Analyzing the consistency in within-activity learning patterns in blended learning

Varshita Sher
Simon Fraser University
vsher@sfu.ca

Marek Hatala
Simon Fraser University
mhatala@sfu.ca

Dragan Gašević
Monash University
dgasevic@acm.org

ABSTRACT

Performance and consistency play a large role in learning. This study analyzes the relation between consistency in students' online work habits and academic performance in a blended course. We utilize the data from logs recorded by a learning management system (LMS) in two information technology courses. The two courses required the completion of monthly asynchronous online discussion tasks and weekly assignments, respectively. We measure consistency by using Data Time Warping (DTW) distance for two successive tasks (assignments or discussions), as an appropriate measure to assess similarity of time series, over 11-day timeline starting 10 days before and up to the submission deadline. We found meaningful clusters of students exhibiting similar behavior and we use these to identify three distinct consistency patterns: *highly consistent*, *incrementally consistent*, and *inconsistent users*. We also found evidence of significant associations between these patterns and learner's academic performance.

CCS CONCEPTS

• **Applied computing → Education**; • **Mathematics of computing → Time series analysis**; • **Information systems → Data mining**; • **Human-centered computing** → Mobile devices;

KEYWORDS

Work Habits, Time Management, Time-series Analysis, Learner performance and Consistency, Regularity, Student Persistence

ACM Reference Format:
Varshita Sher, Marek Hatala, and Dragan Gašević. 2020. Analyzing the consistency in within-activity learning patterns in blended learning. In *Proceedings of the 10th International Conference on Learning Analytics and Knowledge (LAK '20), March 23–27, 2020, Frankfurt, Germany*. ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3375462.3375470

1 INTRODUCTION

An interesting topic of ongoing research in higher education context has been how different learning approaches relate to academic achievement [5]. These learning approaches are generally part of the cyclic processes involved in self-regulated learning (SRL) (i.e. planning a task, monitoring the performance and reflecting on the outcomes and on the learning process) [43] and educational psychologists affirmed that these processes are a key contributor to the academic success of students. Of the many self-initiated actions involved in SRL such as goal setting, self-monitoring, metacognition, physical and social environment management, and effort regulation, time management is known to be a strong predictor of student grades [5, 28]. Time management as a key self-regulatory skill involves scheduling, planning, and managing one's study time, to allocate efforts depending on intensity of work [27].

Although less frequently highlighted, the *consistency* of our study habits that controls regulation of effort, setting specific work load for the week, and behavioral adjustments is also a key dimension of time management. Despite research suggesting time management and effort regulation, i.e. perseverance, positively predict academic grades significantly [4, 9], the analysis of student's work-pattern changes across individual activities has so far only been sparsely studied, in the context of blended and technology-enhanced learning. Our aim in this study is to observe how stable these patterns of work habits (or procrastination, as an extreme) are when students are given the opportunity to acknowledge differences that may arise from considerable variation in successions of the same type of learning activity in a blended learning environment.

Students' learning patterns are dynamically changing entities. Unlike students' demographics or their prior academic record, learning patterns reflect students' current unique engagement levels and learning processes [13]. Analyzing the within-activity variance in online learning-patterns for a student allows us to challenge most 'traditional aggregated evaluation and analysis methods' [19] (say, prediction models), which utilize data aggregated across the entire semester. As a result of the aggregation, these methods fails to consider the variances in course-activity patterns. For instance, a student's aggregated time spent per week in two different weekly assignments might be identical. However, during week 1 the student might have evenly allocated their learning time in the days leading up to the deadline; whereas in week 2, activities might be concentrated in the last two days before the deadline. Although considered equivalent in total performance efforts, the student's varying patterns might be indicative of success or failure at a finer level. While there could be plenty of reasons (which are outside the scope of this paper) for the observed inconsistencies in behaviors, such as active procrastination [41] or excess workload from other courses, it is nonetheless worthwhile to assess if, and when, the course-activity patterns start to deviate from patterns known to be *favorable* for academic success [18].

Our study is motivated by existing research on engagement, which suggests that academic success is highly likely in case of students adopting habit-inducing behavior [11, 25]. Analysis of

LAK '20, March 23–27, 2020, Frankfurt, Germany

ACM ISBN 978-1-4503-7712-6/20/03…$15.00
https://doi.org/10.1145/3375462.3375470

consistency can allow us to further understand if the student behavior remains constant throughout the semester, visible only in the beginning of a course or converges as one progresses in their course. In this paper, consistency is analyzed from two viewpoints – (a) through periodic participation in the discussion forums, mainly reviewing the course-related discussion topics, and (b) recurring engagement with the assignment tasks. These two learning activities were chosen because of the course design that made use of these two main activities in the online setting. Further, since it has been already established that learners make sequential and simultaneous use of various technological modalities such as desktops, mobiles and tablets for learning activities [23, 33] and that these have potential for an impact on their academic achievements [32], we posit that the preferences for a modality may also evolve over time. Therefore, we include modalities in our analysis of students' activity changes across assignments and discussions when the students are given the opportunity to use multiple devices for participating in a learning activity.

In particular, this paper answers the following research questions:

- **RQ1**: How consistent are students' work patterns across subsequent activities of the same type, when engaging with them from multiple modalities? That is, can we identify conceptually and practically meaningful clusters of students with distinct consistency patterns?
- **RQ2**: Is there an association of the identified patterns with students' academic performance?

1.1 Time series analyses of work patterns

In recent years, several implementations of time series analysis have been reported in the field of education and learning analytics. Time-series clustering was applied by Młynarska et al. [24] to identify distinct activity patterns among students, in order to tackle the issue of difficulty in keeping up with deadlines. Based on high activity levels within a three-week timeline, they identified seven groups of students - Procrastinators, Strugglers, Unmotivated, Steady, Hard-workers, Strategists and Experts. Hung et al. [19] demonstrated that time series models (time series data points aggregated daily) were better than traditional data aggregation models (frequencies over the whole semester) at identifying at-risk online students, both earlier and with greater accuracy (misclassification rate below 10%). Brooks et al. [7] employed features created as *n*-grams (n = 2 to 5) over different time periods, from logs of learner interactions with educational resources. They chose four different granularities of timeframes: accesses within a calendar day, a three-calendar day period, a calendar week, and a calendar month; such that 'an *n*-gram with the pattern (*false*, *true*, *false*), the label of *week*, and count of 2 would indicate that a student had two occurrences of the pattern of not watching lectures in one week, watching in the next week, and then not watching again in the third week' [7]. By detecting similar patterns of interaction that lead to learners achieving a passing grade for a course, their designed models were highly accurate (with a misclassification rate below 5%) and generalizable to new real-world dataset with highly accurate results by third week of the course.

Some empirical investigations have also been carried out to elucidate the theoretical mechanisms that link certain activity patterns (extracted from student time-series) to academic success. For instance, Hung et al. [19] exemplified successful learning patterns as *stable and consistent* engagement levels on all basic learning behaviors, and at-risk patterns as unstable engagement levels with high peaks and gaps during the semester. According to Młynarska et al. [24], the most common patterns for students achieving high grades were regular, relatively higher spikes in activity levels or low-level frequent activity with no high spikes around deadline. Unsurprisingly, students achieving low grades exhibited minimal overall effort but larger activity levels closer to the deadline. Hensley et al. [18] identified six time-use patterns from weekly time logs with late-start studying and Sunday cramming indicative of ineffective time use and consistent weekday studying, Saturday studying, consistent bedtime, and consistent wake time indicative of effective time use.

Altogether, the findings suggest the academic relevance of how students manage their learning time. More nuance in future research is necessary, particularly through studies that address consistency of time use patterns thereby providing a detailed view of how students routinely engage with a learning task – whether they are piecing together a routine or simply engaging in a one-off task.

1.2 Consistency in Learning Behaviours

How stable are learning patterns? A systematic review of the seminal works on learning consistency by Vermunt and Vermetten [39] revealed that these studies were conducted from a longitudinal perspective. That is, questionnaires were typically administered at a suitable gap of time to explore students learning patterns in a pre-post test design. For instance, Svensson [34] studied the ways students process learning material using three measurement points over a period of five weeks and found that ways of processing the learning material were rather stable across the three occasions. In a series of studies by another group of researchers, similar questionnaires were administered twice to the same group of students at a gap of 3 months [38] and 6 months [37]. Overall, the results indicated high stability of learning strategies, learning orientations and conceptions of learning on the two occasions.

The aspect of consistency has also been analyzed, albeit briefly, under different contextual conditions and at varying levels of granularity. In the study by Thomas and Bain [35], a 7-item questionnaire was administered to determine whether the students' learning strategies (deep vs. surface) were consistent in tests vs. essays comparison. High level of consistency in the strategies was found; also high levels of achievement on both tests and essays were associated with use of deep strategies.

At course-specific levels, Vermetten et al. [36] assessed strategy use by the same group of students in four different courses. They found not only that students vary their learning strategies for different courses, but also that the learning strategies differed from each other in their degree of variability across courses.

As evidenced by the aforementioned studies, existing research on learning consistency in students has been investigated mainly using questionnaires, measuring consistency in the way users respond about their learning strategies. A large majority of studies have made use of the Inventory of Learning Styles (ILS) questionnaire to measure consistency and variability in students' use of learning strategies (for an in-depth review, see Vermunt and Vermetten [39]). More recently, there have been attempts at empirically investigating the aspects of consistency in everyday learning. Jo et al. [20] studied the impact of login consistency on academic performance and found significant associations between (ir)regularity of learning interval in LMS and final grades, where regularity was calculated using the standard deviation of the login intervals (i.e. average login time into the LMS). Thus, a higher value indicated highly irregular logins.

Similarly, Dvorak and Jia [12] studied the relationship between consistency of time of study and found regular work on assignments to be associated with high grades in course work. They defined regularity as the degree to which the student tends to work at the same time of the day, i.e. whether he or she would start each assignment at the same time of day throughout the term, and operationalized it as the inverse of the standard deviation of the hours before the assignment deadline. Młynarska et al. [24] studied consistency at the activity-specific level by comparing the time series signatures of activity patterns between successive assignments. They found activity patterns were more similar for the same student than those for different students, and moreover, students who changed their behaviour from one assignment to another, exhibited a change in grade too, i.e. the two were positively correlated.

To the best of our knowledge, only Młynarska et al. [24] have attempted to evaluate consistency in learning patterns at an activity-specific level and as such, the topic warrants ongoing attention from the learning analytics community. Additionally, we noticed a distinct lack of studies that assess how the use of multiple devices is associated with consistency of activity and success. Analysis of the use of multiple devices is necessary in this digital era since learners are making sequential and simultaneous use of a combination of devices (like mobiles and tablets) to support formal learning [23]. Further, most previous studies have investigated either variation or consistency in learning processes by looking at variability and consistency of *self-reported* strategies at the same time. In the present paper, however, a different position was taken by looking at variability and consistency in work patterns using log data, which does not suffer from the shortcomings of survey or questionnaire data [40, 42] and thus reflects actual students' behaviours.

2 METHODS

2.1 Study Context

The data analyzed in this study was gathered from the second and third year undergraduate students in two subsequent offerings (2017 and 2018) of two information technology courses (C1 - Multimedia Programming and C2 - Internet Computing Technologies) at a Canadian university. Both courses were similar in structure, having a 2-hour face-to-face lecture per week, a 2-hour in-lab tutorial per week. Tutorial participation contributed 10% towards the final grade, assignments 40% of the grade, quizzes and exams 50% in course C1 and 35% in course C2, and course C2 had three online discussions 5% each for a total of 15%. C1 was used for collecting data related to assignment activity whereas C2 provided the discussion activity data. The activity topics and grading structure for both assignment and discussion activities remained constant over the two years (2017 and 2018) and were taught by same instructors too. Both courses used blended delivery, utilizing the university's learning management system (LMS) to support learning activities and students' overall schoolwork. The students were experienced in using the LMS as they used it on a day-to-day basis in prior courses. In addition to the web-browser versions of the LMS (desktop/laptop/mobile), students had access to the mobile app version provided by the LMS vendor. There was no apparent difference between the features and functionalities offered by the two versions. Log data from the LMS was the main source of data for analyses.

Prior to the analyses, student records were anonymized and assignment reviewing and discussion participation records were extracted from C1 and C2, respectively. Assignments in C1, eight in total, were all individual, comprising of programming tasks of increasing complexity, and developed in the programming environment outside of the LMS. The assignments required students to apply the concepts learnt during lectures and tutorials onto new problems. The assignment specifications were posted in the LMS; students submitted assignments via the LMS, and received feedback and grades in the LMS. The discussion activities in C2, three in total and unrelated to one another, were 10-14 days long, in small groups of 6-8 students, and required conducting research and developing a group statement to an open ended question posed from topics surrounding programming. Quality of post content, building on ideas of others and quality of the group final statement were marked. A minimum of four posts was required for a student to get the full mark. The grades for discussions were posted in the LMS as well.

2.2 Learning traces and time series

The study used the interaction trace data from students' engagement with the LMS. Students self-regulated their participation in the course activities, guided by the course requirements and deadlines. The use of device modalities was a choice of each student. Each student action in the LMS was logged with the following data: student id, course id, type of learning action, user-agent (used for extracting the type of device used for the action), action URL, session number, start time, and end time.

The log data was transformed into a series of equispaced points in time. In our case, a time series is a 11-day timeline – from 10th day before a deadline until the day of submission. Each bucket in these timeline corresponds to activity counts on the (i-1)th day before the deadline (i = 1:11). The count measures were extracted based on the number of times each learning action was performed by each student (i.e., discussion views in case of discussion activity and assignment views in case of assignment activity). A 10-day limit was chosen because even though each assignment was released at least 14 days in advance, most students did not start working 10 days prior to the deadline. Similar observations were made for discussion activity too. The day of deadline (0th day) was included in the timeline since a majority of students (96%) submitted the assignment on the day of the deadline (of these, 73% submitted less than 6 hours before the deadline), meaning they were working very close to the deadline on their assignment tasks. Further, to account for the simultaneous use of multiple modalities in these activities, we created multi-dimensional time series. Thus, for each student, we generate two time series per assignment or discussion task: T1

= x_{10}, x_9, ... x_0 and T2 = y_{10}, y_9, ... y_0, where x_i is the count of (assignment/discussion) views from desktop on the i-th day before the deadline and y_i is the count of views from mobile on the i-th day before the deadline.

To assess consistency between student's temporal patterns during a learning activity, we first addressed the challenge of appropriately measuring the similarity/distance between pairs of series. Euclidean distance was ruled out since it misses similarity between time series if activity peaks are offset in time, a common occurrence especially since learners work according to their own time availability. Instead, we used dynamic time warping (DTW) measure which has been proposed for quantifying similarity between pairs of temporal sequences [1, 14]. DTW, using stretching or compressing segments of temporal data, determines an optimal match between two time series. That is, two series that exhibit similar peaks (or troughs) are considered similar even if they are slightly displaced in time. The extent of warping allowed can be maintained using global constraints [15] in a way that allows more intuitive warpings. For instance, the series with the peak in actions on 10th day before the deadline will be distinguished from a peak in actions one day before the deadline, since the two represent quite different time scheduling patterns from an SRL perspective. For calculating the DTW measure in our study, we implement the *sakoechiba window* [30] for enforcing a global constraint on the envelope of the warping path with the *window size* set to 2. This size was intuitively and carefully chosen since a very small size makes the warping impossible whereas an unnecessarily large size will introduce impossible mappings (or pathological warping). Finally, the computed DTW distances were normalized for warping path length.

2.3 Data Analysis Techniques

To find recurring patterns in the consistency of work habits, for each student we first calculated the similarity between subsequent activities. That is, for a student participating in three discussion activities, we calculate three corresponding distance measures ($D_{i,j}$), one for each pair of discussion task, such that $D_{i,j}$ is the DTW measure between the bi-variate time series obtained from work habits in discussion i and discussion j. Thus, we obtained $D_{1,2}$, $D_{1,3}$ and $D_{2,3}$ measures for each student participating in discussion activity. For the eight assignment activities, we obtained 28 corresponding measures, one for each pair of assignment tasks.

The distance measures computed for each student were used in the cluster analysis (agglomerative clustering based on Ward's method) to group students (N = 55 for discussion activity and N = 162 for assignment activity). All the DTW measures were normalized prior to the clustering; the Euclidean metric was used to compute the distance between vectors. The optimal number of student clusters was obtained from (a) inspection of the resulting dendrogram, and (b) using the "Silhouette statistic" proposed by Rousseeuw [22, 29] and computed using the clValid R package [6]. The Silhouette value measures the degree of confidence in a particular clustering assignment and lies in the interval [-1,1], with well-clustered observations having values near 1 and poorly clustered observations having values near -1.

Each student cluster was summarized by calculating its centroid, which represented the mean value of all cluster members across all clustering variables. Student cluster assignments (representative of their work pattern consistencies) enabled us to group students and identify whether different consistency patterns relate to differences in overall academic performance (operationalized by discussion grades in discussion activity and assignment grades in assignment activity).

To examine if there were significant differences between the identified student groups, we performed two separate analysis of variance (ANOVA) tests. The student cluster assignment was treated as the single, independent variable in each test, along with the respective dependent variables: final discussion grade and final assignment score.

Before running the ANOVA, we checked the homogeneity of variance using Levene's test. The Shapiro-Wilk test was performed to check for normality. In our case, we found significant Levene's test (i.e., the homogeneity of variance assumption was violated), thus, the non-parametric Kruskal-Wallis test was used. Finally, the measure of epsilon-squared (ϵ^2) were used to report the effect sizes for Kruskal-Wallis tests, and interpretations were done using Cohen's primer [8], the most commonly used primer for effect size interpretation. The significant Kruskal-Wallis tests were followed up by pairwise Wilcox test to calculate pairwise comparisons between group levels with Benjamini-Hochberg (BH) corrections for multiple testing.

3 RESULTS

In Section 3.1 and 3.2, we present the results from the clustering of students based on their consistency in assignment activities and assess the impact of consistency on academic achievement. In Section 3.3 and 3.4, we do the same for the discussion activities.

3.1 Clustering of students based on consistency in assignment activities

The solution with five clusters was found as optimal. The resulting clusters indicate five different patterns of consistency in temporal patterns that students tended to display when engaging with the assignment material whilst working towards a deadline, and self-regulating their studies through the LMS.

Figure 1 presents the box-plots for each of the five consistency clusters. The y-axis represents all possible assignment-assignment pairs (starting from assg1-assg2, assg1-assg3, and so on at the bottom, to assg7-assg8 at the very top) and the x-axis denotes the corresponding DTW measures for each pair. The DTW measures were scaled between [0,1] for cross-cluster comparisons, with values closer to 0 representing almost similar time series and values closer to 1 representing highly dissimilar time series. For all the clusters, the box plots denote the five-number summary - whiskers going from (1) minimum to (2) maximum DTW value, middle box representing middle 50% of DTW scores for the group i.e. left and right box-edge representing (3) Q1 (first quartile) and (4) Q3 (third quartile), respectively, and (5) median DTW measure represented by the vertical line going through the box. As can be observed from Figure 1, except for Cluster 5, students in all other clusters had median DTW measures for all 28 assignment-assignment pairs well below the half of the maximum threshold.

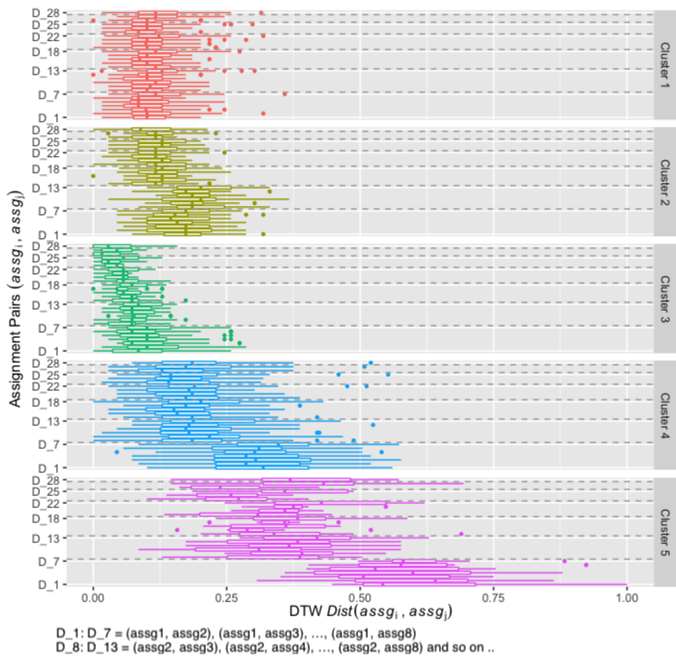

Figure 1: Box plots representing five number summary for the five student clusters. The box plots are color-coded by the student cluster they belong to. The y-axis represents all possible assignment-assignment pairs and the x-axis denotes the corresponding DTW measures for each pair.

From the perspective of the pairwise DTW measures described in Section 2.3, the clusters can be described as follows:

- Student Cluster 1 – *Highly Consistent* (N = 62, 38.27%): This cluster constitutes the largest group of students. This group of students had the least variation in their work patterns in going from one assignment to the other, as exhibited by the low DTW measures.
- Student Cluster 2 – *Delayed Consistent* (N = 25, 15.43%): This group of students' approach in first two assignments was quite different from the six remaining assignments. However, assignment 3 onward their work patterns steadily got more consistent.
- Student Cluster 3 – *Incrementally Consistent* (N = 19, 11.73%): This cluster represents the group of students whose time-series, reflecting engagement with assignment material, became more and more similar as the assignments progressed.
- Student Cluster 4 – *Early Consistent* (N = 48, 29.63%): This cluster is similar to Cluster 2 in that the students' engagement patterns with assignment materials in the very first assignment were less similar to the subsequent assignments but assignment 2 onward, their work patterns steadily became more consistent. However, it never reached the level of consistency of Cluster 2.
- Student Cluster 5 – *Inconsistent Users* (N = 8, 4.94%): This cluster constitutes the smallest group of students. These students exhibited remarkably different temporal work patterns and even though the similarity increased (i.e., DTW measures start getting smaller) as the assignments progressed, they were still relatively large compared to the DTW measures from the previous four clusters.

3.1.1 Analysis of cluster differences based on overall grade. Since we found a high degree of correlation between assignment score and final grades (r = .85), we decided to test any underlying cluster differences on the overall student grades before proceeding to check for differences with respect to assignment grades in particular. In order to do so, we used the ANOVA test due to their robustness to mild violations of normality [16], with cluster assignments as the independent variable and final academic grade as the dependent variable. The analyses of the degree of variation in adopted temporal consistency patterns was found to be significantly associated with the overall academic performance score, with a moderate effect size ($F(4,157) = 5.943$, $p < 0.001$, $\eta^2 = .13$). The pairwise comparison of clusters with respect to the final grade (i.e. *percentage*) revealed that Cluster 3 performed significantly lower than all the other clusters (all $ps < 0.005$), even after adjustments to the p-values using the Tukey HSD procedure. However, the difference between the two highly contrasting groups, i.e. Cluster 1 and 5, was not statistically significant.

3.1.2 Transitions in work patterns at activity-specific level. To inspect whether the transitions in temporal work-pattern were because of switching to different modality (from desktop to mobile, or vice versa), variations in intensity of peaks (higher or lower activity peaks due to procrastination) or a combination of the two, we examined the prototype[1] time-series of the clusters. Since the computation of the optimal prototype poses some challenges, we used DTW Barycenter Averaging (DBA) algorithm [26] to determine the cluster centroids (prototypes). This approach computes an 'average' sequence, called barycenter, such that the sum of squared DTW between the barycenter and the set of considered sequences is minimum. Upon observing the prototypes, we found almost no contribution of varying modalities to the varying consistency patterns within assignment phases. This is because students relied mostly on desktops for all assignments and the use of mobiles for this learning activity was sparse, with under 10% of the class using it at most 2-3 times in the 11-day timeline (in conjunction with desktops) in any given assignment.

Table 1 sheds further light on the shift in work pattern timeseries from one assignment to the other for each of the five clusters described above, with the black trend-line representing the prototype time series. We graphed the number of times assignments were accessed on each day (until the deadline) to demonstrate changing patterns of access over the course. Since, it was observed that the use of mobile modality for this learning activity was sparse, only desktop accesses are plotted. Based on Table 1, we can draw some inferences regarding frequently-occurring patterns which were present across multiple assignments. (Note: Each column in Table 1, representing a cluster, has the same number of time series in each of the 8 corresponding assignments; however, some time-series were composed of all zeroes, i.e. zero-engagement level on any given day

[1] A prototype effectively summarizes the most important characteristics of all series in a given cluster [31].

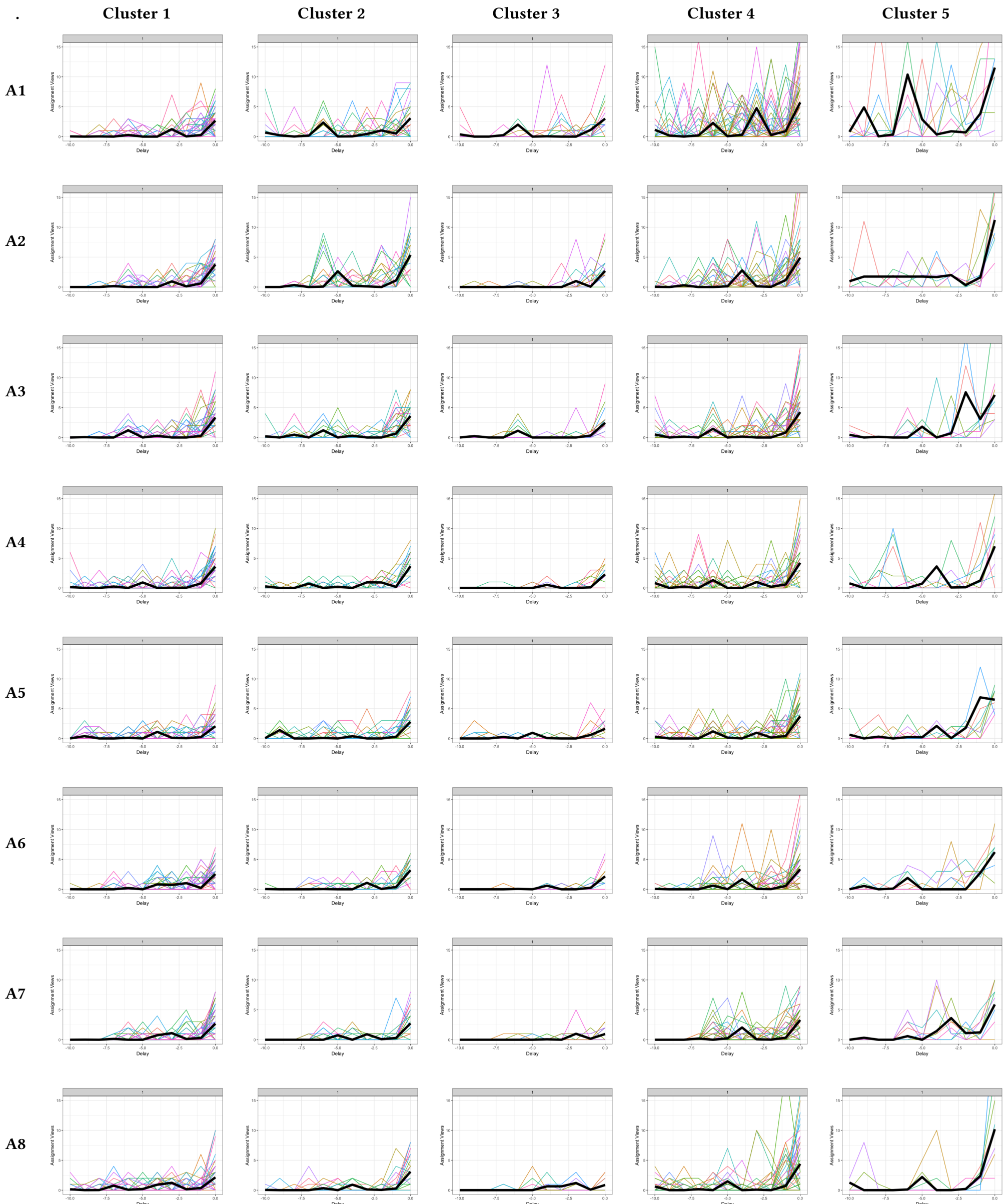

Table 1: Prototype activity patterns for the five student clusters (Cluster1 : Cluster5) at the eight assignment tasks (A1 : A8). The x-axis represents number of days before the assignment, starting from 10th day before the deadline up to the day of submission and the y-axis represents the number of assignment views. To allow cross graph comparison, all graphs have been plotted with x-axis scale [-10,0] and y-axis [0,15].

in timeline, and hence may have been obscured by each other at the bottom of the graph.)

For most assignments, students in cluster 1 were active quite early on (five or six days before the deadline) but their level of engagement with the assignment was low (less than 5 views) and infrequent (at most two peaks in 11 day time-frame). The Cluster 2 students' engagement with the first two assignments differed compared to the later six assignments. While the level of engagement in assignments A1 and A2 was high (5 or more views in a day) and evenly spaced out in the days leading up to the deadline, assignment A3 onward the engagement levels dropped immensely and a high level of activity was witnessed closer to the deadline. The Cluster 3 students were barely active with the assignment activity on the LMS (except for the first two assignments, which were relatively easier in terms of task difficulty). Any engagement with assignment activities was witnessed much closer to the deadline only, thus explaining their incrementally consistent (but rather poor) approach observed. The level of activity revealed in all the assignments in Cluster 4 was higher than that of any other group. Except for assignment A1 where exceptionally large peaks in activity levels (7 or more views in a day) were present throughout the 11-day timeline, the Cluster 4 students were steady in their approach, with preparations starting quite in advance and small peaks in engagement (5 or less views) observed around four-five days before the deadline. The students belonging to the smallest group, Cluster 5, demonstrated unique activity patterns with each assignment. There were some instances (for example assignment A2 and A8) where exceedingly high spikes in activity levels (more than 10 views) were found on the day of submission whereas in other cases (for example assignments A3, A5, and A7) the engagement was regular before the deadline and even higher compared to those found in other groups.

3.2 Analysis of cluster differences based on assignment grade

After examining the differences between clusters based on final grade, we proceeded to further check for the differences between the discovered clusters with respect to their performance in the assignments per se. In total, scores obtained in the eight assignment submissions represented the main data source for analyzing cluster differences.

A non-parametric one-way analysis of variance was conduced with the students' cluster assignment and the final assignment score (average of the eight assignments) as the single independent and dependent variable, respectively. The main effect analyses from the test revealed that the final assignment scores were statistically significantly associated with the learners' consistency profile, with a small-medium effect size ($\chi^2(5) = 17.463$, $p = 0.001$, $\epsilon^2 = .11$). The pairwise comparison of clusters with respect to the assignment grade (i.e. assignment percentage) revealed that Cluster 3 (52.44 ± 32.14) performed significantly lower than all other clusters (all $ps < 0.01$), even after adjustments to the p-values using the Benjamini-Hochberg (BH) procedure. Additionally, performance was significantly better for students in Cluster 5 (92.45 ± 5.86) compared to Cluster 1 (74.72 ± 17.66, $p < 0.01$) and Cluster 4 (75.04 ± 18.67, $p < 0.01$). For completeness, the Cluster 2 percentages were (76.57 ± 21.48).

3.3 Clustering of students based on consistency in discussion activities

The solution with 6 clusters was found as optimal. The resulting clusters indicating the six different patterns of consistency in temporal patterns that students tended to display, when reviewing the discussions in the forum whilst working towards a deadline.

Figure 2 presents the box-plots for each of the six consistency profiles (with the interpretation of the several plot elements same as in Figure 1). As can be observed from Figure 2, except Cluster 2 and 5, the median DTW measures for all three discussion-discussion pairs were well below the half of the maximum threshold. From the perspective of the pairwise DTW measures described in Section 2.3, the clusters can be described as follows:

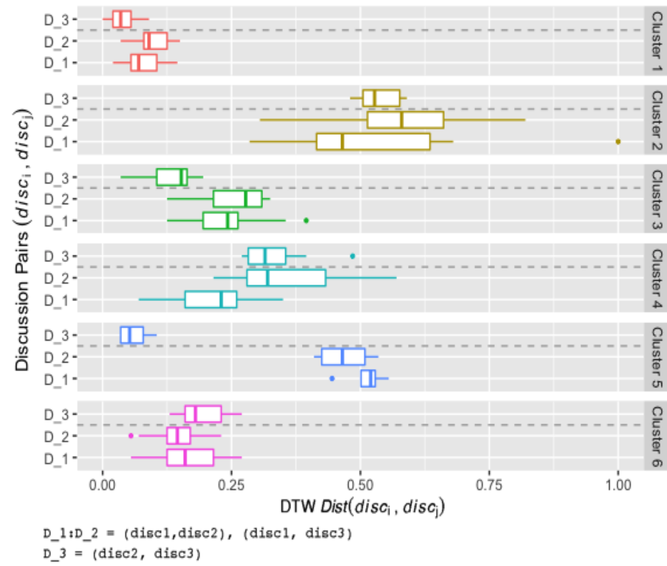

Figure 2: Box plots representing five number summary for the six student clusters. The box plots are color-coded by the student cluster they belong to. The y-axis represents all possible discussion-discussion pairs and the x-axis denotes the corresponding DTW measures for each pair.

- Student Cluster 1 – *Highly Consistent* (N = 9, 16.36%): This group of students had the least variation in their work patterns in going from one discussion to the other, as exhibited by the low DTW measures.
- Student Cluster 2 – *Incrementally Inconsistent (high DTW)* (N = 6, 10.91%): This group of students' patterns became increasingly dissimilar with each successive discussion.
- Student Cluster 3 – *Early Consistent* (N = 12, 21.82%): This cluster constitutes the second largest group of students wherein engagement patterns in the first discussion are less similar to the other two discussions tasks. However, discussion 2 onward, their work patterns became more similar, although it never reached the level of consistency of Cluster 1.
- Student Cluster 4 – *Incrementally Inconsistent (low DTW)* (N = 11, 20%): This cluster is similar to cluster 2 in that the patterns become increasingly dissimilar discussion 2 onward. However, it never reached the level of inconsistency of Cluster 2.
- Student Cluster 5 – *Steep Consistent* (N = 4, 7.27%): In this cluster, students' engagement patterns with discussion posts in the very first discussion were highly dissimilar to those in the subsequent discussions. However, a remarkable level of consistency between timeseries patterns of discussion 2 and 3 was seen (low *D_3* DTW measure) such that Cluster 1's level of consistency was achieved.
- Student Cluster 6 – *Fairly Consistent* (N = 13, 23.64%): This cluster constitutes the largest group of students wherein least variation in their work patterns were observed between subsequent discussion tasks but the overall consistency was slightly lower than that observed in Cluster 1.

3.3.1 Analysis of cluster differences based on overall grade. The analyses of the degree of variation in adopted temporal consistency patterns was found to be significantly associated with the overall academic performance score, with a large effect size ($F(5,49) = 3.381$, $p < 0.01$, $\eta^2 = .32$). The pairwise comparison of clusters with respect to the final grade (i.e. *percentage*) revealed that Cluster 1 performed significantly lower than Clusters 2 and 4 (both $ps < 0.01$) while Cluster 2 performed better than Cluster 3 ($p = 0.04$), even after adjustments to the p-values using the Tukey HSD procedure.

3.3.2 Transitions in work patterns at activity-specific level. Table 2 depicts the shift in work pattern timeseries between discussion tasks for each of the six clusters described above, with the black trend-line representing the prototype time series. Much like the assignment activity, the use of mobile phone modality was sparse for the discussion activity as well and the variations in work patterns are mainly attributed to the change in intensity of engagement levels (from desktops).

For Cluster 1, the high consistency was a result of almost no discussion-viewing activity throughout the 11-day timeline. On the contrary, the consistency achieved by Cluster 6 was achieved as a result of regular evenly spaced out work patterns, with a majority of viewing activity occurring from 3-4 days before the deadline. The work patterns for both Cluster 2 and 4 went from consistent in first two discussions (D1 and D2) to inconsistent in last two discussions (D2 and D3), although the change for Cluster 4 was not that extreme. For Cluster 2, the majority of the discussion-viewing activity in D1 and D2 took place in the middle of the timeline (approximately 10 views in a day) whereas the strategy for the third assignment included preparations starting much in advance (almost 10 days before the deadline) and finishing with another peak in discussion activities just a day before the deadline. For Cluster 4, work patterns in D1 and D2 were fairly consistent but in D3, the students performed discussion-viewing activity on the deadline only. Both Cluster 3 and 5 achieved higher consistency in work patterns as the discussion tasks progressed. However, the extremely high consistency levels witnessed in D2 and D3 in Cluster 5 were a result of students doing minimal work, whereas in Cluster 3 it was due to the similar activity levels (approx 3 views in a day) with students showing larger activity 2-3 days before the deadline.

3.4 Analysis of cluster differences based on discussion grade

The ANOVA analyses revealed statistically significant associations between final discussion scores (average of the three discussions) and learners' consistency cluster, with a large effect size ($\chi^2(5) = 22.73$, $p < 0.001$, $\epsilon^2 = .40$). The pairwise comparison of clusters revealed that performance of Cluster 1 (28.15 ± 32.05) was significantly lower than all other clusters (Cluster 2 = 82.48 ± 14.06, Cluster 4 = 85.45 ± 19.22, Cluster 6 = 64.79 ± 30.3; all ps < .05) except Cluster 3 (54.99 ± 24.03) and Cluster 5 (29.67 ± 15.49) where the differences were non-significant. The differences between the two incrementally inconsistent groups – i.e. Cluster 2 and 4 – were non-significant although the former performed significantly better than Cluster 5 and latter performed significantly better than Cluster 3 and 5. The differences between the two incrementally consistent groups, i.e. Cluster 3 and 5, were again non-significant. Overall, it seems that students who drastically changed their work patterns in the later discussions were more bound to be successful compared to those whose performance remained constant throughout.

4 DISCUSSION

Upon observing the consistency profiles obtained from each of the two activities - discussions and assignments, we found some commonalities. First, for both the learning activities, the preferred technological modality for engaging was primarily desktop. It was initially hypothesized that for discussion activity at least, mobile use would be prominent in the early days to "keep-in-touch" with the forum, followed by a switch to desktop modality as the deadline approached for creating stronger arguments requiring deeper knowledge construction. However, this was not found to be true (as seen in Section 3.1.2 & 3.3.2) and thus, it is safe to assume choice of technological modality remains consistent throughout the learning activity phases. Therefore, we agree with Sher et al. [32]'s recommendation that it is imperative for instructors to educate their students on the benefits of choosing a modality which has been theoretically-established as appropriate for a particular learning activity, as it is likely that once a modality is adopted, it will be continually used.

Second, overall student performance decrease throughout the course, as exemplified by the decline in non-zero time-series from A1 to A8 (in Figure 1) and D1 to D3 (in Figure 2), although the pattern was more prominently visible in assignment activity. This is consistent with recent research findings by Ahadi et al. [2] who found a noticeable decline in the number of students who belong to the high-performing quantile as the semester progresses, partially explainable by the incremental nature of programming.

Third, contrary to the trend in the recent literature implying a far greater degree of time management skills as learners progress in their course [17], the notion of increased consistency in temporal work patterns across subsequent learning tasks was unrelated to improvement in deadline management. In our study, we found a substantial number of students in each activity (12% (Cluster 3) and 24% (Cluster 1 and 5) in assignment and discussion activity, respectively) who maintained high consistency profiles, meaning there was little to no change in their work patterns throughout the semester. However, these were often students performing bare minimal activity in each (assignment or discussion) task or performing activity on the deadline only. In fact, these students scored the lowest in terms of academic achievement in both final grade and activity-specific grades, and qualify for those needing interventions and support the most. This data suggests that high consistency is not always a sign of excellence in learning and relying entirely upon students to ensure good time management practice is not a sign of sound pedagogical practice. Consequently, it is necessary for instructors to carefully sieve out learners with habits of effort and participation that are too similar to detrimental work patterns as identified in the literature [18, 19].

Much like lecture-specific traits are known to be associated with student persistence and engagement [13], we found possibility of some activity-specific traits to be associated with student consistency after comparing the work patterns in the two contexts – discussion and assignment. This is because the profiles obtained

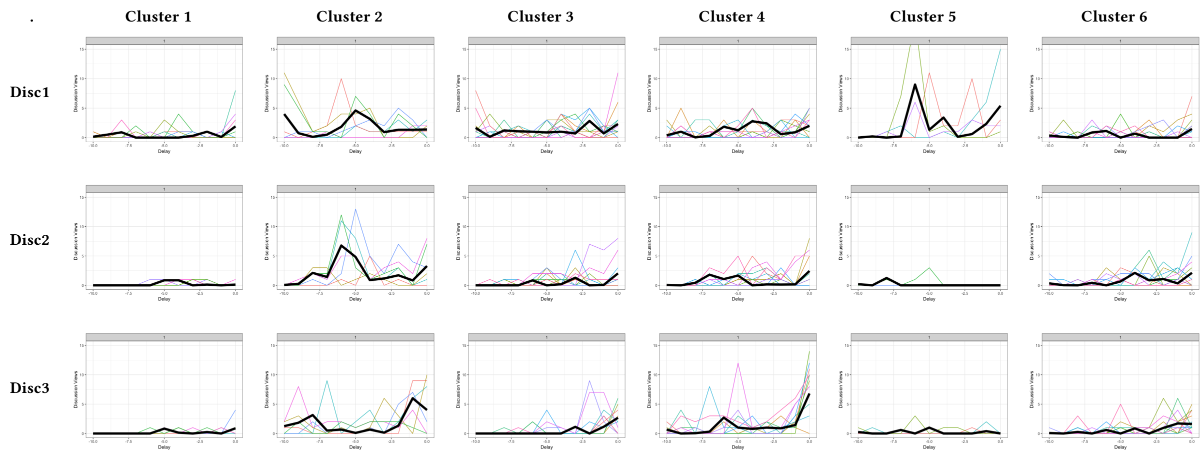

Table 2: Prototype activity patterns for the six student clusters (Cluster1 : Cluster6) at the three discussions tasks (Disc1 : Disc3). The x-axis represents the number of days before the discussion, starting from 10th day before the deadline up to the day of submission and the y-axis represents the number of discussion views. To allow cross graph comparison, all graphs have been plotted with x-axis scale [-10,0] and y-axis [0,15].

from assignment activity all seemed to converge (decreasing DTW measures from A1 to A8 in each of the five profiles in Figure 1). i.e. the engagement patterns in a specific assignment were more or less similar to the next subsequent assignments. The degree of sameness varied depending on the profile as some converged after the very first assignment (Cluster 2) whereas others after the second one (Clusters 4 and 5). On the other hand, the discussion activity witnessed instances where instead of converging, the work patterns turned excessively varied (increase in DTW measure from D1 to D3 for Cluster 2 and 4 in Figure 2). For instance, the top two high performing groups in discussion activity (Clusters 2 and 4) demonstrated very varied consistency in work patterns between the three discussion tasks as seen in Table 2. In discussions D1 and D2, Cluster 4 demonstrated patterns conventionally associated with good performance such as stable participation throughout the 11-day timeline and non-reliance on the day of the deadline to complete task and Cluster 2 showed consistency in terms of when in the 11-day timeline majority of the coursework would take place. However, for discussion D3, both clusters experimented with temporal patterns that did not match the patterns from the preceding discussion tasks, but instead were indicative of active procrastination [41], i.e. their time management strategies involved deliberately delaying task participation until the deadline. These findings suggest a continuing need to understand the nature of consistency, including when and why learners break patterns to opt for widely varied work habits and whether activity-specific traits such as inherent group dynamics, assessment methods and instructional conditions play a role in it.

In summary, for the assignment activity learners' engagement patterns were incrementally consistent with each of the eight assignment tasks. However, for the discussion activity, evidence suggested one third of the class deviated from consistent patterns towards the end (RQ1).

In the discussions surrounding aspects of consistency in learning, question of variability versus consistency does not yield an 'either-or answer' [39], since empirical support for this presumed conceptual structure is limited and conflicting. On the one hand, we can argue sticking to a routine is better while on other assert that strategies must evolve with time and needs. To give some perspective to this discussion on utility of a consistent vs. evolving learning approach, the results show that most successful group of students in terms of academic performance were those who were able to adapt their work patterns to each individual task, as a result of which their weekly patterns were unable to follow consistency (high DTW measures for each activity pair). These results partially contradict the claims by Du et al. [10] that students who study at consistent times outperform those with more varied time patterns. In fact, in our study, students whose consistency levels increased incrementally throughout the course scored lower in academic achievement than those whose consistency levels dropped. This also serves as a cautionary note to researchers using entropy [10, 21] or standard deviation [12, 20] of LMS activity measures (counts and time spent) as a measure of consistency since it could be the case that high consistency could be attributed to series of bad engagement patterns which the learner is not correcting. Thus, achieving consistency is not always analogous to excellence in learning and while time management and effort regulation may positively predicts academic grade [4], time-management skills need not be consistent but evolved enough to stabilize these efforts.

In summary, there is an association between identified patterns of consistency with student's academic performance for both assignment and discussion learning activity, although the associations are not always positive (RQ2).

Lastly, as a by-product from this study we were able to confirm claims by existing research that link certain activity patterns to academic success. The study confirmed that consistent participation is more important than high frequency [19]. Cluster 6 in discussion activities was perhaps the only cluster wherein engagement patterns were highly symmetrical in all activity phases and the high consistency in work patterns was not a result of near-zero activity. Their strategy involved studying a little bit every day which is usually associated with a 'willingness to pursue longer-term academic goals over immediate gratification' ([3], as cited in [18]), thereby explaining highest grade amongst all clusters. We also found evidence of poor academic performance linked to students working minimally on the assigned tasks or working only on deadline, in accordance with claims by Młynarska et al. [24], as exemplified by Cluster 3 in assignment and Cluster 1 and 5 in discussion activity.

5 CONCLUSIONS

The current study analyzed consistency, or persistence in work patterns on an activity-specific level. The results showed that students were incrementally consistent in their work habits over different assignments up to some extent, although students do vary in their consistency of work patterns during the discussion activity. Further, there were significant associations between identified patterns of consistency with student's academic performance for both assignment and discussion learning activity, although the associations were not always positive.

While we acknowledge that time scheduling patterns may be impacted by external factors relevant to a student – e.g., heavy course-load, freshman vs. senior time-management skills, personal commitments – in the current study we assumed that learners render equal importance to the courses used in this study (since these were mandatory and important pre-requisites to future courses) and to any other courses they may be simultaneously enrolled in. In future research, we aim to consider whether external factors are at play which may hamper or promote consistency in work patterns.

REFERENCES

[1] Saeed Aghabozorgi, Ali Seyed Shirkhorshidi, and Teh Ying Wah. 2015. Time-series clustering–A decade review. *Information Systems* 53 (2015), 16–38.

[2] Alireza Ahadi, Raymond Lister, Shahil Lal, Juho Leinonen, and Arto Hellas. 2017. Performance and consistency in learning to program. In *ACM ICPS*.

[3] Héfer Bembenutty. 2009. Academic delay of gratification, self-efficacy, and time management among academically unprepared college students. *Psychol. Rep.* 104, 2 (2009), 613–623.

[4] Jaclyn Broadbent. 2017. Comparing online and blended learner's self-regulated learning strategies and academic performance. *The Internet & Higher Education* 33 (2017), 24–32.

[5] Jim Broadbent and Walter L Poon. 2015. Self-regulated learning strategies & academic achievement in online higher education learning environments: A systematic review. *The Internet and Higher Education* 27 (2015), 1–13.

[6] Guy Brock, Vasyl Pihur, Susmita Datta, Somnath Datta, et al. 2011. clValid, an R package for cluster validation. *J. of Statistical Software* 25, 4 (2011).

[7] Christopher Brooks, Craig Thompson, and Stephanie Teasley. 2015. A time series interaction analysis method for building predictive models of learners using log data. In *Proceedings of the fifth international conference on LAK*. ACM, 126–135.

[8] Jacob Cohen. 1992. A power primer. *Psychological bulletin* 112, 1 (1992), 155.

[9] Jesus Alfonso D Datu, Jana Patricia M Valdez, and Ronnel B King. 2016. Perseverance counts but consistency does not! Validating the short grit scale in a collectivist setting. *Current Psychology* 35, 1 (2016), 121–130.

[10] Xu Du, Mingyan Zhang, Brett E Shelton, and Jui-Long Hung. 2019. Learning anytime, anywhere: a spatio-temporal analysis for online learning. *Interactive Learning Environments* (2019), 1–15.

[11] Charles Duhigg. 2013. *The Power of Habit: Why we do what we do and how to change.* Random House.

[12] Tomas Dvorak and Miaoqing Jia. 2016. Do the Timeliness, Regularity, and Intensity of Online Work Habits Predict Academic Performance?. *J. of Learning Analytics* 3, 3 (2016), 318–330.

[13] Brent J Evans, Rachel B Baker, and Thomas S Dee. 2016. Persistence patterns in massive open online courses (MOOCs). *J. High. Educ.* 87, 2 (2016), 206–242.

[14] Tak-chung Fu. 2011. A review on time series data mining. *Engineering Applications of Artificial Intelligence* 24, 1 (2011), 164–181.

[15] Toni Giorgino et al. 2009. Computing and visualizing dynamic time warping alignments in R: the dtw package. *J. Stat. Softw.* 31, 7 (2009), 1–24.

[16] Gene V Glass, Percy D Peckham, and James R Sanders. 1972. Consequences of failure to meet assumptions underlying the fixed effects analyses of variance and covariance. *Review of educational research* 42, 3 (1972), 237–288.

[17] Keith Gregory and Sue Morón-García. 2009. Assignment submission, student behaviour and experience. *Engineering Education* 4, 1 (2009), 16–28.

[18] Lauren C Hensley, Christopher A Wolters, Sungjun Won, and Anna C Brady. 2018. Academic probation, time management, and time use in a college success course. *J. of College Reading and Learning* 48, 2 (2018), 105–123.

[19] Jui-Long Hung, Morgan C Wang, Shuyan Wang, Maha Abdelrasoul, Yaohang Li, and Wu He. 2015. Identifying at-risk students for early interventions - a time-series clustering approach. *IEEE TECT* 5, 1 (2015), 45–55.

[20] Il-Hyun Jo, Taeho Yu, Hyeyun Lee, and Yeonjoo Kim. 2015. Relations between student online learning behavior and academic achievement in higher education. In *Emerging issues in smart learning*. Springer, 275–287.

[21] Jelena Jovanovic, Negin Mirriahi, Dragan Gašević, Shane Dawson, and Abelardo Pardo. 2019. Predictive power of regularity of pre-class activities in a flipped classroom. *Computers & Education* 134 (2019), 156–168.

[22] Leonard Kaufman and Peter Rousseeuw. 1990. *Finding Groups in Data: An Introduction To Cluster Analysis.*

[23] Greig Emil Krull. 2017. *Supporting seamless learning: Students' use of multiple devices in open and distance learning universities.* Ph.D. Dissertation. Universitat Oberta de Catalunya.

[24] Ewa Młynarska, Derek Greene, and Pádraig Cunningham. 2016. Time series clustering of Moodle activity data. In *24th Irish Conference on Artificial Intelligence and Cognitive Science (AICS'16).*

[25] Suzanne Nolte et al. 2016. *The comparison of behavioural patterns of top achievers from diverse work environments: an analysis of Shadowmatch database benchmarks.* Ph.D. Dissertation. University of Pretoria.

[26] François Petitjean, Alain Ketterlin, and Pierre Gançarski. 2011. A global averaging method for dynamic time warping, with applications to clustering. *Pattern Recognition* 44, 3 (2011), 678–693.

[27] Paul R Pintrich. 2004. A conceptual framework for assessing motivation and self-regulated learning in college students. *Educ. Psychol. Rev* 16, 4 (2004), 385–407.

[28] Michelle Richardson, Charles Abraham, and Rod Bond. 2012. Psychological correlates of university students' academic performance: A systematic review and meta-analysis. *Psychological bulletin* 138, 2 (2012), 353.

[29] Peter J Rousseeuw. 1987. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.* 20 (1987), 53–65.

[30] Hiroaki Sakoe and Seibi Chiba. 1978. Dynamic programming algorithm optimization for spoken word recognition. *IEEE transactions on acoustics, speech, and signal processing* 26, 1 (1978), 43–49.

[31] Alexis Sardá-Espinosa. 2018. Comparing Time-Series Clustering Algorithms in R Using the dtwclust Package.

[32] Varshita Sher, Marek Hatala, and Dragan Gašević. 2019. On multi-device use: Using technological modality profiles to explain differences in students' learning.. In *Proceedings of the Ninth International Conference on LAK*. ACM.

[33] Glenn Stockwell. 2013. Tracking learner usage of mobile phones for language learning outside of the classroom. *CALICO Journal* 30, 1 (2013), 118–136.

[34] Lennart Svensson. 1977. On Qualitative Differences in Learning: III–Study Skill and Learning. *Br. J. Educ. Psychol.* 47, 3 (1977), 233–243.

[35] Patrick R Thomas and John D Bain. 1982. Consistency in learning strategies. *Higher Education* 11, 3 (1982), 249–259.

[36] Yvonne J Vermetten, Hans G Lodewijks, and Jan D Vermunt. 1999. Consistency and variability of learning strategies in different university courses. *Higher Education* 37, 1 (1999), 1–21.

[37] Yvonne J Vermetten, Jan D Vermunt, and Hans G Lodewijks. 1999. A longitudinal perspective on learning strategies in higher education-different view-points towards development. *Br. J. Educ. Psychol.* 69, 2 (1999), 221–242.

[38] Jan D Vermunt. 1998. The regulation of constructive learning processes. *Br. J. Educ. Psychol.* 68, 2 (1998), 149–171.

[39] Jan D Vermunt and Yvonne J Vermetten. 2004. Patterns in student learning: Relationships between learning strategies, conceptions of learning, and learning orientations. *Educ. Psychol. Rev* 16, 4 (2004), 359–384.

[40] Philip H Winne and Dianne Jamieson-Noel. 2002. Exploring students' calibration of self reports about study tactics and achievement. *CEP* 27, 4 (2002), 551–572.

[41] Christopher A Wolters, Sungjun Won, and Maryam Hussain. 2017. Examining the relations of time management and procrastination within a model of self-regulated learning. *Metacognition and learning* 12, 3 (2017), 381–399.

[42] Mingming Zhou and Philip H Winne. 2012. Modeling academic achievement by self-reported versus traced goal orientation. *Learning and Instruction* 22, 6 (2012), 413–419.

[43] Barry J Zimmerman. 2008. Investigating self-regulation and motivation: Historical background, methodological developments, and future prospects. *Am. Educ. Res. J.* 45, 1 (2008), 166–183.