematically and scientifically (De Champlain & André, 2018; Norcini, 2003). Our aim was to strike a balance between interpretability, and validity, as well as to factor in the "consequences of adopting a given cut-score". In healthcare education —the setting of our study— students have to compete for a limited number of residency placements and therefore, we had to resort to equally spaced subgrouping with finite placement in each performance level. Therefore, we have used a three-tiered stratification of performance (high, intermediate, and low) (e.g., Davey, Kaplan, & Claire, 2014; Patel et al., 2020). Three levels of achievement were computed by dividing students into three equal levels of achievement in each course, where each tercile has almost the same number of students. As such, the final classification was Achiever (top 1/3rd of the student in the course), Intermediate (middle 1/3rd), Low (bottom 1/3rd). A student may have a different achievement state in each course, e.g., an Achiever in one course and an Intermediate in another one according to the rank among all students in the course.

3.6. *Contingency test*

To test the association of engagement states and achievement states in all courses, we performed a multi-way contingency mosaic plot. The reason is that we have three engagement states, three achievement states, and fifteen courses. A typical contingency table would have 135 cells (3 × 3*15), which would be very difficult to read. If we added the statistical significance, the number would double to 270 cells. The mosaic plot makes visualizing such high-dimensional data more intuitive. The mosaic plot is a visualization of the proportion of "observed versus expected" frequencies made up of cells similar to contingency tests (e.g., Chi-squared test). Cells can be signed (negative or positive) according to the magnitude and direction of the relationship represented as colors: blue indicates a positive relationship, while red indicates a negative relationship. The height of the cell reflects the proportion of the residual; the width is sized according to the square root of the expected frequency, and the box area indicates the magnitude of the difference between the expected and observed frequencies. A blue cell where the achiever and engaged state meet at course 1 means that there is a statistically significant association: the magnitude of the significance is the depth of color, and the size of the cell is relative to the number of students.

3.7. *Sequence mining*

Two sequences were built for each of the engagement and achievement states using the TramineR R package (Gabadinho, Gilbert, Studer, & Nicolas, 2009). To build the sequence for students, we ordered and synchronized the students' engagement and achievement states in each course. Since courses are chronologically ordered, we used the course starting date to order the sequence of the combined states for the fifteen courses available in the dataset. An example of the sequences can be as follows.

	Course 1		**Course 2**		**Course 3**		**Course 4**				**Course 15**
Engagement sequence	Active	→	Active	→	Disengaged	→	Average	→	...	→	Average
Achievement sequence	Achiever	→	Intermediate	→	Low	→	Intermediate	→	...	→	Achiever
Multi-channel sequence	Active/Achiever	→	Active/Intermediate	→	Disengaged/Low	→	Average/Intermediate	→	...	→	Active/Average

The distribution of sequences was plotted using the sequence distribution plot to show the share of each state at each course. A sequence index plot was plotted to visualize the sequence of states for students at each time point (e.g., course). The index plot represents each student's trajectory as a sequence of stacked colored blocks in a time ordered manner (1–15) to demonstrate the succession of states. The aforementioned plots were created for each channel and for both channels combined.

3.8. *Clustering of trajectories using the Mixture Hidden Markov Model (MHMM)*

We used the Mixture Hidden Markov Model (MHMM) to cluster students' trajectories according to their longitudinal patterns of both engagement and achievement states and to simultaneously study the dynamics within the trajectories. The analysis was performed in R using the package *seqHMM* (Helske & Helske, 2019) for the estimation and visualization of the MHMM. The MHMM can be regarded as a combination of LCA and HMM (Helske & Helske, 2019; Vermunt, Bac, & Jay, 2008). In HMM, observations in one or more channels (here engagement and achievement) are related to a hidden (latent) process which follows a *Markov chain*. In the (first-order) HMM, the status at time t depends on the status at time t-1, and not on any previous states. Furthermore, conditional on the hidden state, the observed state in channel c at time t is independent of the observed states in other channels at time t as well as the observed states in all channels before time t. The hidden states generate or *emit* observed states (engagement and achievement) with varying *emission probabilities*. The interpretation of the hidden state is similar to that of the latent class, but the hidden state can vary in time; we can estimate the *transition probabilities* between the hidden states. The MHMM extends the HMM by adding another layer of hidden time-constant statuses and can be used for clustering multichannel trajectories jointly while estimating the transition probabilities between the hidden states. The model was configured so that students who have a similar state at the beginning of their trajectories are clustered together. By doing so, we are able to study how many students who start the program with a similar engagement-achievement profile keep this profile throughout the program and how many deviate from their original profile. We estimated a number of MHMMs with varying number of hidden states and clusters, the estimation was repeated 1000 using random starting values in search for the global optimum according to (Helske et al., 2018). Finally, we determined the optimal number of clusters and hidden states using the BIC which was 4928.4.

4. Results

The dataset for this study contained the complete online traced-log data of 106 students along 4 years of education, including 1394 course enrollments in 15 distinct courses. After cleaning the data retrieved from the LMS, e.g., by removing logs not related to learning (such as clicks on profiles and chats), the total number of log records was 345,826. The median number of learning-related records per course offering was 19,591 (ranging from 4296 to 44,490). Table 1 shows a summary of the data for all the courses.

Clustering students to identify the distinct levels of engagement in each course resulted in three groups of students with three *engagement states*. Both AIC and BIC fit indices pointed to three clusters as the optimal solution (BIC = 29,435.54; AIC = 28,720.35, Fig. 1). Three clusters also accounted for the most interpretable model (Porcu & Giambona, 2017). See Fig. 2 for means, standard deviation (SD) and distribution of values of each LMS activity indicator for each of the clusters.

An ***Active*** cluster (n = 421, 30.2%), where students were intensely engaged with the course learning resources (FLV), active more days than their peers (AD), more regular (REG), had longer time online (SC and TD) and were more engaged in the forum discussions (FFCs and FFCb). The students in this cluster had levels of activity around the 8th decile (see Fig. 3) in all LMS indicators except for lecture (FLV) and forum contribute (FFCb) indicators where the activity was close to the 7th decile. We refer to these students as having an ***active engagement state.***

An ***Average*** cluster (n = 660, 47.3%), where students were intermediately engaged (between the active and disengaged). So, they had an intermediate level of activity of access to the course materials (FLV), intermediate number of active days (AD), intermediate levels of regularity (REG), time online (SC and TD) and intermediate levels of forum reading or contributing (FFCs and FFCs). The students in the average cluster had activity around the 5th decile in all indicators (see Fig. 3). We refer to these students as having an ***average engagement state.***

A ***Disengaged*** cluster (n = 313, 22.5%), where students showed low levels of access to course materials (FLV), had the least active days compared to their peers (AD), were the least regular (REG), had shortest time online (SC and TD) and were the least engaged in the forum discussions (FFCs and FFCs). The students in this cluster showed activity levels close to the 2nd decile in all indicators except for lecture (FLV) and forum contribute (FFCb) indicators where the activity was around the 3rd decile (see Fig. 3). We refer to these students as having a ***disengaged engagement state.***

The comparison of LMS indicator activity levels among the three clusters (Table 2) was statistically significant with an effect size (ε^2) ranging from 0.278 (FFCb) to 0.798 (SC). All pairwise comparisons were statistically significant ($p < 0.001$).

4.1. Association between engagement and achievement

The association between the *Active* engagement state and the *Achiever* state was positive and statistically significant in most of the courses (indicated as blue rectangles in the top left corner of the plot in Fig. 4). More evidently, the association between the *Achiever*

Table 1
Summary statistics by quantile of the frequency of events per student per course. The total represents the total number of events of all the students in each course.

Course	25%	Median	75%	Total
1	35.50	80.5	187.75	13,018
2	214.00	406.0	577.00	42,753
3	160.50	254.0	359.00	28,635
4	212.00	316.0	422.00	31,888
5	190.00	274.0	418.00	29,828
6	337.00	475.0	643.00	44,490
7	82.00	163.0	262.00	18,749
8	33.00	50.0	355.00	17,452
9	36.50	53.0	68.00	7453
10	31.00	75.0	251.50	13,053
11	81.50	242.0	712.75	38,321
12	22.00	63.0	250.00	12,920
13	22.75	280.0	407.00	23,379
14	22.50	155.0	330.00	19,591
15	18.00	28.0	52.00	4296

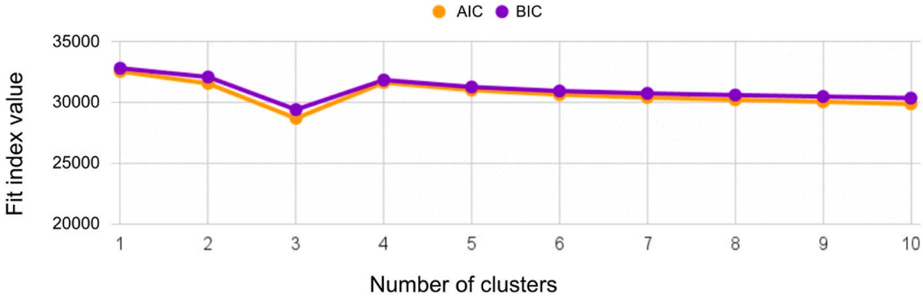

Fig. 1. AIC and BIC values for each number of clusters (LCA). The x-axis represents the number of clusters ranging from 1 to 10. The y-axis represents the value of the fit index for the AIC (in orange) and the BIC (in purple). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

	Active				Average				Disengaged			
	Mean	SD	Boxplot	Histogram	Mean	SD	Boxplot	Histogram	Mean	SD	Boxplot	Histogram
AD	8.51	1.43			4.96	1.74			1.77	0.94		
REG	8.15	1.83			5.11	2.03			1.94	1.03		
FCB	8.46	1.49			4.91	1.83			1.94	1.21		
FFCs	7.95	2.07			4.89	2.21			2.66	1.92		
FFCb	7.25	2.42			5.14	2.51			3.08	2.32		
FLV	7.32	2.41			5.18	2.51			2.90	2.05		
SC	8.71	1.09			4.87	1.55			1.68	0.85		
TD	8.02	2.00			4.94	2.19			2.46	1.69		

Fig. 2. Summary statistics for the LMS indicators for each of the identified clusters.

state and the *Disengaged* state was negative and statistically significant in all courses (indicated as red rectangles in the bottom left corner). Similarly, there was higher statistically significant association between the *Disengaged* state and *Low* achievement state in most of the courses (marked as blue rectangles in the bottom right plot) as well as negative association between the *Disengaged* state and the *Achiever* state (marked as red rectangles in the top right corner). In summary, there was a consistent and statistically significant association between disengagement and low achievement in all the courses, as well as a statistically significant high association between active engagement and high achievement in most of the courses. In other words, low engagement had a consistent association with low achievement.

4.2. Sequence mining

In the previous step, three clusters representing levels of engagement (or e*ngagement states*) were obtained for each student at each course. Such *engagement states* were used to build a sequence of the states of engagement. *Engagement states* were visualized using index plots, where every individual student's sequence is plotted as stacked colored blocks corresponding to course *engagement states* (Fig. 5A). Fig. 5B shows the sequence index plot of students' *achievement states*. Both data channels were temporally aligned and combined, so that each student's engagement state is synchronous with the achievement state in each course (Fig. 5C).

At the top of the three graphs in Fig. 5, there is a distinguishable pattern where students were predominantly disengaged (Fig. 5A) and scoring low (Fig. 5B). Such a combination of disengagement and low scoring is more remarkable in the multi-sequence plot in Fig. 5C. Similarly, another pattern can be observed at the bottom of Fig. 5A, where students were mostly engaged; their corresponding achievement states in Fig. 5B show predominantly high achievement states. Such a combination is further noticeable in the combined multi-channel sequence in Fig. 5C.

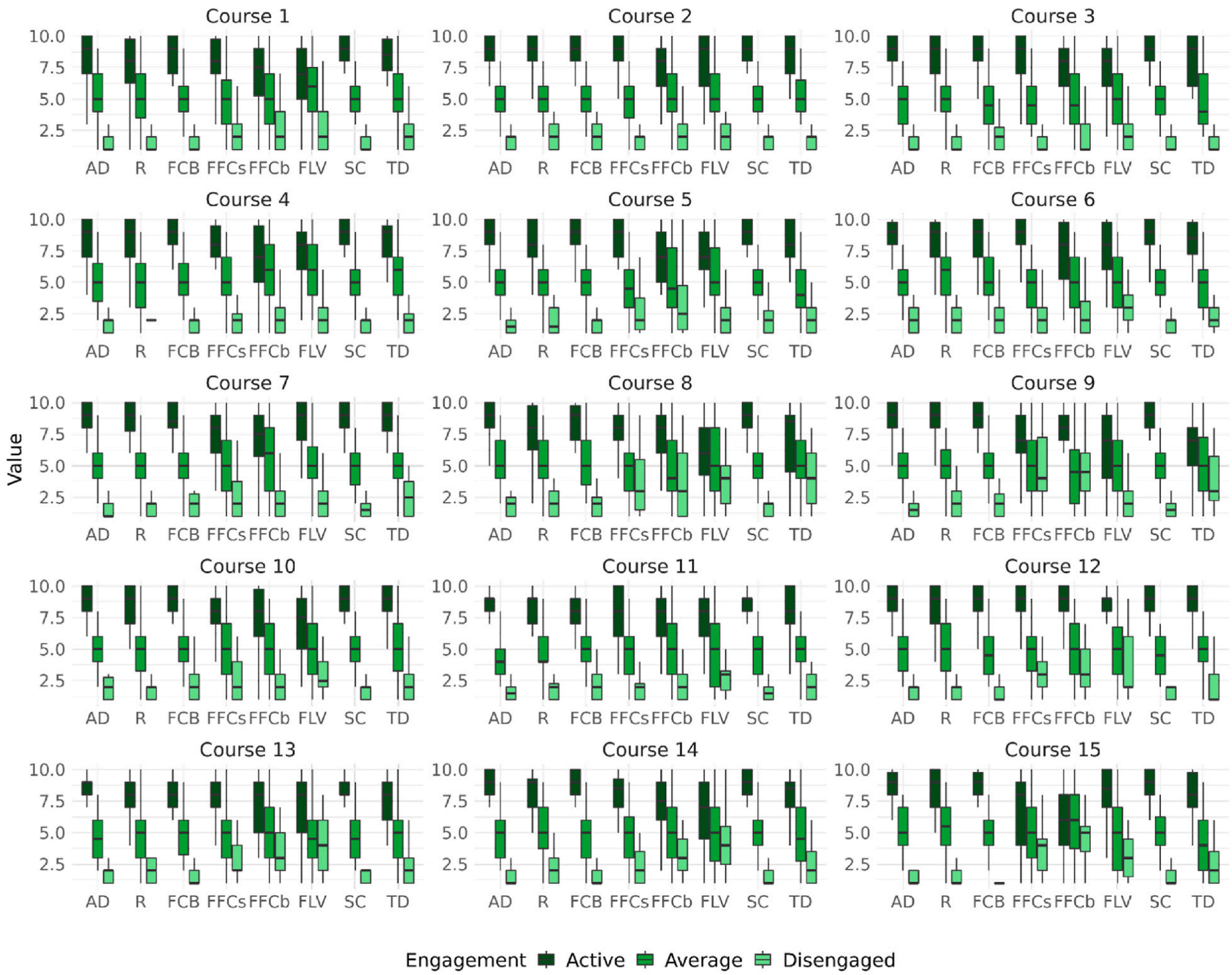

Fig. 3. Box plot of each activity indicator for each engagement state in each course. The indicators are as follows: Active days (AD), Regularity (R), Course browsing frequency (FCB), Forum consuming frequency (FFCs), Forum contributing frequency (FFCb), Lecture viewing frequency (FLV), Session count (SC), Total duration (TD). The x-axis represents each indicator, the y-axis represents the value of each indicator rescaled since each indicator is measured on different units (e.g., duration and frequency), and the color indicates the engagement state. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

Table 2
Kruskal-Wallis and pairwise comparison of the LMS indicators among clusters.

Indicator	χ^2	df	p	ε^2	Active-Average	Active-Disengaged	Average-Disengaged
AD	1014	2	<.001	0.728	<.001	<.001	<.001
REG	853	2	<.001	0.612	<.001	<.001	<.001
FCB	957	2	<.001	0.687	<.001	<.001	<.001
FFCs	637	2	<.001	0.457	<.001	<.001	<.001
FFCb	387	2	<.001	0.278	<.001	<.001	<.001
FLV	431	2	<.001	0.309	<.001	<.001	<.001
SC	1112	2	<.001	0.798	<.001	<.001	<.001
TD	694	2	<.001	0.498	<.001	<.001	<.001

4.3. Clustering of the trajectories

Clustering using MHMM resulted in three distinct groups of students according to their combined engagement-achievement patterns. We named the clusters according to students' profiles at the start of the program: *Engaged high-achieving starters*, *Average starters*, and *Disengaged starters*. The following is a detailed description of each of the three clusters.

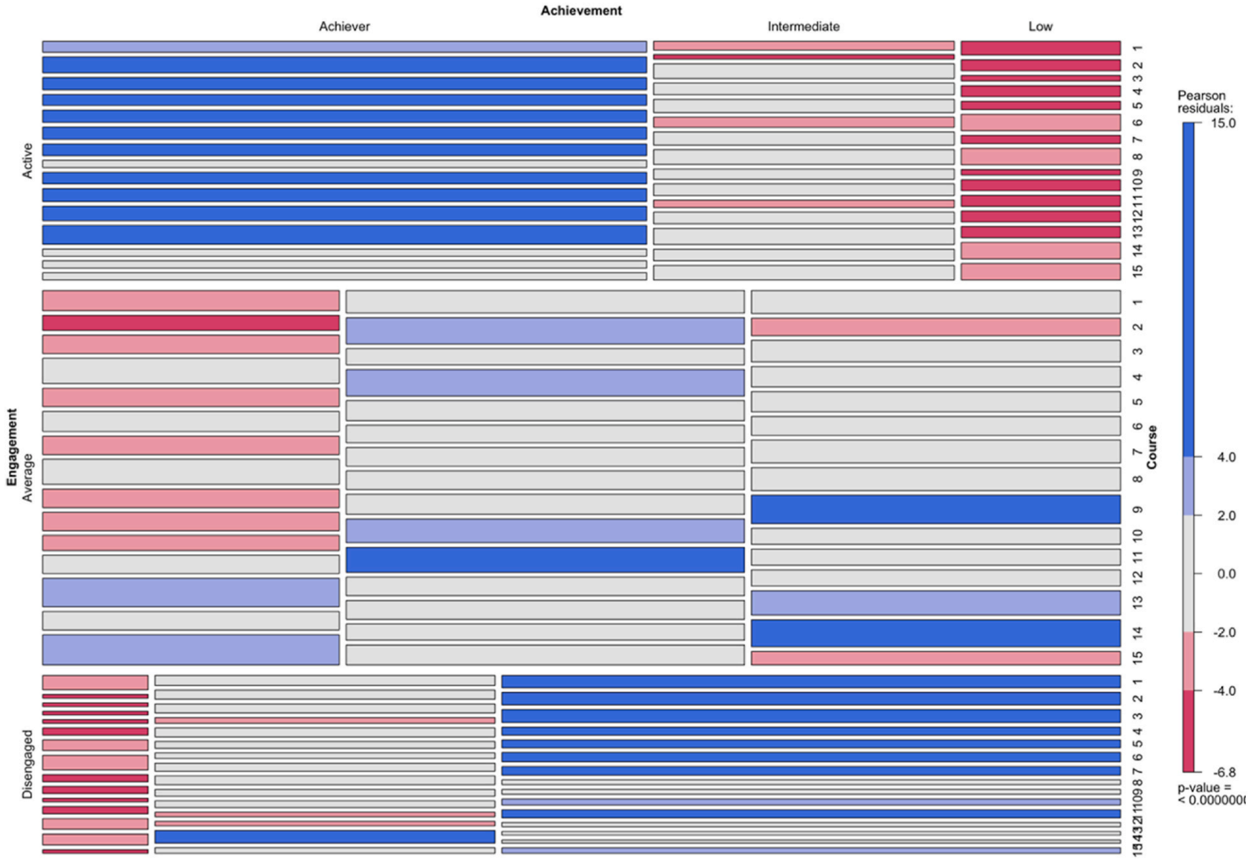

Fig. 4. Mosaic plot showing the association between engagement states and achievement states in each course.

4.3.1. Engaged high-achieving starters (n = 28, 26.4%)

Fig. 6 shows the two channels (engagement at the top and achievement in the middle) for the first cluster, showing predominantly high levels of engagement (*Active* engagement state), as well as high scores, indicated by the dominance of the dark green and dark blue colors. The lower channel represents the combined engagement-achievement trajectory, showing two distinct states. State 1 (dark red) shows a combination of predominantly high engagement (with emission probabilities: 87.4% *Active*, 12.6% *Average)* and high achievement (82.6% *Achiever*, 12.4% *Intermediate*). In turn, State 2 (pink) represents slightly lower levels of engagement (*Active* 32.8%, *Average* 59.1%, *Low* 8.1%) and achievement (*Achiever* 62.5%, *Intermediate* 30.4%*, Low* 7%).

Students in State 1 were more likely to shift to State 2 (transition probability = 11%) as the program advanced, while less likely to shift from State 2 to State 1 (transition probability = 2.1%). Only seven students (25.9%) remained in State 1 throughout the program (fully dark red sequences). The rest remained high achievers but lowered their activity level. For two students, this decrease in activity was temporary and they became active again later in the program.

Thus, in engaged high-achieving starters, decreasing engagement of students in State 2 was not associated with a notable decrease in their achievement level. Therefore, one can conclude that engagement and achievement are predominantly sequentially intertwined with some variability. Put another way, engaged high-achieving starters are able to maintain a rather high achievement level despite occasional moderate decreases in engagement.

4.3.2. Average starters (n = 43, 40.6%)

The second cluster (Fig. 7) represents students with predominantly *Average* or *Disengaged* states (top channel) and various levels of achievement that are mostly *Intermediate* or *Low* (middle channel). Their trajectory shows two distinct combined engagement-achievement states. Students in State 1 (dark purple) had predominantly *Disengaged* (emission probability = 52.2%) or *Average* (43.7%) engagement states, as well as *Intermediate* (51.3%) or *Low* (36.5%) achievement levels.

Students in State 2 (light purple) had comparatively higher engagement with *Average* (58.6%) or *Active* (39%) engagement states, as well as relatively higher levels of achievement (*Intermediate* 43.2%*, Achiever* 36.5%). A total of 15 students (34.9%) of the students remained average-level throughout the program, while the rest improved their engagement and achievement level (most of early in the program, during courses 2–5). In the intermediate group, high or improving engagement was reflected in comparable levels of achievement. However, low levels or non-improving engagement were associated with low achievement and occasional drop-out.

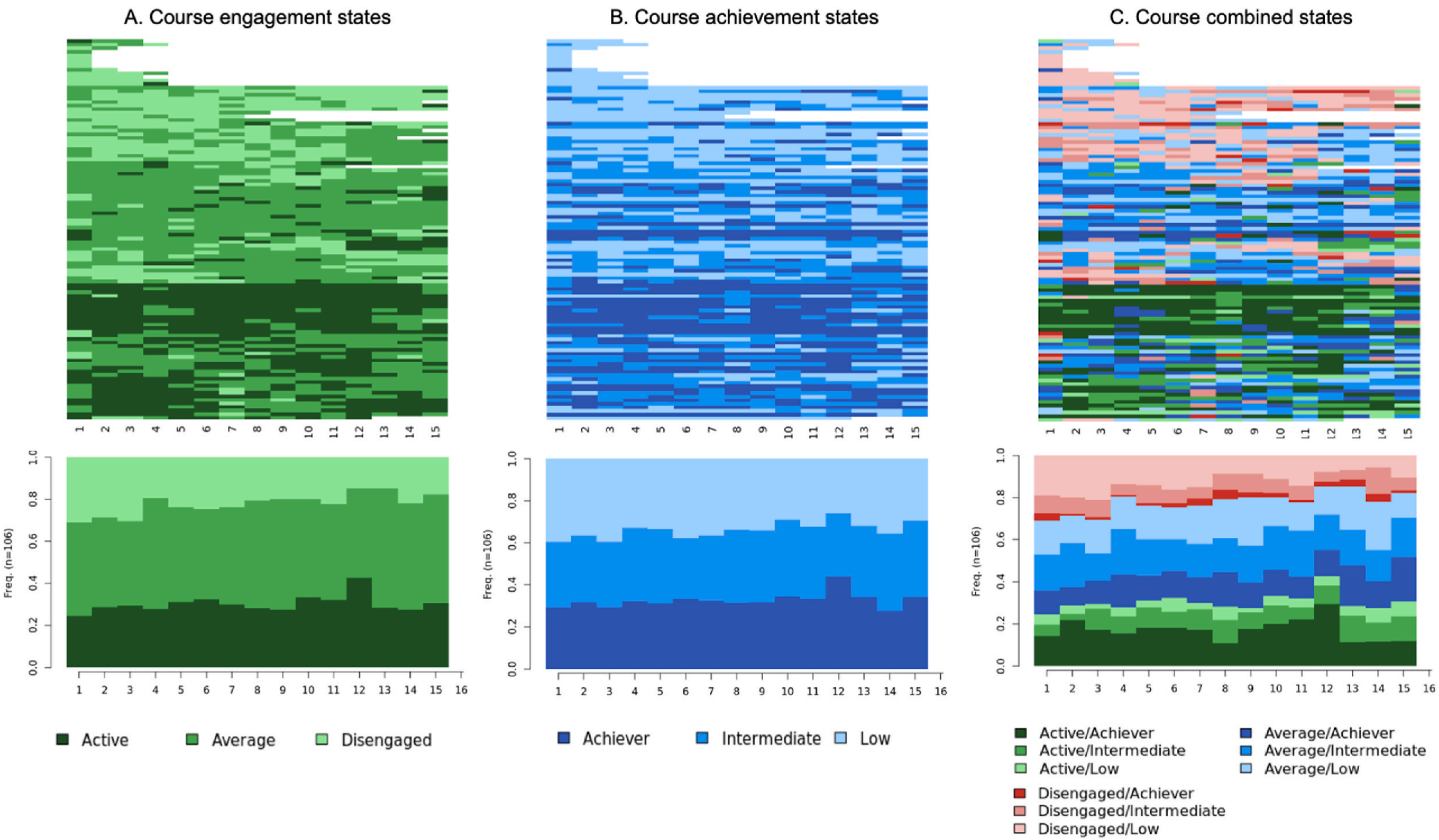

Fig. 5. Sequence index plot (top) and sequence distribution plot (bottom) for A) course engagement states, B) course achievement states, and C) course combined states.

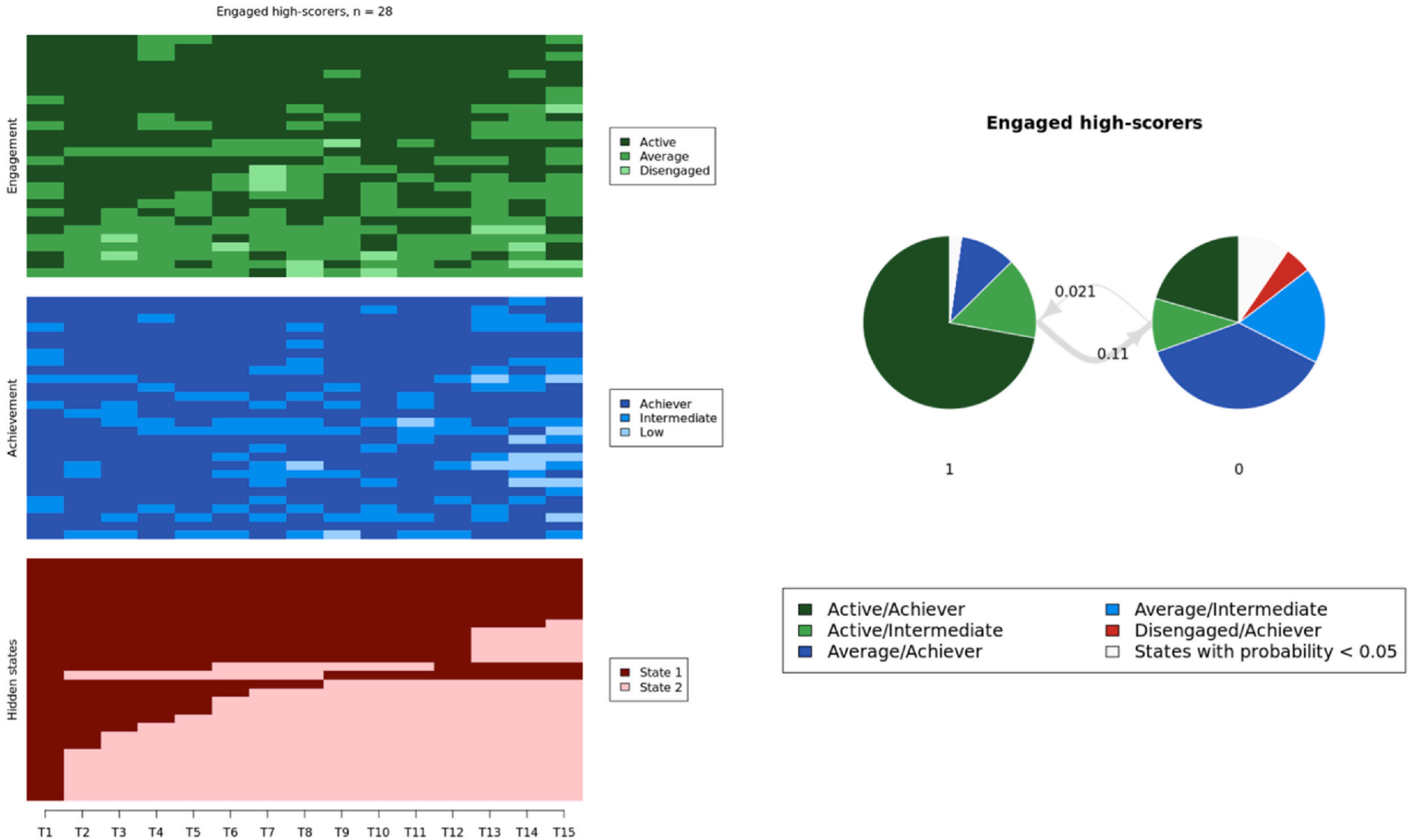

Fig. 6. Characterization of the *Engaged high-achieving starters.* The left side shows the sequence index plot of the engagement states (top), achievement states (middle) and hidden states (bottom). The right side shows the transition diagram.

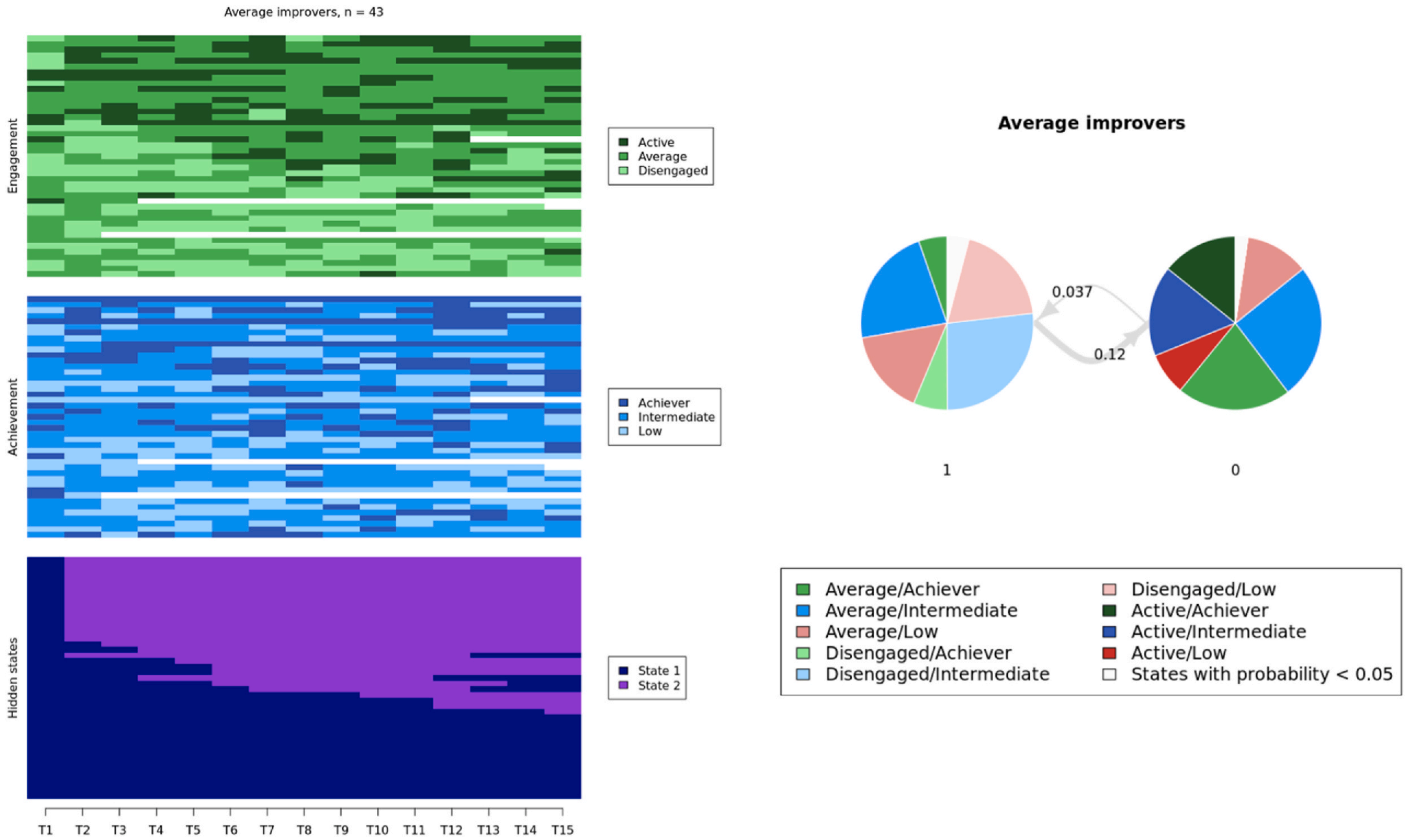

Fig. 7. Characterization of the *Average starters.* The left side shows the sequence index plot of the engagement states (top), achievement states (middle) and hidden states (bottom). The right-hand side shows the transition diagram.

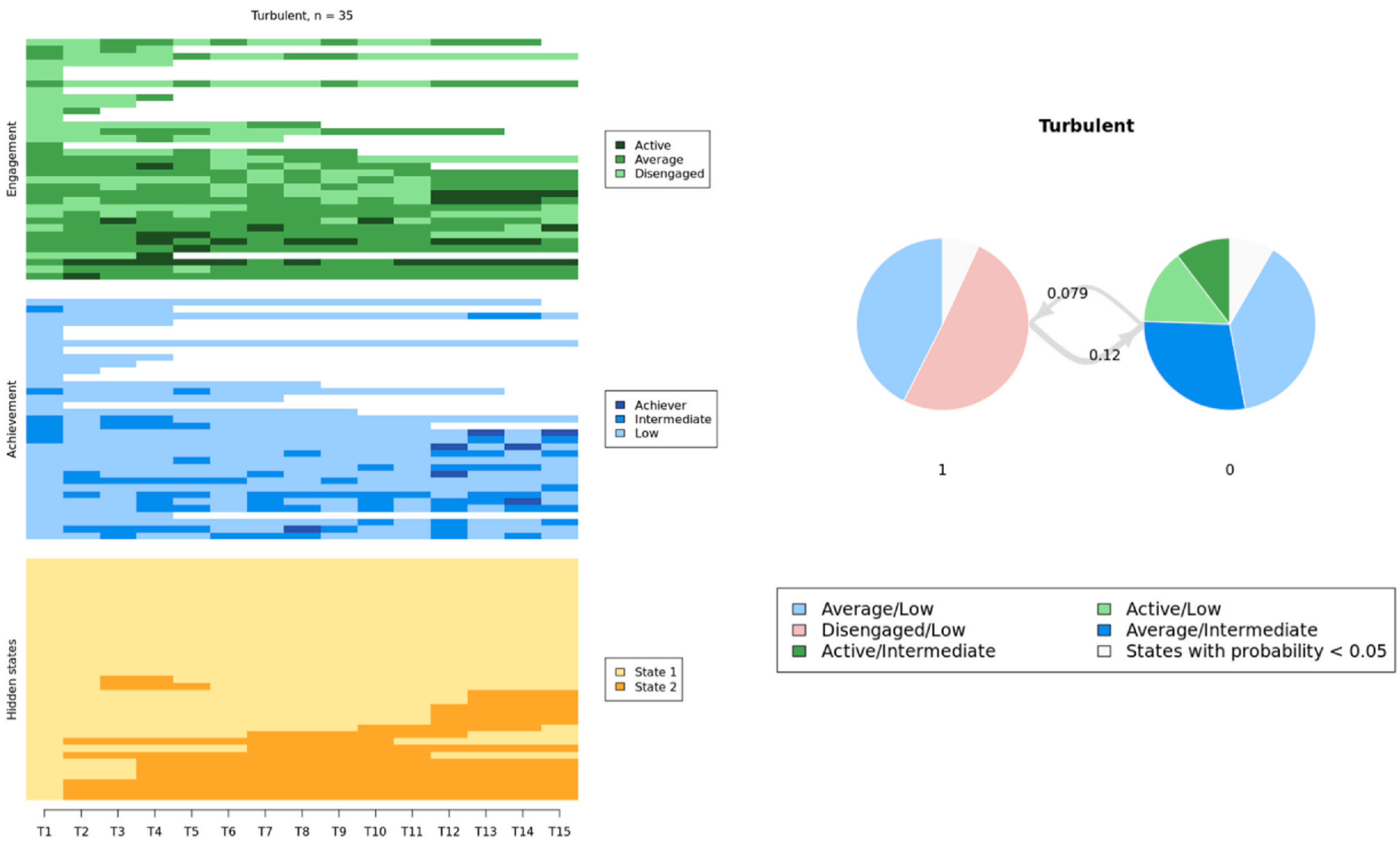

Fig. 8. Characterization of the *Disengaged starters*. The left side shows the sequence index plot of the engagement states (top), achievement states (middle) and hidden states (bottom). The right-hand side shows the transition diagram.

4.3.3. **Disengaged starters** (n = 35, 33.02%)

The third cluster (Fig. 8) represents students who started with predominantly *Disengaged* or *Average* states (top channel) or dropouts and various levels of achievement that are mostly *Low* (middle channel). Their trajectory shows two distinct combined engagement-achievement states. Students in State 1 (the predominant one) had *Disengaged* (54.3%) or *Average* (emission probability = 45.7%) engagement states and were mostly *Low achievers* (93.1%) or *Intermediate* (6.9%).

State 2 represented students with slightly higher levels of both engagement (*Average* 70.6%, *Active* 25.7%, *Disengaged* 3.7%) and achievement (*Low* 54.8%, *Intermediate* 40.4%, *Achiever 0.5%*). Students were unlikely to shift from one state to another (transition probability State 1 to State 2 = 12%, transition probability State 1 to State 2 = 7.9%). Around half of the students in this group (18) were able to shift to a higher engagement level as the program advanced, finishing in a relatively higher achievement state than how they started. Therefore, starting with low engagement is rather consequential, decreasing, or non-improving engagement is associated with a high probability of early drop-out.

5. Discussion

Two issues vie for educators' attention: engagement as a marker for high achievement and disengagement as a prelude to low achievement or possible drop-out. The present work was conducted to study the association between online engagement and achievement across a four-year full program and how such a relationship evolves within students' subpopulations i.e., for whom engagement evolves and for whom it devolves, and how achievement is affected by such changes. We first tested the significance of the statistical association between engagement and achievement states. Then, we modeled the temporal association and studied the variations of associations within students' subgroups.

5.1. *RQ1: How are engagement and achievement intertwined across a full program?*

The first research question aimed at determining to which extent the engagement and achievement states were statistically and consistently associated. Our findings corroborate some of the prior cross sectional research findings (Lei et al., 2018) and extend to novel insights regarding the temporal association and consistency aspects. Our study has shown that the association between online disengagement and low achievement is longitudinal, consistent, and predictable. Nonetheless, high engagement was not always associated with high achievement; average engagement showed mixed results, fluctuating between low achieving (most times), and infrequent high achievement (few times). The consistency and predictability of results are rarely studied topics as most of the existing papers are cross-sectional studies at one time point. Hence, our results show the importance of online disengagement as a consistent indicator of achievement and, on the same token, the weakness of high levels of online engagement as indicators of achievement. Of course, one should factor in the idea that online data is far from comprehensive; notwithstanding, it is unobtrusive, less burdensome, and inclusive of all students across all time points, i.e., does not suffer attritional issues. Although the reported results do not support causal inference, they are actionable. As Hughes et al. (2008) posit, engagement and achievement constitute a dynamic system such that early intervention can alter students' learning trajectories.

5.2. *RQ2: How do student engagement and achievement evolve across a full program across different students' subgroups?*

Our second research question, and the main motivation behind this study, was to model the longitudinal relationship between engagement and academic achievement. This is an issue that has received little attention, especially within the context of online and blended learning. Our findings have indicated that the students were a heterogeneous group, where several clusters of similar behavior exist with distinct trajectories. The *Engaged high-achieving starters* diverged into two subgroups: a group which remained consistently highly engaged, and another group that continued to be moderately engaged. Both groups continued to achieve at comparable levels. A possible explanation might be that students in this group may have "learned how to learn" (Archambault & Dupéré, 2017; Zhen, De Liu et al., 2020), or acquired better learning strategies (Biggs, 1979; Marton & Säljö, 1976) that allowed them to obtain good grades while investing less effort in online learning, or perhaps students' efforts did not change and only their online activities did. Put another way, high achievers are expected to be highly engaged, although a moderate drop in high achievers' engagement is expected to be rather non-consequential.

A similar engaged cluster has been reported by other researchers in online settings using cross-sectional data e.g., (Jovanović et al., 2017; López-Pernas, Saqr, & Viberg, 2021). In face-to-face education research, longitudinal evolution has received more attention. Recent research emphasized the heterogeneity of longitudinal engagement profiles (Archambault & Dupéré, 2017; Saqr and López-Pernas, 2021, 2022, Saqr, López-Pernas, Jovanović, & Gašević, 2023; Zhen, De Liu, et al., 2020); that is, different subgroups exist with different engagement profiles and distinct patterns of temporal evolution, i.e., some students are consistently engaged across time, others have increasing engagement and others have a declining pattern, e.g., (Borup et al., 2020). Yet, the association between achievement and engagement is always expressed as the final outcome at the end of the program using e.g., regression models. A contribution of our study is tracking the evolution across time, and across subgroups which regression models do not offer.

The second cluster in our study was the *Average starters*, which included students who started in an intermediate engagement and achievement state. This group exhibited two subgroups: *improvers* whose engagement and achievement improved over time with lower dropout rates, and another subgroup that remained averagely engaged, with moderate achievement and some dropouts. An intermediate engagement cluster has been reported in cross-sectional research, e.g., as "get-it-done" or "satisfying module requirements" (Mirriahi, Jovanovic, Dawson, Gašević, & Pardo, 2018). In longitudinal research, both improving and declining levels of engagement have been reported (Archambault & Dupéré, 2017; Zhen, De Liu et al., 2020), while others described a “fluid” pattern where students alternate between higher and lower levels of engagement (Saqr and López-Pernas, 2021). Whereas our results can be comparable to such studies regarding engagement. Our findings point to a tied combination in engagement and achievement in this group, i.e., if engagement improves, it is likely that achievement follows, lack of improvement would be probably associated with average achievement and carries a risk of dropping out. The description of such groups, their trend of evolution, and how achievement follows improvement is another contribution of this study.

The last group, the *Disengaged starters,* was —similar to the previous two clusters— composed of two subgroups: a subgroup that improved their engagement level and obtained relatively acceptable grades, and another group that was troubled and had several dropouts. Disengagement in this group was the most consequential and led to attrition while, in the previous two groups, declining engagement was not associated with such a high attrition rate. A disengaged cluster has been described in most longitudinal studies (Archambault & Dupéré, 2017; Saqr and López-Pernas, 2021; Zhen, De Liu, et al., 2020). Our results showed that maintaining or improving engagement in this group is likely to be related to acceptable levels of achievement, as well as lower levels of attrition.

In summary, we describe the trajectories of interactions between engagement and achievement which were not described in previous research. We found that a trajectory of high achievement and engaged students was a rather stable and consistent trajectory. The two other trajectories were less engaged and had comparatively lower achievement levels. In the two said trajectories, the succession of states was important. That is, the ability of the students to improve was the decisive factor. Students who were able to improve their engagement were able to graduate and maintain a reasonable academic trajectory. Students who failed to catch up had an ominous trajectory and dropped out in larger proportions.

5.3. Discussion of the methods

Our study has applied the novel methods of multi-channel sequence analysis with MHMM to combine and align temporal engagement and achievement events. These methods have several advantages. First, they provide a summarizing approach to highly dimensional data. While statistically complex, they have a visually intuitive and interpretable toolset that enable researchers to make sense of the results (Gauthier et al., 2010; Helske et al., 2018). MHMM takes temporality into account in all statistical processes such as clustering, so that the discovered trajectories are based on temporal alignment. Thirdly, MHMM is a multi-state model which allows students to transition between clusters (or states) allowing us to see a more dynamic view of students' groups. For example, our findings of multiple states within the clusters highlighted the subgroups of improvers and their transitions. This approach is more plausible and realistic than traditional methods (e.g., k-means) which assumes fixed cluster memberships with no transitions between clusters or states over time (Helske, Steele, Kokko, Räikkönen, & Eerola, 2015, 2018).

Perhaps the most important potential of MHMM models is that they allow co-variates to be included in the model (Helske et al., 2018). Therefore, researchers can use MHMM to understand cluster membership, moving from a descriptive approach to a more explanatory approach. More importantly, using MHMM, researchers can predict future student trajectories and therefore apply remedial measures when a turbulent trajectory is predicted. For instance, Helske et al. (2021) used a large dataset of Finnish registry data to predict the employment trajectory of Finnish individuals using socio-economic and health data. Their results showed that individuals with a late entry into the labor market or an advantageous background are more likely to have more stable employment.

When preparing this study, we have experimented with a long list of longitudinal methods. Most existing longitudinal methods —e.g., longitudinal k-means, group-based trajectory modeling (GBTM), and Growth Mixture Modeling (GMM) — are well-suited for univariate modeling (Herle et al., 2020). Furthermore, such methods assign a student to a single cluster over time, i.e., a disengaged student will always be within the disengaged cluster. In doing so, they overlook the fact that students can —and do— change over time. Therefore, we had to resort to sequence methods that harness the temporal features and have potential for explaining the changes across time. The same methods have been recently used in similar research problems (e.g., Saqr & López-Pernas, 2022).

A common thread of learning analytics research has focused on creating models to predict students' course performance (Bergdahl, Jalal Nouri, Karunaratne, Afzaal, & Saqr, 2020; Jovanović et al., 2021). Although many studies have created well-performing predictive models within the examined course, the transfer of the models to other courses proved challenging even within the same context (Conijn, Snijders, Kleingeld, & Matzat, 2017). In this study, we have opted for a different approach based on student's relative activity within the course and examined how such activity is indicative of future course activity and of the overall program trajectory. While we have not solved the issue of transferring predictive models to future courses, we have shown one way learning analytics data could work and possibly alert us to the students who may need help. What is more, longitudinal studies in face-to-face settings relied on data collection through surveying students about their engagement: while this method has brought immense insights, it requires effortful data collection, which is prone to attrition of research subjects. Our study has shown that data extracted from online environments —while imperfect— can offer a reasonable alternative.

Capturing the full gamut of engagement is an effortful process that entails multiple concerted data channels from several sources such as video recording, self-reports, trace data, eye tracking, think aloud, physiological sensors (Azevedo, 2015). Collecting such multi-channel data over extended periods while retaining a representative sample of students is rather difficult in longitudinal studies. A balance has to be struck between length of study, resolution of data, and feasibility.

5.4. Limitations of our study

Our study has limitations related to data, methods, and analysis. While online environments offer an unobtrusive method for collecting data about students, they have several limitations regarding scope (e.g., being able only to track clicks), accuracy (e.g., issues with online time and multitasking), and applicability (e.g., limited to activity in front of the screen). In our case, modeling online data was a reasonable indicator of engagement. Nonetheless, it fell short of completely reflecting the full picture or explaining all its aspects (e.g., high achievement with average online engagement not explained with current information at hand).

Our study does not account for emotional engagement, which is an important aspect of engagement. Including emotional engagement may broaden our understanding of aspects of engagement and achievement. Nevertheless, there is a trade-off between ease of collecting unobtrusive data (e.g., online) and collecting comprehensive data of all students' activities as well as all dimensions of engagement (e.g., multimodal and video data). However, the latter approach may be difficult in longitudinal studies and thus automated methods could help. Future research could explore the longitudinal influence of emotions on students' engagement and achievement.

We have used a clustering method to group the students into engagement states which has limitations. For instance, the clustering algorithm has to assign students to one cluster, and therefore, students who may have small differences in activity may have been assigned different clusters or different states. While this has possibly occurred in some clusters, the trajectories in our study rely on fifteen longitudinal cluster assignments (in fifteen courses) and, therefore, it is very unlikely that this has affected a significant part of any student trajectory that led to misclassification. Our study was based on healthcare education and therefore generalization to another context remains to be investigated. Furthermore, the clustering method offers a reductionist approach to a student's overall activity by compressing all his/her activity in just one cluster (e.g., engaged). Nonetheless, a reductionist approach was necessary to study four years of data without being overly complex. Lastly, the MHMM method used in this study has known limitations, mainly being sophisticated, as well as computationally inefficient in memory and processing time. However, recent implementations have offered several solutions to such problems (Helske et al., 2018).

6. Conclusions

This study was performed to investigate the sequential succession of engagement states and their association with achievement, how they unfold over time and for whom such changes happen. A prime question we attempted to answer was if changes in engagement reflect a change in achievement and, if so, how consistent are changes in online engagement interacting with the trajectory of longitudinal achievement.

Our results showed that online engagement at any single time-point (course in our case) is not a consistent indicator of high achievement (RQ1). It takes more than a single point of time to reliably forecast high achievement. As a corollary, longitudinal engagement, improving engagement, or previous high achievement are all signs of future high achievement. When longitudinally highly engaged students occasionally decline to an *average* level of engagement, they are able to return to their high levels of engagement. Longitudinally high-achieving students maintained their level of academic achievement even when they had an average level of engagement later in the program. As such, longitudinal high grades, or longitudinal high levels of engagement (either separately or combined) were indicators of a stable academic trajectory in which students remained engaged —at least on average— and had a higher level of achievement.

On the other hand, disengagement at any time point was consistently associated with lower achievement. Improving to a higher level of engagement was associated with —at least— acceptable achievement levels and rare dropouts. On the same token, students who failed to improve were at risk, having the most dropouts. Lack of improvement or "catching up" may be a more ominous sign that should be proactively addressed. Therefore, students who had lower levels of engagement or achievement (at any time point) and, in particular, at the initial stages of the program should be closely followed for signs of improvement, or lack thereof, as such early courses may be critical in the long term. Students with average levels of engagement should similarly be closely monitored, as such a group was the most fluid.

There are several opportunities for future research to uncover the dynamics of engagement and achievement. Transition models with covariates can be used to examine heterogeneity of changes across time, subgroups as well as to investigate the factors that influence students transitioning to a higher or lower engagement state. Another opportunity lies in predicting future states using the past. That is, predicting the trajectory a student would take across the program and proactively students in need. Within-student (idiographic) or personal trajectories are a future direction that has started to emerge in the literature and would probably lead to sharpening our focus on precise education and personalization.

Credit author statement

Mohammed Saqr: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Writing – original draft, Writing – review & editing, Sonsoles López-Pernas: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Validation, Visualization, Writing – original draft, Writing – review & editing, Satu Helske: Formal analysis, Investigation, Methodology, Project administration, Software, Validation, Visualization, Writing – original draft, Writing – review & editing, Stefan Hrastinski: Writing – original draft, Writing – review & editing

Data availability

The authors do not have permission to share data.

Acknowledgement

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 the first author. The First and second authors would also like to thank January Collective for their generous support.

References

Agrawal, R., & Srikant, R. (1995). Mining sequential patterns. In *Proceedings of the eleventh international conference on data engineering* (Vols. 3–14)IEEE.
Alves, P., Morais, C., & Miranda, L. (2017). Learning analytics in higher education: Assessing learning outcomes. In *Proceedings of the European conference on E-learning, ECEL, edited by mesquita A. Peres P. Mesquita A., 2010-octob:25–32*. Academic Conferences Limited.
Archambault, I., & Dupéré, V. (2017). Joint trajectories of behavioral, affective, and cognitive engagement in elementary school. *The Journal of Educational Research, 110*(2), 188–198.
Asikainen, H., & Gijbels, D. (2017). Do students develop towards more deep approaches to learning during studies? A systematic review on the development of students' deep and surface approaches to learning in higher education. *Educational Psychology Review, 29*(2), 205–234.
Azevedo, R. (2015). Defining and measuring engagement and learning in science: Conceptual, theoretical, methodological, and analytical issues. *Educational Psychologist, 50*(1), 84–94.
Barthakur, A., Kovanovic, V., Joksimovic, S., George, S., Richey, M., & Dawson, S. (2021). Assessing program-level learning strategies in MOOCs. *Computers in Human Behavior, 117*(April), Article 106674.
Bergdahl, N., Jalal Nouri, Karunaratne, T., Afzaal, M., & Saqr, M. (2020). *Learning analytics for blended learning: A systematic review of theory, methodology, and ethical considerations* (Vol. 2). International Association of Online Engineering (IAOE).
Biggs, J. (1979). Individual differences in study processes and the quality of learning outcomes. *Higher Education, 8*(4), 381–394.
Borup, J., Graham, C. R., Leanna Archambault, R. E. W., & Spring, K. J. (2020). Academic communities of engagement: An expansive lens for examining support structures in blended and online learning. *Educational Technology Research & Development: ETR & D, 68*(2), 807–832.
Bryan, C. J., Tipton, E., & Yeager, D. S. (2021). Behavioural science is unlikely to change the world without a heterogeneity revolution. *Nature Human Behaviour, 5*(8), 980–989.
Christenson, S. L., Wylie, C., & Reschly, A. L. (2012). *Handbook of research on student engagement*.
Conijn, R., Snijders, C., Kleingeld, A., & Matzat, U. (2017). Predicting student performance from LMS data: A comparison of 17 blended courses using Moodle LMS. *IEEE Transactions on Learning Technologies, 10*(1), 17–29.
Davey, D. D., Kaplan, D. R., & Claire, W. M. (2014). Strong performance on the progressive evaluation of competency fellowship final examination predicts American board of pathology certification. *Journal of the American Society of Cytopathology, 3*(5), 269–273.
De Champlain, & André, F. (2018). Standard setting methods in medical education. In *Understanding medical education, 347–59*. Chichester, UK: John Wiley & Sons, Ltd.
Dewar, A., Hope, D., Jaap, A., & Cameron, H. (2021). Predicting failure before it happens: A 5-year, 1042 participant prospective study. *Medical Teacher, 0*(0), 1–9.
Dziuban, C., Moskal, P., Cavanagh, T., & Watts, A. (2012). Analytics that inform the university: Using data You already have. *Journal of Asynchronous Learning Networks, 16*(3), 21–38.
Fincham, E., Whitelock-Wainwright, A., Kovanović, V., Joksimović, S., van Staalduinen, J.-P., & Gašević, D. (2019). Counting clicks is not enough. In *Proceedings of the 9th international conference on learning analytics & knowledge*. New York, NY, USA: ACM. https://doi.org/10.1145/3303772.3303775.
Firke, S. (2020). *Janitor: Simple tools for examining and cleaning dirty data*. R Package Version.
Frishkoff, G. A., Collins-Thompson, K., Hodges, L., & Scott, C. (2016). Accuracy feedback improves word learning from context: Evidence from a meaning-generation task. *Reading and Writing, 29*(4), 609–632.
Froiland, J. M., & Oros, E. (2014). Intrinsic motivation, perceived competence and classroom engagement as longitudinal predictors of adolescent reading achievement. *Educational Psychology Review, 34*(2), 119–132.
Fryer, L. K. (2017). (Latent) transitions to learning at university: A latent profile transition analysis of first-year Japanese students. *Higher Education, 73*(3), 519–537.
Gabadinho, A., Gilbert, R., Studer, M., & Nicolas, S. M. (2009). *Mining sequence data in R with the TraMineR package: A users guide for version 1.2* (Vol. 1). Geneva: University of Geneva.
Gašević, D., Jovanović, J., Pardo, A., & Dawson, S. (2017). Detecting learning strategies with analytics: Links with self-reported measures and academic performance. *Journal of Learning Analytics, 4*(2), 113–128.
Gauthier, J. A., Widmer, E. D., Bucher, P., & Notredame, C. (2010). Multichannel sequence analysis applied to social science data. *Sociological Methodology, 40*(1), 1–38.
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.
Halverson, L. R., & Graham, C. R. (2019). Learner engagement in blended learning environments: A conceptual framework. *Online Learning, 23*(2). https://doi.org/10.24059/olj.v23i2.1481
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
Helske, S., Helske, J., & Eerola, M. (2018). *Combining sequence analysis and hidden Markov models in the analysis of complex life sequence data*.
Helske, S., Keski-Säntti, M., Kivelä, J., Juutinen, A., Kääriälä, A., Gissler, M., et al. (2021). Predicting the stability of early employment with its timing and childhood social and health-related predictors: A mixture Markov model approach. *SocArXiv*. https://doi.org/10.31235/osf.io/qkcxs
Helske, S., Steele, F., Kokko, K., Räikkönen, E., & Eerola, M. (2015). Partnership Formation and dissolution over the life course: Applying sequence analysis and event history analysis in the study of recurrent events. *Longitudinal and Life Course Studies, 6*(1), 1–25.
Henrie, C. R., Halverson, L. R., & Graham, C. R. (2015). Measuring student engagement in technology-mediated learning: A review. *Computers & Education, 90*, 36–53.
Herle, M., Micali, N., Mohamed, A., Loos, R., Bryant-Waugh, R., Hübel, C., et al. (2020). Identifying typical trajectories in longitudinal data: Modelling strategies and interpretations. *European Journal of Epidemiology, 35*(3), 205–222.
Hickendorff, M., Edelsbrunner, P. A., McMullen, J., Schneider, M., & Kelly, T. (2018). Informative tools for characterizing individual differences in learning: Latent class, latent profile, and latent transition analysis. *Learning and Individual Differences, 66*(vember 2017), 4–15.
Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics, Theory and Applications, 6*(2), 65–70.
Hughes, J. N., Luo, W., Kwok, O.-M., & Loyd, L. K. (2008). Teacher-student support, effortful engagement, and achievement: A 3-year longitudinal study. *Journal of Educational Psychology, 100*(1), 1–14.
Jimerson, S., Egeland, B., & Teo, A. (1999). A longitudinal study of achievement trajectories: Factors associated with change. *Journal of Educational Psychology, 91*(1), 116–126.
Jishan, S. T., Rashu, R. I., Haque, N., & Rahman, R. M. (2015). Improving accuracy of students' final grade prediction model using optimal equal width binning and synthetic minority over-sampling technique. *Decision Analytics, 2*(1), 1–25.
Joksimović, S., Poquet, O., Kovanović, V., Dowell, N., Mills, C., Gašević, D., et al. (2018). How do we model learning at scale? A systematic review of research on MOOCs. *Review of Educational Research*. https://doi.org/10.3102/0034654317740335
Jovanović, J., Dawson, S., Joksimović, S., & George, S. (2020). Supporting actionable intelligence: Reframing the analysis of observed study strategies. In *Proceedings of the tenth international conference on learning analytics & knowledge, 161–70. LAK '20*. New York, NY, USA: Association for Computing Machinery.

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*(April), 74–85.

Jovanović, J., Mirriahi, N., Gašević, D., Dawson, S., & Pardo, A. (2019). Predictive power of regularity of pre-class activities in a flipped classroom. *Computers in Education*. https://doi.org/10.1016/j.compedu.2019.02.011

Jovanović, J., Saqr, M., Joksimović, S., & Gašević, D. (2021). Students matter the most in learning analytics: The effects of internal and instructional conditions in predicting academic success. *Computers & Education, 172*(April), Article 104251.

Kinai, T., Ndambuki, P., & Peter, M. (2019). *Causal attributions as correlates of secondary school students' academic achievement*. Education Research International. https://doi.org/10.1155/2019/1950753, 2019.

King, R. B. (2015). Sense of relatedness boosts engagement, achievement, and well-being: A latent Growth model study. *Contemporary Educational Psychology, 42*, 26–38.

Kovanović, V., Gašević, D., Joksimović, S., Hatala, M., & Adesope, O. (2015). Analytics of communities of inquiry: Effects of learning technology use on cognitive presence in asynchronous online discussions. *The Internet and Higher Education, 27*, 74–89.

Kovanović, V., Joksimović, S., Gašević, D., James, O., Anne-Marie, S., & Amy, W. (2016). Profiling MOOC course returners: How does student behavior change between two course enrollments?. In *Proceedings of the third (2016) ACM conference on learning @ scale, 269–72. L@S '16*. New York, NY, USA: Association for Computing Machinery.

Kovanović, V., Joksimović, S., Poquet, O., Hennis, T., Pieter de Vries, Hatala, M., et al. (2019). Examining communities of inquiry in massive open online courses: The role of study strategies. *The Internet and Higher Education, 40*(January), 20–43.

Lei, H., Cui, Y., & Zhou, W. (2018). Relationships between student engagement and academic achievement: A meta-analysis. *Social Behavior and Personality: International Journal, 46*(3), 517–528.

Li, Y., & Lerner, R. M. (2011). Trajectories of school engagement during adolescence: Implications for grades, depression, delinquency, and substance use. *Developmental Psychology, 47*(1), 233–247.

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.

Lyubovnikova, J., Napiersky, U., & Vlachopoulos, P. (2015). How are task reflexivity and intercultural sensitivity related to the academic performance of MBA students? *Studies in Higher Education, 40*(9), 1694–1714.

Martin, F., & Borup, J. (2022). *Online learner engagement: Conceptual definitions, research themes, and supportive practices*. Educational Psychologist.

Martin, F., Sun, T., & Westine, C. D. (2020). A systematic review of research on online teaching and learning from 2009 to 2018. *Computers & Education, 159* (December), Article 104009.

Marton, F., & Säljö, R. (1976). On qualitative differences in learning: I-outcome and process. *British Journal of Educational Psychology, 46*(1), 4–11.

Matcha, W., Gašević, D., Ahmad Uzir, Nora'ayu, Jovanović, J., Pardo, A., Lim, L., et al. (2020). Analytics of learning strategies: Role of course design and delivery modality. *Journal of Learning Analytics, 7*(2), 45–71.

Matcha, W., Gašević, D., Ahmad Uzir, Nora'ayu, Jovanović, J., Pardo, A., Maldonado-Mahauad, J., et al. (2019). Detection of learning strategies: A comparison of process, sequence and network analytic approaches. In *Transforming learning with meaningful technologies, 525–40*. Springer International Publishing.

Mirriahi, N., Jovanovic, J., Dawson, S., Gašević, D., & Pardo, A. (2018). Identifying engagement patterns with video annotation activities: A case study in professional development. *Australasian Journal of Educational Technology, 34*(1). https://doi.org/10.14742/ajet.3207

Norcini, J. J. (2003). Setting standards on educational tests. *Medical Education, 37*(5), 464–469.

Ostertagova, E., Ostertag, O., & Kováč, J. (2014). Methodology and application of the kruskal-wallis test. In *, Vol. 611. Applied mechanics and materials* (pp. 115–120). Trans Tech Publ.

Patel, M. D., Tomblinson, C. M., Benefield, T., Ali, K., DeBenedectis, C. M., England, E., et al. (2020). The relationship between us medical licensing examination step scores and ABR Core examination outcome and performance: A multi-institutional study. *Journal of the American College of Radiology: JACR, 17*(8), 1037–1045.

Porcu, M., & Giambona, F. (2017). Introduction to latent class analysis with applications. *The Journal of Early Adolescence, 37*(1), 129–158.

Postareff, L., Anna, P., & Lindblom-Ylänne, S. (2015). Factors contributing to changes in a deep approach to learning in different learning environments. *Learning Environments Research, 18*(3), 315–333.

Postareff, L., Lindblom-Ylänne, S., & Anna, P. (2014). Explaining university students' strong commitment to understand through individual and contextual elements. *Frontline Learning Research*. https://doi.org/10.14786/flr.v2i1.63

R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing Version 4.0. 2. .

Redmond, P., Lindy Anne Abawi, Brown, A., Henderson, R., Heffernan, A., Lindy Anne Abawi, Brown, A., & Henderson, R. (2018). An online engagement framework for higher education. *Online Learning, 22*(1), 183–204.

Redmond, P., Heffernan, A., Abawi, L., Brown, A., & Henderson, R. (2018). An online engagement framework for higher education. *Online Learning, 22*(1). https://doi.org/10.24059/olj.v22i1.1175

Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance: A systematic review and meta-analysis. *Psychological Bulletin, 138*(2), 353–387.

Rosato, N. S., & Judith, C. B. (2012). Latent class analysis: A method for capturing heterogeneity. *Social Work Research, 36*(1), 61–69.

Rotgans, J. I., & Schmidt, H. G. (2011). Cognitive engagement in the problem-based learning classroom. *Advances in Health Sciences Education: Theory and Practice, 16* (4), 465–479.

Sander, Pete, and Information Services. (2016). *Using learning analytics to predict academic outcomes of first-year students in higher education* (Vol. 1277, pp. 2–41). CAPSTONE REPORT Pete Sander Manager, Information Services Oregon State University University of Oregon Applied Information Management Program Spring, 800.

Saqr, M., & López-Pernas, S. (2021). The longitudinal trajectories of online engagement over a full program. *Computers & Education, 175*, Article 104325.

Saqr, M., Fors, U., & Tedre, M. (2017). How learning analytics can early predict under-achieving students in a blended medical education course. *Medical Teacher, 39* (7), 757–767. https://doi.org/10.1080/0142159X.2017.1309376

Saqr, M., & López-Pernas, S. (2022). How CSCL roles emerge, persist, transition, and evolve over time: A four-year longitudinal study. *Computers & Education, 189* (104581), Article 104581.

Saqr, M., López-Pernas, S., Jovanović, J., & Gašević, D. (2023). Intense, turbulent, or wallowing in the mire: A longitudinal study of cross-course online tactics, strategies, and trajectories. In *, Vol. 57. The Internet and Higher Education*. Elsevier BV, Article 100902. https://doi.org/10.1016/j.iheduc.2022.100902.

Sinatra, G. M., Heddy, B. C., & Lombardi, D. (2015). The challenges of defining and measuring student engagement in science. *Educational Psychologist, 50*(1), 1–13.

Schneider, M., & Preckel, F. (2017). Variables associated with achievement in higher education: A systematic review of meta-analyses. *Psychological Bulletin, 143*(6), 565–600.

Tomczak, M., & Tomczak, E. (2014). The need to report effect size estimates revisited. An overview of some recommended measures of effect size. *Trends in Sport Sciences, 21*(1), 19–25.

Vermunt, J. K., Bac, T., & Jay, M. (2008). Latent class models in longitudinal research. *Handbook of Longitudinal Research: Design, Measurement, and Analysis*, 373–385.

Wang, M.-T., & Degol, J. (2014). Staying engaged: Knowledge and research needs in student engagement. *Child Development Perspectives, 8*(3), 137–143.

Wang, M.-T., & Eccles, J. S. (2013). School context, achievement motivation, and academic engagement: A longitudinal study of school engagement using a multidimensional perspective. *Learning and Instruction, 28*(December), 12–23.

Weller, B. E., Bowen, N. K., & Faubert, S. J. (2020). Latent class analysis: A guide to best practice. *Journal of Black Psychology, 46*(4), 287–311.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., et al. (2019). Welcome to the tidyverse. *Journal of Open Source Software, 4*(43), 1686.

Wigfield, A., Eccles, J. S., Ulrich, S., Roeser, R. W., & Davis-Kean, P. (2007). Development of achievement motivation. In *Handbook of child psychology*. Hoboken: John Wiley & Sons, Inc.

Wood, D. F. (2003). Problem based learning ABC of learning and teaching in medicine: Topic collections ABC of learning and teaching in medicine problem based learning. *BMJ, 326*(7384), 328–330.

Wylie, C., & Hodgen, E. (2012). Trajectories and patterns of student engagement: Evidence from a longitudinal study. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (Vols. 585–99). Boston, MA: Springer US.

You, S., & Jill Sharkey. (2009). Testing a developmental–ecological model of student engagement: A multilevel latent Growth curve analysis. *Educational Psychology, 29*(6), 659–684.

Zhang, D., Lee, K., & Lee, I. (2018). Hierarchical trajectory clustering for spatio-temporal periodic pattern mining. *Expert Systems with Applications, 92*(February), 1–11.

Zhen, R., De Liu, R., Wang, M. T., Ding, Y., Jiang, R., Fu, X., et al. (2020). Trajectory patterns of academic engagement among elementary school students: The implicit theory of intelligence and academic self-efficacy matters. *British Journal of Educational Psychology, 90*(3), 618–634.