An analytics-based framework to support teaching and learning in a flipped classroom

An Analytics-based Framework to Support Teaching and Learning in a Flipped Classroom

Jelena Jovanovic, Negin Mirriahi, Abelardo Pardo, Shane Dawson, Dragan Gašević

1. Introduction

Flipped Classroom (FC) is a particular form of active learning design that requires students to complete assigned activities prior to attending class (i.e. face-to-face sessions with the instructor). The completion of preparatory work affords greater opportunity for class time to be dedicated to active learning tasks. As models of FC practices are increasingly being adopted in higher education there is a corresponding increase in the level of research investigating the impact on student learning. To date, the research has demonstrated that FC models can lead to improvements in students' performance, as measured on pre- and post-tests and course exams (O'Flaherty & Phillips, 2015). However, uptake and acceptance of this type of active learning design has been countered by instructor anxiety regarding a student's level of engagement with, and completion of, the preparatory activities. The challenge for students relates to the high degree of autonomy required by FC models (Kim, Kim, Khera, & Getman, 2014). Hence this form of active learning and teaching requires a higher level of proficiency in self-regulation skills than traditional 'lecturing' model.

To promote adoption of the FC model and support students in regulating their learning, it is important to shed some light on learning strategies students adopt as they undertake and complete learning activities in a FC. This is particularly relevant for the preparatory learning activities as these activities are crucial for the overall success of the FC model (Rahman et al., 2015). In addition, it is important to examine how the adopted learning strategies change over the duration of a course. Students demonstrating a high level of self-regulated learning proficiency will modify their learning behaviour to better adapt to the specific course design and requirements (Winne & Hadwin, 1998; Winne, 2006).

To date there has been scant research investigating student learning strategies in FC settings (Anonymous, 2017a). What is particularly lacking is research into dynamics of students' strategies, and the associated question of the level of granularity that would be appropriate for examining this dynamic. The purpose of this chapter is to address this gap by proposing a learning analytics framework for identifying discrete learning strategies and their dynamics in the context of a FC. The area of Learning Analytics focuses on the use of data captured during learning experiences to increase the level of understanding of learning processes and improve the context in which they occur (Siemens & Gasevic, 2012). FC is an ideal setting for the use of learning analytics techniques since preparation activities are usually mediated by technology, thus providing detailed data about individual student interactions with learning resources.

The proposed framework is based on the consolidated model of learning analytics by Gašević, Kovanović and Joksimović (2017), which posits that in order to produce results that are effective and valid for both research and practice, learning analytics approach has to be based on three mutually connected dimensions: theory, design and data science. From the theoretical perspective, the proposed framework is grounded in Winne and Hadwin's model of self-regulated learning (SRL) (Winne & Hadwin, 1998). From the data science perspective, the framework relies on methods for the analysis of trace (log) data concerning students' engagement with online preparatory activities. The analysis of trace data involves the use of different unsupervised and temporal data science techniques such as clustering, latent class analysis, and sequence analysis. Finally, from the design perspective, the framework adopts design-based research (DBR) (Reimann, 2011) as the study method, and FC as the general learning design. DBR is considered as the study design of choice for learning analytics, as both (DBR and learning analytics) are examining learning in context and both are aimed at improving the practice and advancing the theory of learning (Reimann, 2016; Gašević et al., 2017).

1.1 Flipped classroom

Flipped or inverted classroom is a form of blended learning that requires students' active participation in learning activities both before and during face-to-face teaching sessions (Lage, Platt, & Tregua, 2000). The FC model assumes that students prepare for a teaching session by completing a set of assigned learning activities before scheduled class time (Strayer, 2012). Initial FC approaches typically relied on reviewing online lecture recordings as the primary activity assigned to students prior to face-to-face sessions. However, with the increasing sophistication of educational technologies the types of activities and resources incorporated into FC models have become more diverse and engaging involving specialized videos, audio clips, interactive exercises and simulations. Through these preparatory activities, students develop the basic knowledge and skills required for active engagement in the subsequent face-to-face session with the teacher. Face-to-face teaching sessions typically include individual or collaborative problem solving, as well as critically discussing and/or debating issues related to the topics introduced in the preparatory activities.

Prior studies on FC have tended to focus on student satisfaction with this mode of teaching and its impacts on the overall course performance (O'Flaherty & Phillips, 2015). Most studies have relied on questionnaires and interviews to collect students' opinion and perceptions of FC. Pre- and post-tests and final course grades have been used to assess the impact of FC on academic performance. The majority of the studies have reported positive educational outcomes associated with the FC design. Benefits include increased student satisfaction (e.g., Forsey, Low, & Glance, 2013), improved course grades (e.g., Pierce & Fox, 2012), and higher levels of attendance (e.g., Prober & Khan, 2013).

An aspect of FC that has received limited attention relates to the kinds of learning strategies that students apply in this particular kind of course design. Considering that FC encourages student ownership of learning, and at the same time differs considerably from the conventional teaching practices, it is important to understand how students approach and manage their learning in a model that requires considerable autonomy. That is, there is a need for research to investigate the kinds of strategies students adopt and how they regulate those strategies throughout the duration of a course. The relevance for this work stems from research noting that students frequently lack the necessary skills and proficiency to adapt their learning strategies to the specificities of newly encountered learning situations (Winne & Jamieson-Noel, 2003; Lust, Elen, & Clarebout, 2013a). This is particularly the case when the acquired study skills need to be transferred to substantially different learning contexts - this problem has been referred to as "far transfer" (Hattie, Biggs, & Purdie, 1996). The significant differences between FC and traditional lecturing models suggests that students who have experienced lecturing as the primary teaching method would need to make a "far transfer" of their previously developed study skills. Hence, students are likely to encounter difficulty with learning strategy regulation. Consequently, it is reasonable to assume that students new to the FC model will employ suboptimal learning strategies.

1.2 Learning strategies

A student's adopted study approach or learning strategy includes "*any thoughts, behaviours, beliefs or emotions that facilitate the acquisition, understanding or later transfer of new knowledge and skills*" (Weinstein, Husman, & Dierking, 2000, p. 227). This suggests that learning strategy is a latent construct that can neither be directly observed nor measured. Prior studies examining learning strategies have relied on learners' self-reports (e.g. questionnaires or think-aloud protocols) as the primary means of identifying strategies adopted by learners in a particular setting (see e.g., Hill & Hannafin, 1997; Bannert, Reimann, & Sonnenberg, 2013). However, this data source suffers from several deficiencies, such as inaccuracy due to the poor recall of prior behaviour, in the case of questionnaires (Winne & Jamieson-Noel, 2002), and high level of cognitive load placed on the participants, in the case of think-aloud protocols (Winne, 2013).

When students engage in a set of learning activities in digital settings, their learning strategies can be identified from the user trace data (i.e. data collected from the tools and services the students have interacted with during the learning process). As a data source for studying learning strategies, trace data have some important advantages over self-reports (Winne, 2013). For example, the use of trace data allows for an analysis of the learners' actual behaviour in lieu of their perceptions and recall of events. Furthermore, the data collection does not interfere with the learning process. The analytics-based framework proposed in this chapter draws on trace data to examine the various learning strategies students adopt when preparing for FC sessions.

The proposed analytics-based framework is based on Winne and Hadwin's model of self-regulated learning (SRL) (Winne & Hadwin, 1998), which positions learners as active agents who use cognitive, physical and digital tools to operate on raw information in order to create products of learning. A key feature of learner agency is the regulation of learning behaviour through the continuous evaluation of the quality of the learning products and the effectiveness of the adopted study tools and tactics. This process of metacognitive monitoring is influenced by internal and external conditions. The former include, for example, a learner's level of motivation, prior knowledge, and affective states; the latter are primarily determined by the elements of the instructional design (e.g., the teacher's role, course requirements, and availability of feedback). Students regulate their learning by changing their study tactics and strategies in response to the changes in the external and internal conditions. For example, students are often enrolled in several courses at the same time, and high workloads in some of those courses may require them to reduce their involvement in other courses for a certain period of time. Therefore, we find it important to study how learning strategies – derived from the collected trace data – change from one study unit to the next to account for the potential changes in external and/or internal conditions of the learners.

2. Analytics-based framework for examining learning strategies in FC

The proposed analytics-based framework is illustrated in Figure 1(a) and consists of four stages: (1) selection of data sources, (2) feature creation, (3) building of (unsupervised) analytical models, and (4) interpretation of the results. The structure and content of these stages are informed by the underpinning learning theory, the learning design, and the research design. In particular, the theory underlying the framework is the Winne and Hadwin's model of self-regulated learning (SRL) (Winne & Hadwin, 1998). It directly informs the selection of variables (Stage 2) and interpretation of the results (Stage 4). In addition, by informing the learning design, the learning theory has an indirect impact on practically all the stages of the framework. The learning design shapes external conditions (Winne & Hadwin, 1998) and thus needs to be carefully considered when creating features for data analysis (Stage 2), and also when interpreting the results of the analysis (Stage 4) (Lockyer, Heathcote, & Dawson, 2013). Finally, the selection of data (Stage 1), the creation of features (Stage 2), and the choice of analytical methods (Stage 3) depend on the research objective, that is the learning construct/process one wants to better understand and/or the aspect of the learning experience one aims to optimize. This research component of the framework, namely DBR has also its practical objective - to address problems as they emerge in practice. Through the four stages of the framework, we can understand effectiveness of the decisions made in the learning design and consequently feedback results of analysis to inform next offerings of the course. Thus, the interpretation of the results (Stage 4) has a potential to inform revisions of the learning design. This iterative nature of the framework, with specific intervention in each iteration, promises to not only lead to the improvements in teaching/learning practice, but eventually to also contribute to the learning theory.

Stage 1 consists of the selection of data sources that are going to be used for the analysis. In order for this framework to adequately support the FC context, the data obtained has to contain significant information about student engagement with the learning tasks as envisioned in the learning design. A substantial component of technology mediation is assumed in the framework that translates into an equally rich set of data sources available. In order to provide the right level of support for the rest of the stages in the framework, these data sources must provide the appropriate level of granularity in their observations. For instance, fine granularity data sources such as clickstreams allow for the detection of student engagement in almost real time.

Stage 2 of the framework is the most delicate because it includes an informal process by which the data collected in the previous stage is combined into higher level indicators of student engagement. This process requires the expertise of instructors and designers alike (Gašević, Dawson, Rogers, & Gašević, 2016). A single set of events captured in Stage 1 removed from the context may have no significance, whereas, a specific set of events during a specific time frame can be identified as the indicators to be used as features for a data analysis model. Examples of such features include engagement with video resources (e.g. time watching a video clip, number of times a video is reviewed, etc.), interaction with formative assessment elements (e.g. quizzes with automatic grading), resource annotation (e.g. comments in a wiki or a post), participation in a discussion (statistics extracted from a discussion forum), etc.

An outcome of Stage 2 is the establishment of a learning experience that is represented by a rich set of features that can be used as reliable indicators of student engagement. Stage 3 involves the application of discrete algorithms to process these comprehensive sets of features about student engagement and detect latent constructs that are not possible to detect manually. For this step the framework suggests the use of unsupervised data analysis methods (e.g., clustering, sequence analysis) as these have proven successful in unveiling latent constructs from data (see e.g. Kovanovic, Gašević, Joksimović, Hatala, & Adesope, 2015; Lust et al., 2013a; Perera, Kay, Koprinska, Yacef, & Zaïane, 2009). In the context of this framework, these methods return, for example, a characterisation of various student groups each of them with similar features.

The results produced in Stage 3 include the characterisation of groups of observations based on values of the selected features. This characterisation is interpreted in Stage 4 in the context of the overall learning experience and transformed into insights. As in the case of Stage 2, this interpretation requires substantial support from instructors and designers to define the elements related to the pedagogical intent of the learning design (Lockyer et al., 2013). This way the result of the algorithms can be translated into corresponding definitions of student learning strategies.

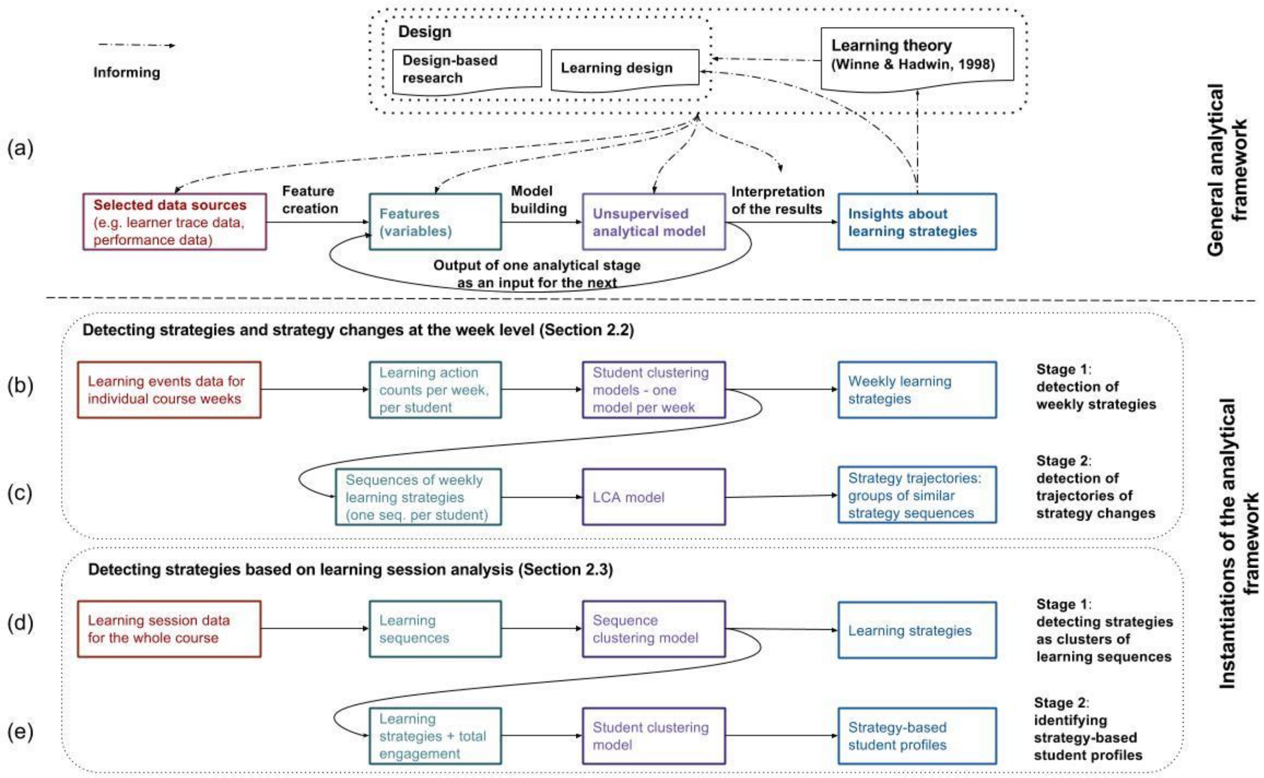

Figure 1. The proposed analytics-based framework for examining learning strategies (a); an instance of the framework for detecting strategies and strategy changes at the level of a course week (b, c); another instance of the framework for detecting strategies through learning session analysis (d, e)

The model has three implicit loops that allow for its use in an iterative manner. The first loop is represented by the connection from Stage 3 back to Stage 2. This connection represents the possibility of using the information extracted from the unsupervised analysis and create new features. For example, if students are grouped based on their strategies to engage with multimedia content, those strategies themselves can be used to form a feature for the next iteration of analysis to detect patterns that take into account this type of engagement. The second loop is captured by the connection between the result of Stage 4 and the structure of the learning design. This relationship represents the possibility of using the insights derived from the enacted design to prompt changes in the design itself. Finally, the third loop is established through the connection between the interpreted results (Stage 4), that is the newly acquired insights and the learning theory. This loop has a potential to, for example, improve our understanding of how learning design, as a major component of external conditions (Winne & Hadwin, 1998), influences the learner performance or even the overall learning experience.

The generality of the framework makes it suitable to be applied to any learning scenario in which a rich set of data provides a detailed account of student interaction, and there is a set of features identified by instructors and designers that are useful for the overall improvement of the experience. Conventional design methodologies that rely on evaluation mechanisms to assess the design after being enacted and decide improvements can also be modelled by the proposed framework. However, this is just a special case. As described, the framework can be used for learning scenarios in which the insights extracted in Stage 4 can be almost immediately reflected on the experience thus increasing the immediacy effect. It is for this reason that FC designs are ideal candidates to instantiate the framework. The overall FC model assumes sustained student engagement, and as a consequence, analysing features that help instructors to gain insight in the learning process are essential for the success of the overall experience.

To better demonstrate the potentials of the framework, we present it in the context of a particular FC course where it has been successfully applied. The design of this course is briefly introduced in the next section to aid comprehension of the framework and the associated findings.

2.1 An application context: an undergraduate course with FC design

The framework was applied to a study of first year undergraduate students enrolled in a Computer systems course at a large Australian research-intensive university. The course enrolled 290 students (81.5% male, 18.5% female) and ran for 13 weeks (1 semester). The course adopted a FC design, whereby students were provided with a set of weekly online tasks to prepare for the face-to-face teaching sessions. These tasks were designed to provide students with the basic knowledge required for active participation in weekly collaborative problem solving tasks undertaken in the face-to-face session.

The weekly preparation tasks included: short videos followed by multiple-choice questions (MCQs) covering the concepts discussed in the video; reading materials with embedded MCQs; and problem (exercise) sequences. MCQs associated with the course videos and those embedded in the reading materials were framed as formative assessment, whereas exercises served as summative assessment. Students were also provided with real-time feedback on their level of engagement with the preparation activities and their activity scores via an analytics dashboard (Anonymous, 2016).

A detailed description of the course learning design, including examples of preparation tasks and real-time feedback via learning analytics dashboard, is outlined in (Anonymous, 2017b).

2.2 Detecting strategies and strategy changes at the week level

From the theoretical background (Winne, 2006; Winne & Hadwin, 1998), we posit that learning strategies change over time, as learners regulate their study approach to the changes in internal and/or external conditions. As course designs in higher education are commonly based on weekly cycles, it is reasonable to expect that learners will adapt their strategies on the weekly level, in response to the week's topic of study, requirements and other week-specific conditions. To examine this, the first instantiation of the proposed analytics-based framework was set towards detecting the strategies that students adopt in each week of the course, as well as identifying patterns of strategy changes during the course (Figure 1 (b, c)). More specifically, considering the particular learning context of FC, this instance of the framework is focused on identifying weekly strategies and patterns of strategy changes in relation to the preparation tasks in a FC (i.e., the tasks that students are requested to complete to prepare for the face-to-face teaching sessions). The details of the chosen data analysis methods and their application to the referenced FC course are given in (Anonymous, 2017c).

2.2.1 Detection of weekly strategies

The first iteration of analysis (Figure 1b) consists of clustering students using the trace data that originate from students' engagement with the preparation tasks in each week of the course. In particular, agglomerative hierarchical clustering based on the Ward's method (Ward, 1963) is used in conjunction with a set of features derived from the students' weekly interactions with various kinds of learning resources made available to them for the preparation tasks (e.g. short videos, reading materials, multiple-choice questions, exercises). This clustering algorithm is chosen as it has proven to be particularly suitable for detecting student groups in online learning contexts (Kovanovic et al., 2015). The clustering algorithm leads to the detection of groups (clusters) of students with similar engagement patterns, that is, learning strategies in each week of the course. By repeating the clustering procedure with the trace data from each week of the course, we obtain for each student, a sequence of weekly cluster assignments. For instance, for a FC course of 12 weeks, at the end of this data analysis stage, each student is described with a vector of 12 numerical values, each one representing the cluster assignment in the corresponding week of the course (i.e. identifying the student's learning strategy in the respective week). These results are further examined and interpreted in the following two steps. First, the clusters identified in each week are interpreted (based on the cluster-specific feature values), and each cluster is assigned a label characterising the strategy the cluster represents. Next, the (interpreted) clusters from all the course weeks are further analysed in order to identify clusters (i.e. strategies) from different weeks that share similar features, and assign common labels to such clusters.

When applied to the trace data collected in the examined FC setting (Section 2.1), the clustering method described above led to the detection of 4-6 clusters for each of the 12 active weeks of the course. Having interpreted the weekly clusters and identified the clusters that represented similar strategies in different weeks of the course, we obtained a total of eight distinct clusters, that is, week-level learning strategies:

- *Zero-engagement* - represents the strategy of those students who, in the given week, did not engage at all with the preparation activities.
- *Performance-orientation* - represents the strategy of students with high level of interaction with summative assessment exercises and very few interactions with the rest of resources. We identified two variants of this general strategy, and considered each as a separate learning strategy: *Performance-oriented – summative focused* and *Performance-oriented – summative and video focused*. The former is characterized by a slightly better ratio of correct answers to summative assessment and a slightly higher interaction with the course videos than the latter one.
- *Minimalist – summative focused* and *Minimalist – summative and video focused* strategies are characterized by student engagement with all kinds of preparation activities but with a moderate to low frequency. The difference between these two strategies is primarily in the level of engagement with the course videos.
- *Task-oriented* strategy is characterized by the student engagement with all preparation activities and a high percentage of correct answers to both formative and summative assessment items.
- *Intensive and ineffective* strategy is that of students who, in the given week, had the highest level of engagement with all preparation activities but a lower percentage of correct answers (than *Task-oriented* students).
- *Selective and effective* strategy, present in only two weeks of the course, reflects an irregular pattern of engagement of students who interacted only with summative exercises, and achieved an unusually high percentage of correct answers. This may be the strategy of high achieving students with solid prior knowledge, or of those that are obtaining the answers through other sources and engaging with the system solely to submit them.

These strategies serve as the input for the next iteration of the analysis, namely the detection of common trajectories of strategy changes.

2.2.2 Detection of trajectories of strategy changes

Post detection of the student learning strategies typical for individual course weeks, the second iteration of the analysis (Figure 1c), applies these strategies as the input to identify common trajectories of learning strategy changes within the course. More specifically, the analyses extend to identify groups of students with similar sequences of learning strategies. To this end, we use latent class analysis (LCA), a statistical technique that allows for finding groups or subtypes of cases in multivariate categorical data (McCutcheon, 1987). LCA can be thought of as a soft-clustering technique since it assumes that each case belongs to each group with a certain probability, and computes these probability distributions. The optimal number of groups is determined based on statistical indicators such as Akaike information criterion (AIC) and Bayesian information criterion (BIC).

Vectors of student cluster assignments in each week of the course (i.e. sequences of strategies within the course) are used as feature vectors for building an LCA model. The LCA computes the similarity of the input feature vectors to identify groups of students who were likely to have common trajectories of strategy changes over the course weeks. We refer to the LCA results as trajectories since each one represents a trajectory or a sequence of learning strategies common to a group of learners who exhibited similar dynamics in their engagement patterns within the course.

When the analytics-based framework is applied to the examined FC course (Section 2.1), the LCA technique identified six trajectories of learning strategy use, each with a distinct pattern of strategy changes over the duration of the course. The trajectories included:

- The trajectory of *Non-users* (56, 19.0%) gathers students who have hardly completed any preparation activity in most of the course weeks.
- The trajectory of *Performance-oriented minimalist* (87, 29.7%) mostly consists of students who follow the *Performance-oriented* and *Minimalist* strategies, i.e. those having low levels of engagement with a limited set of activities.
- The *Disenchanted* trajectory (45, 16.0%) is that of students who mostly followed the *Task-oriented* and *Intensive and ineffective* strategies; over time these students gradually reduced their level of engagement with the preparation tasks.
- The *Mixed* trajectory (28, 10.1%) gathers students who experimented with a broad mix of strategies, and who increased amount of activity and effectiveness towards the end of the course.
- The *Intensive* trajectory (60, 20.3%) is characterised by highly active students for whom the most likely strategies during the course were *Task-oriented* and *Intensive and ineffective*.
- The *Intensive and ineffective* trajectory (14, 4.8%) is dominated by students with a very high level of engagement and high probability of using the *Task-oriented* strategy throughout the course.

The comparison of these trajectories with respect to the students' academic performance revealed two distinct groups of trajectories: i) those that lead to low academic performance: Non-users, Performance-oriented minimalists, and Intensive and ineffective; and ii) those associated with high academic performance: Disenchanted, Mixed, and Intensive.

2.3 Detecting strategies based on learning session analysis

The second instance of the proposed framework (Figure 1(d, e)) is based on an analysis of students' learning sessions. This approach relies on different level of granularity of the input data and examines different kind of dynamics than the previously described instance (Section 2.2). In particular, it is focused on an analysis of sequences of learning actions and the dynamics of learning strategies at the level of learning sessions. The details of this data analysis method and its application to the examined FC course are given in (Anonymous, 2017a).

Learning sessions are extracted from the trace data and encoded as sequences of consecutive learning actions performed by students while interacting with learning resources provided to them for the preparation tasks. In particular, each session is encoded as a sequence of discrete states, where each state corresponds to one learning action. There are as many states as there are learning actions relevant for a particular learning design (i.e. the design of activities that should prepare students for face-to-face teaching sessions). For instance, the following sequence:

```
(VIDEO_PLAY,1)-(MCQ_IN,2)-(MCQ_CO,2)-(MC_EVAL,1)
```

encoded in a format suitable for subsequent analysis (Section 2.3.1), starts with an action of activating a course video (`VIDEO_PLAY`), which is followed by two consecutive actions of incorrectly solving formative multiple choice questions (`MCQ_IN`), which is in turn followed by two correctly solved formative multiple choice questions (`MCQ_CO`), and terminates with an action of accessing the dashboard (`MC_EVAL -` metacognitive evaluation action).

Learning sessions represented in this way are referred to as learning sequences. They are used as input for the the first iteration of analysis: grouping (clustering) of similar learning sequences to detect patterns in students' learning behaviour that are indicative of the students' learning strategies (Figure 1d). The identified patterns in learning sequences (i.e. strategies) serve as the input for the next iteration of analysis, which is about detecting groups of students based on the students' distinct use of learning strategies (Figure 1e). In both iterations of analysis, groups are identified through agglomerative hierarchical clustering.

2.3.1 Detection of strategies through clustering of learning sequences

Clustering of learning sequences, to detect patterns indicative of students' learning strategies (Figure 1d), requires computation of similarity of sequences. Although there are various similarity measures that can be applied to calculate similarity (or distance) of sequences, the one that is most frequently used for the kind of sequences we are dealing with (i.e., sequences of discrete states) is the optimal matching (OM) metric (Studer & Ritschard, 2016). According to this metric, the distance between any two learning sequences is the minimal cost - in terms of insertions, deletions and/or substitutions of learning actions - required for transforming one sequence into another. Whereas insertions and deletions costs are often set to a constant value (typically 1), the substitution costs are computed based on the probability of observing state (i.e. learning action) j at time $t + 1$ given that the state (i.e. learning action) i has been observed at time t (Gabadinho et al., 2011). By applying the OM metric in conjunction with the Ward's method for hierarchical clustering, we identify groups (clusters) of learning sequences that represent different patterns in student engagement with the preparation tasks.

When applied to the learning sequences derived from the trace data of the examined FC (Section 2.1), this clustering method led to the detection of four distinct clusters, that is, manifestations of the learning strategies that students adopted when preparing for face-to-face teaching sessions in the studied FC. The detected clusters and the corresponding strategies are as follows:

- *Formative focused* strategy is dominated by formative assessment actions, with summative assessment almost absent. Access to the course content is not frequent, though it tends to be more present at the beginning and towards the end of the learning sessions. This is the smallest cluster indicating that the corresponding strategy is the least frequently chosen one.

- *Summative focused* strategy is the most dominant one with a clear preference for actions related to summative assessment exercises. It is characterized by trial-and-error learning approach, which is evident in the considerably higher number of incorrect than correct exercise attempts in the corresponding learning sequences. Another feature of sequences from this cluster is that they often end with metacognitive evaluation actions, that is, access to the course dashboard.

- The strategy of *Low active engagement* is characterized by a low level of active engagement with the learning resources (i.e. passive 'consumption' of the provided materials). It predominantly consists of accessing the course content and only a fraction of formative assessment. Learning sequences representative of this strategy tend to be shorter than sequences of the other strategies, and typically end by playing course videos.

- *Multiple activities* strategy consists primarily of playing course videos, then doing the follow-up multiple-choice questions (i.e. formative assessment), and finally trying the exercises (i.e. summative assessment). Another distinct feature is the presence of metacognitive actions at the beginning of sessions corresponding to this strategy.

2.3.2 Identifying student profiles based on the learning strategies

Clustering of learning sequences done in the first iteration of analysis (Section 2.3.1) produces a vector of discrete numerical variables, $seq.clust_I$, i=1:n_c, for each student, where $seq.clust_I$ is the number of learning sequences in cluster (i.e. learning strategy) *i* for a particular student, and n_c is the number of clusters. These variables plus the variable (seq.total) representing the total number of learning sequences per student are used as the input for the second iteration of analysis, namely clustering of students (Figure 1e). The objective of this iteration is to identify distinct student profiles based on the learning strategies the students used throughout the course. When applied to the examined FC, this analysis produced the following five student strategy profiles:

- *Intensive* (19, 6.6%): students with high level of engagement, who undertook a variety of learning strategies, among which the strategy of *Low active engagement* and the *Summative focused* strategy were the most prominent.

- *Strategic* (35, 12.1%): similar to the *Intensive* group, but with a lower level of engagement and a reversed level of importance placed on the *Summative focused* and *Low active engagement* strategies. The group's primary focus on assessment activities (both summative and formative) suggests that these students were performance-oriented, whereas their overall level of engagement indicates a preference for efficiency. Since the course performance of this group did not significantly differ from the best performing (*Intensive*) group, it is reasonable to conclude that these students were strategic in choosing their strategies.

- *Highly strategic* (50, 17.2%): this group is similar to the *Strategic* group in terms of their performance orientation and preference for efficiency. The difference is in a significantly lower adoption of the *Formative focused* and *Multiple activities* strategies. However, in spite of their lower level of overall engagement, the students in this group achieved the performance level of the other two high performing groups (*Intensive* and *Strategic*), thus proving to be even more successful in regulating their learning (i.e., in choosing and/or adapting their strategies).

- *Selective* (128, 44.1%): this is the largest group with a strong preference for the *Summative focused* strategy, though these students also experimented with other learning strategies. The group's overall level of activity and exam scores were significantly lower than those of the previous three groups.
- *Highly selective* (58, 20.0%): the main features of this group is the lowest level of activity and almost exclusive focus on summative assessment (*Summative focused* strategy). These students demonstrated the lowest performance on the course exams.

These five groups also differ in terms of how and to what extent their learning strategies varied throughout the course. The *Intensive* and *Strategic* groups engaged with all learning strategies until a week after the midterm exam (week 7); from that week until the end of the course, they abandoned the *Formative focused* strategy, and even though they retained the other three strategies, they showed preference for the *Summative focused* strategy. *Highly strategic* students used different learning strategies only during the first two weeks of the course and when preparing for the midterm exam (week 6); in all other weeks, they opted for the *Summative focused and Low active engagement* strategies, with the former being the preferred one. Finally, the *Selective* and *Highly selective* groups had strong preference for the *Summative focused* strategy throughout the course. In fact, *Selective* students occasionally undertook the strategy of *Low active engagement* in the weeks prior to the midterm exam, whereas for *Highly selective,* the *Summative focused* strategy was the only choice during the entire course.

3. Discussion

The proposed analytics-based framework and the two presented instances of the framework have so far been studied and evaluated in the context of one particular FC design (Section 2.1). The obtained results are largely consistent with previous findings in educational research (as discussed in detail in (Anonymous, 2017a)). For instance, the identified strategy-based student profiles (Sections 2.2.1 and 2.3.2) correspond well to those reported in previous research (e.g., Lust et al., 2013a; Lust, Elen, & Clarebout, 2013b; del Valle & Duffy, 2009; Kovanovic et al., 2015). This suggests that the findings might be transferable to other domains and FC designs. However, as given in the propositions of our analytics-based framework, any learning analytics finding has to be interpreted in the context of the learning design the learner trace data originate from. Therefore, further investigation is required before the transferability of the findings can be confirmed. As for the proposed analytics-based framework and the two presented instances, these can be reused in other FC settings provided an appropriate mapping between the data captured during the experience and the factors to be analysed is provided. In other words, the collected data need to reflect the interactions relevant for the specific learning design of the module / course under study. In fact, this requirement is one of the main premises of the analytics-based framework itself.

Further, the main purpose of the framework is to provide a model in which learning strategies are interrogated and identified in the context of a specific learning design in order to reveal insights about learning (stage 4 in the framework) and, subsequently, inform the further development or establishment of learning theory, a core component of DBR. As posited by Reimann (2016), DBR and Learning Analytics (LA) strategies provide opportunity for synergies whereby the fine-grained and nuanced data used in LA methods can advance learning theory and inform the development of pedagogical interventions which is the heart of design based research. The introduced framework helps formalise the relationship between LA data collection and analysis strategies with the interpretation and application of findings to inform changes to the learning design and, when applicable, SRL, the learning theory underpinning the framework. The course instructors and/or designers play a critical role in the process by providing expertise (Gašević et al., 2016) in interrogating and interpreting the data in their learning context and advancing the learning design and theory.

While the presented data analysis methods allow for the detection of learning strategies as regularities (patterns) in students' learning behaviour, they are limited in their ability to explain that behaviour and the identified patterns. This is a general problem of approaches based on the analysis of trace data (Reimann, Markauskaite, & Bannert, 2014). To move beyond detection and description of learning strategies and towards full understanding of student learning behaviour, we need to examine connections between the strategies adopted by learners (i.e. the identified strategy-based learner profiles) and learners' motivation and goal orientation. Self-report measures have traditionally been used for gaining an insight into students' motivation. However, they have limited ability in capturing the dynamics of students' learning motivation and goals (Zhou & Winne, 2012), which are generally stable, but still prone to change along with changes in external conditions (Fryer & Elliot, 2007). Furthermore, there is the question of the students' ability to report on their goal orientations in an objective manner (Richardson, 2004). All this implies a need for extending learning environments with instruments that would allow for seamless and unobtrusive collection of data about the dynamics of students' learning motivation and goal orientation. Such developments would set the grounds for complementing the presented data analysis methods with additional ones capable of addressing the question of *why* students may have opted for particular strategies or why they may have abandoned or continued with such strategies during the course.

4. To the Classroom

In this section, we discuss how educators can apply the proposed framework to their own context in order to better understand their students learning behaviours and make informed changes to their teaching approach and course design.

Using the framework, two different approaches of identifying learning strategies were explored. The first approach focused on how students engaged with the course materials and activities in each week of the course, and how they subsequently performed on their summative assessments. The first phase of exploring the weekly learning strategies focused on identifying clusters of students with the same type of learning behaviour in each week of the course. In our case study, eight learning strategies were identified which reveal that not all students engaged with the course resources in the same way or to the same extent. For example, some students achieved high scores on the summative exercises despite only rarely accessing the course materials (*Selective and effective* strategy), while other students had a high level of engagement with the course materials and activities but low number of correct responses on the summative assessment (*Intensive and ineffective* strategy). For the educator, this sort of analysis reveals that there may be a mismatch between the preparation activities (e.g. video and readings) and the summative exercises, suggesting that future iterations of the course may need adjustments to one or the other. The same sort of analysis can be conducted in a different learning context by identifying the course weeks the analysis should focus on and the types of activities and resources students are expected to access.

The second phase of studying students learning strategies focuses on exploring changes in students learning behaviour over the duration of the course. Using the clusters of learning behaviour detected in the first phase, the second phase employs another data analysis technique, LCA, to identify the trajectories of learning behaviour across time and how they correspond with academic performance. In other words, this analysis reveals groups of students with similar learning behaviours throughout the course and how their behaviours relate to their academic performance, thus helping educators identify the types of support required for particular groups of students. When educators are aware of the ways in which students are engaging with course materials and activities over time and how this impacts their academic performance, they can choose how they will adapt their teaching. In some instances, educators may choose to design interventions such as ways in which they will communicate with students who exhibit certain types of trajectories. For example, by using the proposed analytics-based framework, a half way through the semester, educators may wish to identify the groups of students who are less engaged with the course materials and activities and give them a nudge, particularly if it will support their academic performance. Conversely, educators may wish to redesign components of their course materials or activities for the following iteration of the course after being aware that certain groups of students (e.g. the *Disenchanted*), lost interest in engaging with the course materials and activities part way through the course and yet were successful in their summative assessments.

The second type of learning strategy that the proposed analytics-based framework can help detect relates to students' learning session data or, in other words, the sequences of actions students take when accessing course materials and completing formative and summative assessments. This is a two-phase analysis whereby the first phase focuses on clustering student learning sequences to identify the types of strategies students use when completing course tasks, and the second phase reveals clusters of students based on the identified learning strategies. For example, in our case study, four strategies were detected with the second strategy being the most common one where students focused on the summative assessment and had a higher number of incorrect attempts compared to correct ones. The clustering stage then revealed groups of students based on the strategies they opted to use. For example, the largest cluster of students (i.e. the *Selective* students), focused largely on their summative exercises and had limited activity with the course materials. However, their assessment scores were lower than other groups of students, particularly students who were more strategic when engaging with the course materials or intensively applied a variety of learning strategies. This sort of information confirms that the course design (e.g. the videos, readings, and formative activities), when used effectively can lead to greater success in the summative assessments.

Applying the analytics-based framework in a similar manner to any FC context can help educators interrogate their learning design by investigating their students' learning behaviour using various complementary methods. The two instantiations of the framework presented in this chapter illustrate how different combinations of data analysis methods - consistent with the learning theory and informed by the adopted learning design - can offer different but complementary insights into student behaviour. For instance, the first instance of the framework (Section 2.2; Figure 1 (b,c)) allows for more fine-grained strategy detection than the second instance (Section 2.3; Figure 1 (d,e)), and thus was able to identify more nuances in the student behaviour.

The application of the framework in the examined FC course (Section 2.1) has demonstrated that students tend to change their learning strategies over time and turn to less effective strategies. This finding is consistent with previous research findings in self-regulated learning (Bjork, Dunlosky, & Kornell, 2013). It implies that instructors should occasionally, and especially after some important course milestones (e.g. midterm exam), remind their students about the importance of regulating their learning towards effective learning strategies, that is strategies based on active engagement with learning resources (e.g., different forms of formative assessment). Furthermore, it is important to remember that learning strategies are skills that only some students have as innate abilities, while the majority have to develop, and all have to practice to reach and maintain proficiency (Ericsson, Krampe, & Tesch-Romer, 1993; Winne, 2013). The analyses presented in this chapter can inform modifications of the learning design aimed at scaffolding the development of the desired learning strategies. It can also provide grounds for adaptive/personalized feedback (e.g., hints, guidelines) to help students improve their study skills. Such feedback is particularly important in FC settings, since research has shown that students in a FC environment often require more awareness of their learning process than students in more traditional settings (Frederickson, Reed, & Clifford, 2005; Strayer, 2012). Thus, students would benefit from personalized feedback that would make them aware of their learning strategies, and how those strategies compare to the strategies of well performing peers.

References

Anonymous. (2016). Details omitted for double-blind reviewing.

Anonymous. (2017a). Details omitted for double-blind reviewing.

Anonymous. (2017b). Details omitted for double-blind reviewing.

Anonymous. (2017c). Details omitted for double-blind reviewing.

Bannert, M., Reimann, P., & Sonnenberg, C. (2013). Process mining techniques for analysing patterns and strategies in students' self-regulated learning. *Metacognition and Learning, 9*(2), 161–185. http://dx.doi.org/10.1007/s11409-013-9107-6.

Bjork, R. A., Dunlosky, J., & Kornell, N. (2013). Self-Regulated Learning: Beliefs, Techniques, and Illusions. *Annual Review of Psychology*, *64*(1), 417–444. doi:10.1146/annurev-psych-113011-143823

del Valle, R. D., & Duffy, T. M. (2009). Online learning: Learner characteristics and their approaches to managing learning. *Instructional Science, 37*(2), 129–149. doi:10.1007/s11251-007-9039-0

Ericsson, K. A., Krampe, R. T., & Tesch-Romer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review, 100*, 363–406. doi:10.1037//0033-295X.100.3.363

Forsey, M., Low, M., & Glance, D. (2013). Flipping the sociology class: Towards a practice of online pedagogy. *Journal of Sociology, 49*(4), 471–485.

Frederickson, N., Reed, P., & Clifford, V. (2005). Evaluating web-supported learning versus lecture-based teaching: Quantitative and qualitative perspectives. *Higher Education*, *50*, 645–664. doi:10.1007/s10734-004-6370-0

Fryer, J. W., & Elliot, A. J. (2007). Stability and change in achievement goals. *Journal of Educational Psychology*, 99(4), 700–714. http://dx.doi.org/10.1016/j.cedpsych.2010.08.002

Gabadinho, A., Ritschard, G., Müller, N. S., & Studer, M. (2011). Analyzing and visualizing state sequences in R with TraMineR. *Journal of Statistical Software, 40*(4), 1–37. http://dx.doi.org/10.18637/jss.v040.i04

Gašević, D., Dawson, S., Rogers, T., & Gašević, D. (2016). Learning analytics should not promote one size fits all: The effects of course-specific technology use in predicting academic success. *The Internet and Higher Education*, 28, 68–84. http://dx.doi.org/10.1016/j.iheduc.2015.10.002

Gašević, D., Kovanović, V., & Joksimović, S. (2017). Piecing the Learning Analytics Puzzle: A Consolidated Model of a Field of Research and Practice. *Learning: Research and Practice, 3*(1).

Hattie, J., Biggs, J., & Purdie, N. (1996). Effects of learning skills interventions on student learning: A meta-analysis. *Review of Educational Research, 66*, 99–136. https://doi.org/10.3102/00346543066002099

Hill, J. R., & Hannafin, M. J. (1997). Cognitive strategies and learning from the world wide web. *Educational Technology Research and Development, 45*(4), 37–64. http://dx.doi.org/10.1007/BF02299682.

Kim, M. K., Kim, S. M., Khera, O., & Getman, J. (2014). The experience of three flipped classrooms in an urban university: An exploration of design principles. *The Internet and Higher Education, 22*, 37–50. http://dx.doi.org/10.1016/j.iheduc.2014.04.00

Kovanovic, V., Gašević, D., Joksimović, S., Hatala, M., & Adesope, O. (2015). Analytics of communities of inquiry: Effects of learning technology use on cognitive presence in asynchronous online discussions. *The Internet and Higher Education*, *27*, 74–89. http://dx.doi.org/10.1016/j.iheduc.2015.06.002

Lage, M. J., Platt, G. J., & Tregua, M. (2000). Inverting the classroom: A gateway to creating an inclusive learning environment. *The Journal of Economic Education, 31*(1), 30-43.

Lockyer, L., Heathcote, E., & Dawson, S. (2013). Informing pedagogical action: Aligning learning analytics with learning design. *American Behavioral Scientist, 57*(10), 1439–1459. https://doi.org/10.1177/0002764213479367

Lust, G., Elen, J., & Clarebout, G. (2013a). Regulation of tool-use within a blended course: Student differences and performance effects. *Computers & Education, 60*(1), 385–395. http://dx.doi.org/10.1016/j.compedu.2012.09.001

Lust, G., Elen, J., & Clarebout, G. (2013b). Students' tool-use within a web enhanced course: Explanatory mechanisms of students' tool-use pattern. *Computers in Human Behavior*, *29*(5), 2013–2021. http://dx.doi.org/10.1016/j.chb.2013.03.014

McCutcheon, A. L. (1987). Latent class analysis. Quantitative Applications in the Social Sciences Series No. 64. Thousand Oaks, California: Sage Publications. ISBN 0-521-59451-0.

O'Flaherty, J., & Phillips, C. (2015). The use of flipped classrooms in higher education: A scoping review. *The Internet and Higher Education*, 25, 85–95. http://dx.doi.org/10.1016/j.iheduc.2015.02.002

Perera, D., Kay, J., Koprinska, I., Yacef, K., & Zaïane, O. R. (2009). Clustering and sequential pattern mining of online collaborative learning data. *IEEE Transactions on Knowledge and Data Engineering, 21*(6), 759–772. https://doi.org/10.1109/TKDE.2008.138

Pierce, R., & Fox, J. (2012). Instructional design and assessment: Vodcasts and active learning exercises in a "flipped classroom" model of a renal pharmacotherapy module. *American Journal of Pharmaceutical Education, 76*(10), 1–5.

Prober, C., & Khan, S. (2013). Medical education reimagined: A call to action. *Academic Medicine, 88*, 1407–1410.

Rahman, A. A., Aris, B., Rosli, M. S., Mohamed, H., Abdullah, Z., & Zaid, N. M. (2015). Significance of preparedness in flipped classroom. *Advanced Science Letters, 21*(10), 3388–3390. https://doi.org/10.1166/asl.2015.6514

Reimann, P. (2011). Design-Based Research. In L. Markauskaite, P. Freebody, & J. Irwin (Eds.), Methodological Choice and Design (pp. 37–50). Springer Netherlands. Retrieved from http://link.springer.com/chapter/10.1007/978-90-481-8933-5_3

Reimann, P. (2016). Connecting learning analytics with learning research: the role of design-based research. *Learning: Research and Practice, 2*(2), 130–142. https://doi.org/10.1080/23735082.2016.1210198

Reimann, P., Markauskaite, L., & Bannert, M. (2014). e-Research and learning theory: What do sequence and process mining methods contribute? *British Journal of Educational Technology*, *45*(3), 528–540. doi:10.1111/bjet.12146

Richardson, J. T. E. (2004). Methodological issues in questionnaire-based research on student learning in higher education. *Educational Psychology Review*, *16*, 347-358. doi:10.1007/s10648-004-0004-z

Siemens, G., & Gasevic, D. (2012). Guest Editorial-Learning and Knowledge Analytics. *Educational Technology & Society, 15*(3), 1–2.

Strayer, J. F. (2012). How learning in an inverted classroom influences cooperation, innovation and task orientation. *Learning Environments Research, 15*, 171-193.

Studer, M., & Ritschard, G. (2016). What matters in differences between life trajectories: a comparative review of sequence dissimilarity measures. *Journal of the Royal Statistical Society: Series A (Statistics in Society), 179*(2), 481–511. https://doi.org/10.1111/rssa.12125

Ward, J. H. J. (1963). Hierarchical Grouping to Optimize an Objective Function. *Journal of the Americal Statistical Association, 58*(301), 236-244.

Weinstein, C. E., Husman, J., & Dierking, D. R. (2000). Self-regulation interventions with a focus on learning strategies. In P. R. Pintrich, & M. Boekaerts (Eds.), *Handbook on self-regulation* (pp. 727–747). New York: Academic. http://dx.doi.org/10.1016/B978-012109890-2/50051-2

Winne, P. H. (2006). How Software Technologies Can Improve Research on Learning and Bolster School Reform. *Educational Psychologist, 41*(1), 5-17. doi:10.1207/s15326985ep4101_3

Winne, P. H. (2013). Learning strategies, study skills, and self-regulated learning in postsecondary education. InM. B. Paulsen (Ed.), *Higher education: Handbook of theory and research* (pp. 377–403). Netherlands: Springer. http://dx.doi.org/10.1007/978-94-007-5836-0_8.

Winne, P. H., & Hadwin, A. F. (1998). Studying as self-regulated learning. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 277–304). Mahwah, NJ, U.S.: Lawrence Erlbaum Associates Publishers

Winne, P. H., & Jamieson-Noel, D. (2002). Exploring students' calibration of self reports about study tactics and achievement. *Contemporary Educational Psychology, 27*(4), 551–572. http://dx.doi.org/10.1016/S0361-476X(02)00006-1

Winne, P. H., & Jamieson-Noel, D. L. (2003). Self-regulating studying by objectives for learning: Students' reports compared to a model. *Contemporary Educational Psychology, 28*, 259–276. http://dx.doi.org/10.1016/S0361-476X(02)00041-3

Zhou, M., & Winne, P. H. (2012). Modeling academic achievement by self-reported versus traced goal orientation. *Learning and Instruction*, *22*, 413–419. http://dx.doi.org/10.1016/j.learninstruc.2012.03.004