The trajectory of early disengagement: A longitudinal learning analytics study over a full program

Mohammed Saqr (✉ mmas3@kth.se)

KTH Royal Institute of Technology <https://orcid.org/0000-0001-5881-3109>

Sonsoles López-Pernas

Universidad Politécnica de Madrid <https://orcid.org/0000-0002-9621-1392>

Research Article

Keywords: Learning analytics, early disengagement, trajectories of engagement

Posted Date: August 31st, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-855846/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Research on online engagement is abundant. However, most of the available studies have focused on a single course. Therefore, little is known about how students' online engagement evolves over time. Previous research in face-to-face settings has shown that early disengagement has negative consequences on students' academic achievement and graduation rates. This study examines the longitudinal trajectory of students' online engagement throughout a complete college degree. The study followed 99 students over 4 years of college education including all their course data (15 courses and 1383 course enrollments). Students' engagement states for each course enrollment were identified through Latent Class Analysis (LCA). Students who were not engaged at least one course in the first term were labeled as "Early Disengagement", whereas the remaining students were labeled as "Early Engagement". The two groups of students were analyzed using sequence pattern mining methods. The stability (persistence of the engagement state), transition (ascending to a higher engagement state or descending to a lower state), and typology of each group trajectory of engagement are described in this study. Our results show that early disengagement is associated with higher rates of dropout, lower scores and lower graduation rates whereas early engagement is relatively stable. Our findings indicate that it is critical to proactively address early disengagement during a program, watch the alarming signs such as presence of disengagement during the first courses, declining engagement along the program or history of frequent disengagement states.

1 Introduction

Engagement is concerned with students' participation, efforts and time investment to learn, develop and optimize their learning outcomes [1, 2]. The society aspirations for better education as well as the increasing interests in graduating students have prompted an increasing interest in engagement as a malleable quality that holds the hope for better education [3]. Research has consistently confirmed a positive link between engagement and academic achievement and disengagement has been linked to worse outcomes, dropout and behavioral problems [4].

Engagement has several temporal dimensions, e.g., engagement with a learning task, engagement within a course, or engagement through a full program [5]. Longitudinal engagement –the subject of this study– seems to be the least studied, as collecting data for the same students over the years during a complete undergraduate study is not a trivial task and requires exhaustive work over several years [6]. The timeline of longitudinal engagement is often referred to as engagement trajectory, while a trajectory is a sequence of engagement states that unfold over time [6]. Research on engagement trajectories of face-to-face education has shown that engagement has a cross-time correlation pattern, i.e., engagement states are sequentially alike. Therefore, students who are engaged in a course, are more likely to be engaged in the next [6]. Research has also shown that the cross-time correlation varies among subgroups of students (e.g., consistent in engaged students while turbulent in disengaged students) [7].

While research on online engagement is abundant [8], most of the studies have investigated engagement within a course or two [9, 10]. Studies in the higher education context that have studied the longitudinal trajectory of online engagement are almost non-existent. As online learning has become increasingly common and witnessed an exponential growth during the COVID-19 pandemic [11], it has become increasingly important to focus on longitudinal engagement or the lack thereof. Our study addresses this gap, i.e., investigates the longitudinal engagement trajectory throughout a complete university program over four years. Our interest is in how early disengagement or engagement unfolds, and how accurately it predicts the completion of the program.

We split students into two groups: *Early engagement* and *Early disengagement*. We then follow the students along their whole program and study their trajectory of engagement or disengagement, the stability or the lack thereof (transitions between trajectories), and how these trajectories of student engagement predict persistence in the program. Our research questions are:

- **RQ1:** What are the engagement states that students have over a full program and what are the characteristics of such engagement states?
- **RQ2:** What are the differences between the early engaged students and early disengaged students regarding their longitudinal engagement trajectory, transitions between engagement states and stability?
- **RQ3:** What is the relationship between early engagement and students' outcome in terms of graduation (or dropping out) and final grades?

2 Background

The concept of student engagement started to gain mainstream acceptance around the eighties as a means for supporting students' involvement, achievement as well as preventing alienation and dropout [1, 2]. Back then, engagement was defined as "the student's psychological investment in and effort directed toward learning, understanding, or mastering the knowledge, skills, or crafts that academic work is intended to promote" [1]. While agreement on a unified definition of engagement may be far from realizable, there is wide agreement that engagement is a multi-dimensional concept that includes a behavioral dimension (behavior related to learning activities and learning process), an emotional dimension (feelings about school, studies, and peers), and a cognitive dimension (investment in learning and going beyond the required). Such dimensions are dynamically closely interrelated, e.g., positive feelings about school motivate behavioral engagement in school activities, and cognitive engagement [3, 5, 12]. Researchers have demonstrated that all domains of engagement (behavioral, emotional and cognitive) are significant catalysts of academic achievement. Such relationship has been repeatedly confirmed in all levels of education [4, 13, 14].

Engagement has obvious qualities that can be observed, tracked and easily understood by teachers [8, 15, 16]. More importantly, engagement is malleable and therefore, disengaged students are amenable to intervention. Underachievement, and consequently dropout, could be traced back to events which showed

obvious “distress signals” of disengagement [16, 17]. Students may show signs of low effort, truancy, declining grades that may lead to failing a course or more. Several other distress signals can be observed, e.g., missing assignments, procrastination or lack of participation, disinterest in school or social activities [3, 5, 12, 18]. Since persistence in the program or school, and consequently graduation, are the ultimate goals of educational institutions, strategies at keeping learners engaged or at prevention of disengagement have gained increasing attention across the years [3, 19].

2.1 Trajectories of engagement

Research on the trajectories of engagement is inconclusive. Some studies have shown a cross-time correlation between engagement states, in which students who are engaged in a course or an academic year are more likely to remain engaged in the next. However, these cross-time correlations should not be interpreted as an indication that engagement is a fixed trait. In fact, research on the dynamics of engagement offers evidence on the contrary, i.e., that engagement is both malleable and dynamic. Some researchers have reported that engagement declines with time [20], while others claimed that engagement can grow with time under optimal conditions [21]. Such variability has been recently explained by the heterogeneity of students’ subgroups, where subgroups of students have a stable trajectory, while others have a declining or troubled trajectories [7, 22]. The previous studies on the trajectories of engagement have all been conducted in face-to-face settings.

Most of the studies on online engagement have looked into a course or two while the longitudinal online engagement remains an uncharted territory especially in higher education settings [18, 23]. Very few exceptions can be mentioned here. Lust et al. investigated students’ engagement with an online learning tool in two consecutive offerings of a course [10]. They found that students used different strategies in each iteration of the course. However, the students were different in each iteration. In another recent example, which studied a full MOOC program, the authors grouped students into distinct engagement profiles and followed the changes in such profiles across the program. The three clusters they identified were 1) *consistent*: who are engaged are more likely to remain engaged, 2) *get-it-done*: assessment-oriented but still able to follow up with the program, 3) *disorganized*: who are mostly disengaged. Similar results were reported by [24], who identified three trajectories along a professional development course.

2.2 Learning analytics and modeling the engagement

As online learning environments became increasingly common, along with all types of technology-enhanced learning [25], a wealth of online data has driven interest in learning Analytics (LA) for understanding and supporting learning and teaching. Research in LA has produced valuable insights regarding, e.g., engagement patterns, students’ profiles and the relationship between engagement and achievement [9, 26–28]. Within LA, Sequence Pattern Mining (SPM) is among the methods that was established to model the sequential patterns of students learning [18, 29]. SPM accounts for the chronological, time ordered relationships between events [30]. Therefore, the method has been embraced widely by social scientists in modeling life course studies, e.g., the trajectories of careers, marital status or

health trajectories [31]. Our study takes advantage of the developments of LA, life course studies, and data-mining methods to model the longitudinal trajectories of engagement

3 Methods

3.1 Context

The study included a full blended Problem-Based Learning (PBL) higher education program. The learning management system (LMS) is the platform for online collaboration (weekly PBL forum discussions), distribution of lectures, announcements, and interaction with the teachers and peers. The online weekly PBL discussions are mandatory, in which students in small groups discuss a problem scenario that follows the objectives of the course. The courses are similar in educational underpinning (PBL), arranged sequentially (therefore, referred to as blocks). Some practical courses are longitudinal (over the full year) and therefore were excluded as they did not share the same educational underpinning or evaluation method[32].

3.2 Data collection

Every student who joined the program in the academic year 2014–2015 was included in the study ($n = 99$). The logs of the students in the 15 sequential courses (1,383 enrollments) over four years of education (2014–2015, 2015–2016, 2016–2017, 2017–2018) were retrieved from the Moodle LMS. Only learning-related events were collected. We included only students who completed the first term of the college to avoid including students who were dis-interested in the type of program or had early problems moving to the city or any other problems. Seven students were excluded as they early withdrew during the first course and left the program. *Early disengagement* ($n = 43$) group comprised students who had at least one course with a disengagement state in the first term (first two courses). This definition was adopted since the first course is introduction about the teaching and learning methods of the program, whereas the second course is the first subject matter course. As such, disengagement in either of these courses could be cortical. The remainder ($n = 56$) of students belong to the group of *Early engagement*.

Online engagement in learning was operationalized following the approach proposed in [8], through collecting logs of “frequency of logins to website; number and frequency of postings, responses, and views ... or other website resources accessed; time spent creating a post; and time spent online.” To capture the full breadth of students’ activities in the program, three types of indicators were collected: I) indicators representing the **frequencies of engagement** with the course learning resources: a) frequency of access to lectures, links to lectures, or course materials (e.g., videos), b) frequency of forum discussion contribution (posting, updating, or rating), c) frequency of forum reading, and d) frequency of course browsing and thus getting course updates, announcements, and links to resources; II) indicators of **activity level and regularity** of activities: a) number of active days where the student had at least made one click on learning related activities, and b) regularity of course access calculated according to [33]; III) general indicators representing the total **time and effort** of online activities: a) number of unique sessions—a session was considered as a an uninterrupted sequence of online actions (see [34] for details of the

computation method)—, and b) the total time spent online accessing learning related activities. Since the data were collected from several courses, the indicators were discretized into equal width bins, so all indicators have a similar measurement scale, are easy to compare, and the outliers are neutralized [35].

3.3 Analysis

To classify students according to their engagement states at the course-level (**RQ1**), Latent Cluster Analysis (LCA) was performed using the LMS indicators (frequencies, activities and time indicators) [36]. LCA offers a robust method for clustering of educational data (see [9] for a review and advantages). To select the optimum number of clusters we relied on best Akaike information criterion (AIC), Bayes information criterion (BIC), as well as the separation of clusters with best effect size [9, 36, 37]. Students were clustered according to their engagement level in each course offering (we refer to clusters as *engagement states*). The resulting clusters were compared with Kruskal–Wallis non-parametric one-way analysis of variance (ANOVA) [38]. To test the magnitude of the obtained results and the quality of separation of clusters, we calculated the epsilon-squared effect size [37]. Post-hoc pairwise comparisons were also performed through Dunn’s test using Holm’s correction for multiple testing [39].

To answer **RQ2**, SPM was performed to study and visualize the longitudinal trajectory of engagement states throughout the whole program, patterns of transitions, stability of engagement states and the characteristics thereof. The *TramineR* R package [40] was used to construct a state sequence object from the chronologically ordered course engagement states. Sequence distribution plots were used to demonstrate the ratio of each course engagement state at each time point. Sequence index plots were used to visualize the longitudinal timeline of engagement states for each single student. Mean time plots were used to demonstrate the number of courses each student spent in each engagement state on average. To test how homogenous the sub-groups of engagement are, intra-students stability was measured using entropy [41]. Chi-squared test with Bonferroni correction was computed to compare the patterns of transitions of engagement states in each subgroup.

To find if there are subgroups or pattern within the groups, we implemented clustering, which a standard sequence mining technique. Each group (the early engaged and early disengaged) was clustered using agglomerative hierarchical clustering based on Ward’s method using distances based on optimal matching (see [42] for details of the method).

To answer **RQ3**, a Kaplan-Meier (KM) survival curve was performed to estimate the probability of students persisting in the program (survival). For each group, we report the number of events (of dropouts) and the survival probability at each time point. To compare the difference between trajectories, we performed the recommended tests: Log-rank, Gehan, Tarone-Ware and Peto-Peto [43]. Lastly, a non-parametric Wilcoxon-Mann-Whitney (WMW) was used to compare students Grade Point Average (GPA) at the end of the program which is the total grades of all courses [44].

4 Results

The extracted data included 1,052,807 recorded logs, which became 790,956 log records after removal of non-learning events (e.g., clicks on profile pages). The median number of students per course was 92 and ranged from 84 to 99. The median number of events per course offering (in the final data set after removal of non-learning events) was 15,345 and ranged from 3,925 to 36,417. The median number of forum consumption events was 48.5 per student per course offering; the median of forum contribution per student per course offering was 50; the median of sessions per student per course offering was 48, and the median duration of online time was 4.99 hours per student per course offering.

Table 1
Comparison (ANOVA) of the three clusters according to their mean activity indicators

Indicator	State	Mean	SD	p	ϵ^2
Frequency of course browsing	Actively engaged	8.47	1.487	< .001	0.686
	Averagely engaged	4.91	1.820		
	Disengaged	1.95	1.211		
Frequency of lecture access	Actively engaged	7.31	2.413	< .001	0.304
	Averagely engaged	5.17	2.506		
	Disengaged	2.94	2.057		
Frequency of forum reading	Actively engaged	7.95	2.074	< .001	0.456
	Averagely engaged	4.89	2.218		
	Disengaged	2.67	1.901		
Frequency of forum contribution	Actively engaged	7.25	2.421	< .001	0.277
	Averagely engaged	5.13	2.509		
	Disengaged	3.09	2.298		
Active days	Actively engaged	8.51	1.432	< .001	0.726
	Averagely engaged	4.95	1.738		
	Disengaged	1.79	0.938		
Session count	Actively engaged	8.71	1.095	< .001	0.796
	Averagely engaged	4.87	1.553		
	Disengaged	1.70	0.851		
Total online time	Actively engaged	8.02	1.995	< .001	0.500
	Averagely engaged	4.94	2.187		
	Disengaged	2.45	1.663		
Regularity	Actively engaged	8.15	1.832	< .001	0.609
	Averagely engaged	5.10	2.034		
	Disengaged	1.96	1.032		

To answer **RQ1**, an LCA model was fitted using students' learning activities. We tested models with two to ten clusters (similar to [9]). The model with three clusters had the lowest AIC = 43,748.7 and BIC = 44,891

as well as had clearly separable clusters and high epsilon-squared effect size (ranging from 0.3 to 0.8). Table 1 shows the three identified clusters, which can be described as follows:

- **Actively engaged cluster:** Students in this cluster had the highest values of activity indicators (between the 7th and 9th decile), frequent access to course resources, frequent forum posting and reading. They had the highest frequency of active days, longer sessions, and the highest regularity. These indicators ranged between the 7th and 9th decile.
- **Averagely engaged cluster:** Students in this cluster had moderate (mostly around the 5th decile) values of activity indicators: average frequency of access to course resources, forum posting, forum reading, active days and regularity.
- **Disengaged cluster:** Students in this cluster had the lowest levels of activities that lied between the 1st decile and the 3rd decile.

Throughout this manuscript we will refer to each cluster as an engagement state as it describes the students' state of engagement in each course.

To answer **RQ2**, the engagement states were used to construct a sequence object in each of the two study groups (*Early engagement* and *Early disengagement*) which was mined used SPM. We then compared the two groups and the characteristics of each of their trajectories.

Early disengagement group: The sequence index plot (Fig. 1.A) shows each trajectory of a student as a sequence of horizontally colored stacked bars: the colors reflect their engagement states. The hierarchical clustering (Fig. 1.A) shows two distinct groups of students: G1, a mostly disengaged group ($n = 9$) who eventually drop out, G2 with fluctuating engagement trajectories which can be further divided into two subgroups: G2a with fluctuating engagement trajectory dominated with disengagement states, and G2b, a relatively stable group dominated with active engagement states. Overall, among the group of *Early disengagement*, the total number of students who dropped-out was two students at the 3rd course, six at the 6th course, nine students at the 9th course, and 13 students by the 15th course. A noticeable observation that students who were able to catch up and engage again in the second course were more likely to maintain such an engaged state (11/13), while students who were disengaged for two successive courses had more dropouts (11/28). The distribution plot (Fig. 1.B) shows the distribution of engagement states at each time point, highlighting that dropout occurred immediately after the first term. The mean time plot (Fig. 1.C) shows that these students spent an average of 7.7 courses as disengaged or dropout and 7.3 as averagely or actively engaged.

Early engaged group: The index plot (Fig. 2.A) shows that the students in this group had more stable trajectories that were dominated with average and active engagement states with very infrequent disengagement states (an average of 13.6 courses were active or average engagement per student, Fig.

2.C). Two distinct subgroups can be revealed with hierarchical clustering: G1, a subgroup with mostly average engagement with frequent transition to engagement states, and G2, a subgroup with mostly engaged states with infrequent transition to other engagement states. The distribution plot (Fig. 2.B) confirms that students in this group were mostly highly engaged or averagely engaged. Only a single student dropped at the course 12.

The dynamics of engagement trajectories

Comparing both trajectories helps understand the dynamics of events (transitions, sequences of transitions, stability and persistence. Firstly, we compare the transition probabilities between engagement states to investigate how each group changes across time. Secondly, we compare the frequent transition subsequences that are characteristic of each group. Thirdly, we compare the stability of trajectories in each group using transversal-entropy curves.

Students in the *Early disengagement* group were more likely to descend from the “actively engaged” state to the “disengaged” state (transition probability 28%) compared to the *Early engagement* group (transition probability 11%), which highlights the vulnerability of the former group of students to transition to disengagement states in the future (Table 2). Similarly, students in the *Early disengagement* group were also more likely to stay in a “disengaged” state with a probability of 59% compared to 39% in the *Early engagement* group. The top statistically significant discriminating subsequences in the *Early disengagement* group were characterized by persisting in a “disengaged” state, descending from an engagement state (averagely or actively engaged) to a “disengaged state”, or ascending to an engagement state.

Table 2

Transition probability matrix among engagement states for each engagement group.

<i>Early disengagement</i>	Actively engaged	Averagely engaged	Disengaged	Dropout
Actively engaged >	0.60	0.09	0.28	0.03
Averagely engaged >	0.30	0.55	0.11	0.04
Disengaged >	0.35	0.04	0.59	0.02
Dropout >	0.00	0.00	0.00	1.00
Early engagement	Actively engaged	Averagely engaged	Disengaged	Dropout
Actively engaged >	0.65	0.24	0.11	0.00
Averagely engaged >	0.30	0.69	0.02	0.00
Disengaged >	0.45	0.16	0.39	0.00
Dropout >	0.00	0.00	0.00	1.00

Figure 3, shows the Chi-squared test for discriminating subsequences. The *Early disengagement* group had more frequent and statistically significant (Disengaged) and (Disengaged)-(Disengaged > Actively engaged) subsequences, while less likely to have (Averagely engaged > Actively engaged) subsequences. The transversal-entropy plot show how stable the groups are (Fig. 4). The *Early disengagement* group had higher entropy values at each time point highlighting the instability of *Early disengagement* group. In summary, the *Early disengagement* group were more likely to persist in a “disengaged” state, descend from an engagement state to a disengagement state as well as showed an unstable trajectory. These findings add to the previous findings, that not only being in a disengagement state is an alarming distress signal, but also the persistence or transition to a disengaged state is similarly alarming.

To investigate the likelihood to persist in the program, we estimate the survival probability (probability of completing the program) using KM curves, comparing *Early engagement* and *Early disengagement* groups (Fig. 5.A). The survival probability of the *Early disengagement* group at the end of the 1st year was 0.86 CI (0.76:0.97), at the end of 2nd year the survival probability dropped to 0.81 CI (0.7:0.94), while it was 1.00 in the *Early engagement* group. By the end of the program, the survival probability in the *Early disengagement* group was 0.7 CI (0.57:0.85), while it was 0.98 CI (0.95:1.0) in the *Early engagement* group. The Log-rank, Gehan, Tarone-Ware and Peto-Peto tests were all statistically significant at the level of $p < 0.001$ emphasizing the difference between the groups. In summary, the *Early disengagement* group had a higher and statistically significant probability of dropping out of the program.

Lastly, the results of the WMW test revealed that the performance (measured as GPA) students in the *Early engagement* group (85.15/100) was significantly higher than that of students in the *Early*

disengaged group (79.73/100) with a medium effect size (rank-biserial correlation coefficient of -0.38). These findings indicate that early engagement does not only predict persisting in the program, but it is also a catalyst of higher performance.

5 Discussion

There are several studies that have addressed students' online engagement. The majority of such studies have focused on engagement with a learning task or within a course [8, 9, 26]. Notwithstanding the wealth of insights that we have learned over the years, the longitudinal engagement remained a much-needed area of research. Our study has sought to address such a gap and, in particular, we investigated the longitudinal trajectories of early engaged versus early disengaged students.

Our **RQ1** sought to identify the course-level engagement states of the students over a full program using LCA clustering method. While few studies have looked into a full program, the three engagement states at the course level identified in our study were in line with the existing literature [9, 24, 45]. Our three identified engagement clusters were: 1) an *actively engaged* cluster similar to the intense, highly active groups reported by others [9, 24, 45]. The *actively engaged* cluster invests a significant amount of energy and effort in learning which has been linked to *deep* learning [5, 16]. 2) An *average engagement cluster*, constituted by selective learners investing a moderate amount of effort. Such cluster is similar to *get-it-done cluster* reported in other studies [9, 24]. 3) A *disengaged cluster*, who invest the least amount of effort, focus on assessed resources, or use a *surface* approach to learning [16]. Our clusters may have been more uniform than which can be explained by the fact that our context has a uniform course design, homogeneous course structure and similar underpinnings.

Our **RQ2** investigated the longitudinal trajectories of early disengaged students in comparison to their early engaged counterparts. The results have shown that students in the *Early disengagement* group are not a uniform group but rather a heterogeneous group with two main identifiable subgroups: a group with rapid declining engagement that eventually ended by dropping out early in the program, and another group that maintained a turbulent course with occasional dropping-out. In addition to frequent disengagement states, this group has shown frequent alarming transitions too, that is, transitions from an engagement state to a disengagement state (declined). These findings highlight the importance of course to course transitions as another alarming sign that should alert educators of the imminent risk that these students may have (dropping out). The fact that *Early disengagement* group had a high number of dropouts is an indication of the high risk that early disengagement incurs on students' outcomes. The presence of alarming signs is a sign of tractability that should be taken seriously. Similarly, the *Early engagement* group was not a uniform group of learners, but rather presented two subgroups with two engagement patterns.

The findings of this study align with face-to-face studies [7, 22] that reported the presence of engagement heterogeneity among students, as well as the presence of subgroups of students with different trajectories. A stable subgroup who remain engaged most of the program has been reported by [7, 22],

those students were referred to as “Stable High Trajectory” or “Persistent engagement group” and were around 70% of all enrolled students [7, 22]. A declining trajectory has also been reported by studies in face-to-face settings, referred to as “Descending engagement group” and were around 17% in [7] and around 8% in [22]. While studies online engagement are scarce, the results by [9], who reported the “consistent” and “get-it-done” groups to be around 67% of students if combined and a *disorganized* group (33%) who are mostly disengagement.

RQ3 investigated the relationship between early disengagement and graduation using survival analysis. The results showed that early disengagement projects a substantial risk on students’ graduation and is linked to frequent and statistically significant dropout. Since troubled, unstable or early disengagement have been strongly linked to dropping out of school, the early identification and support of such students is thus of paramount importance to educators [3, 5, 12].

This study contributes to the body of the literature regarding longitudinal engagement, an area that is rarely studied. We show that early online engagement –regardless of the grades– is a significant predictor of persistence in the program. More importantly, we describe the trajectories of early disengaged students and show that in addition to the state of disengagement, the transitions and their stability can also be alarming. Another contribution of this study is a methodological one: we have used SPM methods in modeling the longitudinal engagement, a use that is novel in LA and education.

The generalizability of this study is subject to investigation and replication within different contexts. The quality of data captured from the LMS are far from perfect and students’ usage may differ. Therefore, further confirmation with different samples is required.

6 Conclusions

This study followed 99 students over four years of college education including all their course data (15 courses and 1383 course enrollments). The results have shown that early disengagement is associated with significant higher rates of dropouts, lower scores and lower graduation rates. Our findings indicate that it is critical to proactively watch the alarming signs of disengagement during the first program courses, such as low engagement with the course materials, declining engagement along the program or history of frequent disengagements. Most of existing approaches to disengagement focus on a single course; such approaches are far from optimal as they miss an important aspect of engagement i.e., program-level engagement. This study points to the dire consequences of ignoring program-level disengagement and the loss it incurs in terms of attrition and loss of resources. Therefore, we call for institutions to have program-level monitoring mechanisms that early identify at-risk students and offer a timely relevant support.

Declarations

The authors have nothing to declare

References

1. Newmann FM. Student engagement and achievement in American secondary schools. ERIC; 1992.
2. Fredricks JA, Blumenfeld PC, Paris AH. School engagement: Potential of the concept, state of the evidence. *Rev Educ Res.* 2004;74:59–109. doi:10.3102/00346543074001059.
3. Finn JD, Zimmer KS. Student Engagement: What Is It? Why Does It Matter? In: *Handbook of Research on Student Engagement*. Boston, MA: Springer US; 2012. p. 97–131. doi:10.1007/978-1-4614-2018-7_5.
4. Lei H, Cui Y, Zhou W. Relationships between student engagement and academic achievement: A meta-analysis. *Soc Behav Pers.* 2018;46:517–28.
5. Azevedo R. Defining and Measuring Engagement and Learning in Science: Conceptual, Theoretical, Methodological, and Analytical Issues. *Educ Psychol.* 2015;50:84–94.
6. Skinner EA, Pitzer JR. Developmental Dynamics of Student Engagement, Coping, and Everyday Resilience. In: Christenson SL, Reschly AL, Wylie C, editors. *Handbook of Research on Student Engagement*. Boston, MA: Springer US; 2012. p. 21–44. doi:10.1007/978-1-4614-2018-7_2.
7. Zhen R, Liu R De, Wang M Te, Ding Y, Jiang R, Fu X, et al. Trajectory patterns of academic engagement among elementary school students: The implicit theory of intelligence and academic self-efficacy matters. *Br J Educ Psychol.* 2020;90:618–34.
8. Henrie CR, Halverson LR, Graham CR. Measuring student engagement in technology-mediated learning: A review. *Comput Educ.* 2015;90:36–53. doi:10.1016/j.compedu.2015.09.005.
9. Barthakur A, Kovanovic V, Joksimovic S, Siemens G, Richey M, Dawson S. Assessing program-level learning strategies in MOOCs. *Comput Human Behav.* 2021;117 January 2020:106674. doi:10.1016/j.chb.2020.106674.
10. Lust G, Elen J, Clarebout G. Regulation of tool-use within a blended course: Student differences and performance effects. *Comput Educ.* 2013;60:385–95. doi:https://doi.org/10.1016/j.compedu.2012.09.001.
11. Saqr M, Wasson B. COVID-19: Lost opportunities and lessons for the future. *Int J Health Sci (Qassim).* 2020;14:4–6. <http://www.ncbi.nlm.nih.gov/pubmed/7269621>.
12. Fredricks JA, Blumenfeld PC, Paris AH. School engagement: Potential of the concept, state of the evidence. *Rev Educ Res.* 2004;74:59–109.
13. Wang M Te, Degol J. Staying engaged: Knowledge and research needs in student engagement. *Child Dev Perspect.* 2014;8:137–43.
14. King RB. Sense of relatedness boosts engagement, achievement, and well-being: A latent growth model study. *Contemp Educ Psychol.* 2015;42:26–38.
15. Fredricks JA, McColskey W. The Measurement of Student Engagement: A Comparative Analysis of Various Methods and Student Self-report Instruments. In: Christenson SL, Reschly AL, Wylie C, editors. *Handbook of Research on Student Engagement*. Boston, MA: Springer US; 2012. p. 763–82. doi:10.1007/978-1-4614-2018-7_37.

16. Redmond P, Abawi LA, Brown A, Henderson R, Heffernan A. An online engagement framework for higher education. *Online Learn J*. 2018;22:183–204.
17. Balfanz R, Herzog L, Mac Iver DJ. Preventing student disengagement and keeping students on the graduation path in urban middle-grades schools: Early identification and effective interventions. *Educ Psychol*. 2007;42:223–35.
18. López-Pernas S, Saqr M, Viberg O. Putting It All Together: Combining Learning Analytics Methods and Data Sources to Understand Students' Approaches to Learning Programming. *Sustainability*. 2021;13:4825–43. doi:10.3390/su13094825.
19. Saqr M, Fors U, Tedre M. How learning analytics can early predict under-achieving students in a blended medical education course. *Med Teach*. 2017;39:757–67.
20. Wigfield A, Eccles JS, Schiefele U, Roeser RW, Davis-Kean P. Development of achievement motivation. *Handb child Psychol*. 2007;3.
21. You S, Sharkey J. Testing a developmental-ecological model of student engagement: A multilevel latent growth curve analysis. *Educ Psychol*. 2009;29:659–84.
22. Archambault I, Dupéré V. Joint trajectories of behavioral, affective, and cognitive engagement in elementary school. *J Educ Res*. 2017;110:188–98. doi:10.1080/00220671.2015.1060931.
23. Nouri J, Larsson K, Saqr M. Bachelor Thesis Analytics: Using Machine Learning to Predict Dropout and Identify Performance Factors. *Int J Learn Anal Artif Intell Educ*. 2019;1:116. doi:10.3991/ijai.v1i1.11065.
24. Mirriahi N, Jovanovic J, Dawson S, Gašević D, Pardo A. Identifying Engagement Patterns with Video Annotation Activities: A Case Study in Professional Development. *Australasian Journal of Educational Technology*. 2018;34:57–72.
25. Palvia S, Aeron P, Gupta P, Mahapatra D, Parida R, Rosner R, et al. Online Education: Worldwide Status, Challenges, Trends, and Implications. *J Glob Inf Technol Manag*. 2018;21:233–41. doi:10.1080/1097198X.2018.1542262.
26. Joksimović S, Poquet O, Kovanović V, Dowell N, Mills C, Gašević D, et al. How do we model learning at scale? A systematic review of research on moocs. 2018.
27. Vytasek JM, Patzak A, Winne PH. Analytics for student engagement. *Intelligent Systems Reference Library*. 2020;158:23–48. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85063756448&doi=10.1007%2F978-3-030-13743-4_3&partnerID=40&md5=f5afcc9de2870afd690d75d0b4740269.
28. Saqr M. Learning analytics and medical education. *Int J Health Sci (Qassim)*. 2015;9.
29. Uzir NAAA, Gašević D, Jovanović J, Matcha W, Lim L-AA, Fudge A, et al. Analytics of time management and learning strategies for effective online learning in blended environments. In: *ACM International Conference Proceeding Series*. New York, NY, USA: Association for Computing Machinery; 2020. p. 392–401. doi:10.1145/3375462.3375493.
30. Agrawal R, Srikant R. Mining sequential patterns. In: *Proceedings of the eleventh international conference on data engineering*. IEEE; 1995. p. 3–14.

31. Ritschard G, Studer M. Sequence Analysis: Where Are We, Where Are We Going? 2018. p. 1–11. doi:10.1007/978-3-319-95420-2_1.
32. Saqr M, Nouri J, Jormanainen I. A Learning Analytics Study of the Effect of Group Size on Social Dynamics and Performance in Online Collaborative Learning. In: Scheffel M, Broisin J, Pammer-Schindler V, Ioannou A, Schneider J, editors. Lecture Notes in Computer Science. Cham: Springer, Cham; 2019. p. 466–79. doi:10.1007/978-3-030-29736-7_35.
33. Jovanović J, Mirriahi N, Gašević D, Dawson S, Pardo A. Predictive power of regularity of pre-class activities in a flipped classroom. Comput Educ. 2019. doi:10.1016/j.compedu.2019.02.011.
34. Jovanović J, Mirriahi N, Gašević D, Dawson S, Pardo A. Predictive power of regularity of pre-class activities in a flipped classroom. Comput Educ. 2019.
35. Hung J-L, Shelton BE, Yang J, Du X. Improving Predictive Modeling for At-Risk Student Identification: A Multistage Approach. IEEE Transactions on Learning Technologies. 12:148–57.
36. Goodman LA. Exploratory latent structure analysis using both identifiable and unidentifiable models. Biometrika. 1974;61:215–31.
37. Tomczak M, Tomczak E. The need to report effect size estimates revisited. An overview of some recommended measures of effect size. Trends Sport Sci. 2014;21:19–25.
38. Ostertagova E, Ostertag O, Kováč J. Methodology and application of the Kruskal-Wallis test. In: Applied Mechanics and Materials. Trans Tech Publ; 2014. p. 115–20.
39. Holm S. A Simple Sequentially Rejective Multiple Test Procedure. Scand J Stat. 1979;6:65–70.
40. Gabadinho A, Ritschard G, Müller NS, Studer M. Analyzing and Visualizing State Sequences in R with TraMineR. J Stat Softw. 2011;40.
41. Gabadinho A, Ritschard G, Studer M, Nicolas SM. Mining sequence data in R with the TraMineR package: A users guide for version 1.2. Geneva Univ Geneva. 2009;1.
42. Fincham E, Gašević D, Jovanović J, Pardo A. From Study Tactics to Learning Strategies: An Analytical Method for Extracting Interpretable Representations. IEEE Trans Learn Technol. 2019;12:59–72.
43. Blossfeld HP, Golsch K, Rohwer G. Event history analysis with stata. Routledge; 2007.
44. McKnight PE, Najab J. Mann-Whitney U Test. Corsini Encycl Psychol. 2010;:1.
45. Kovanović V, Joksimović S, Poquet O, Hennis T, de Vries P, Hatala M, et al. Examining communities of inquiry in Massive Open Online Courses: The role of study strategies. Internet High Educ. 2019;40:20–43. doi:10.1016/j.iheduc.2018.09.001.

Figures

Figure 1

The sequence plots of the Early disengagement group: A) index plot, B) distribution plot, and C) mean time plot.

Figure 2

The sequence plots of the Early engagement group: A) index plot, B) distribution plot, and C) mean time plot.

Figure 3

The top statistically significant discriminating subsequences in the Early disengagement group

Figure 4

The transversal-entropy curves of both groups, lower values indicate a stable trajectory.

Figure 5

A) KM survival curves comparing both groups. B) WMW test comparing the GPA of both groups