Tracing learning strategies in online learning environments: A learning analytics approach

Wannisa Matcha

Doctor of Philosophy

Institute for Adaptive and Neural Computation
School of Informatics
University of Edinburgh

2020

Wannisa Matcha:

Tracing learning strategies in online learning environments: A learning analytics approach

Doctor of Philosophy, 2020

PRINCIPAL SUPERVISOR:

Professor Dragan Gašević, Ph.D., *The University of Edinburgh, United Kingdom.*

ASSISTANT SUPERVISORS:

Professor Abelardo Pardo, Ph.D., *University of South Australia, Australia.*

Yi-Shan Tsai, Ph.D., *The University of Edinburgh, United Kingdom.*

EXAMINATION COMMITTEE:

Professor Marcus Specht, Ph.D., *Technical University of Delft, Netherlands*

Areti Manataki, Ph.D., *The University of Edinburgh, United Kingdom.*

Dedicated to my father

Abstract

Learning has expanded beyond formal education; yet, students continue to face the challenge of how to effectively direct their learning. Among the processes of learning, the selection and application of learning tactics and strategies are fundamental steps. Learning tactics and strategies have long been considered as key predictors of learning performance. Theoretical models of self-regulated learning (SRL) assert that the choice and use of learning tactics and strategies are influenced by the internal (cognitive) and external (task) conditions. These conditions are consistently updated when students receive internal/external feedback. However, internal feedback generated based on students' evaluation of their own performance against the expectation and/or learning goal is not accurate. Guiding students to apply appropriate learning strategies i.e. providing external feedback, hence, could enhance the students' learning. Recent research literature suggests that learning analytics can be leveraged to support students in the selection and use of effective learning tactics and strategies. However, there has been limited literature on the ways this can be achieved. This thesis aims to fill this gap in the literature.

This thesis begins by exploring the state of the art regarding how students receive learning analytics-based support for the selection and application of learning tactics and strategies. The systematic literature review on this topic reveals that students rarely receive feedback on learning tactics and strategies with learning analytics dashboards. One of the barriers to providing feedback on learning tactics and strategies is the difficulty in detecting learning tactics and strategies that students used when interacting with learning activities. Hence, this thesis proposes a novel analytics-based approach to detect learning tactics and strategies based on digital trace data recorded in learning environments. The proposed analytics-based approach is based on process, sequence mining and clustering techniques. To validate the results of the proposed approach and the credibility of the automatically detected learning tactics and strategies, associations with academic performance and different feedback conditions are explored. To further validate the approach, the efficacy of each proposed approach in the detection of learning tactics and strategies is investigated. In addition, the thesis explores the alignment of the automatically detected learning tactics and strategies with relevant models of SRL. This is done by examining the association between the internal conditions and external conditions. Specifically, internal conditions are represented by the disposition of students based on self-reports of personality traits, whereas external conditions are represented by course instructional designs and delivery modalities. The thesis is concluded with a discussion of the implications of the proposed analytics methodology on research and practice of learning and teaching.

Lay summary

This thesis presents novel approaches for detecting learning tactics and strategies. Relying on the theory of self-regulated learning and approaches to learning, we demonstrate the use of analytics-based approaches, developed based on process mining, sequence mining, and network analytics to capture learning tactics and strategies based on trace data. Drawing on educational theories, we investigate how the detected learning tactics and strategies are associated with learning constructs as defined in the SRL theory. In addition, we demonstrate how the proposed analytics-based approach could be applied across different technology-enhanced learning contexts. This study contributes to the understanding of students' learning process and it proposes a novel approach to automatically detect learning tactics and strategies. This approach can, therefore, be used to inform instructors and learners of learning processes.

Acknowledgment

I would love to take this opportunity to express my great appreciation to those who have immeasurably supported me on this challenging but profound journey. It has been a privilege to work with the greatest minds.

First and foremost, I would like to express the deepest and sincere gratitude to my supervisor, Professor Dragan Gašević, not only for the immense knowledge and ideas he has guided me with, but also for the support, motivation, and understanding during the hardest time of my life. His support is far beyond the supervisory role. I cannot describe how much I appreciate to have Professor Gašević as my mentor.

I also want to thank my assistant supervisors, Professor Abelardo Pardo and Dr. Yi-Shan Tsai, whose constructive feedback, comments, and words of encouragement were essential for the work presented in this thesis. I would also like to offer special thanks to one of our close collaborators, Professor Jelena Jovanović, who shared with me many great ideas which later became the foundation of this research work. Her valuable contributions to the works presented in this thesis are much appreciated.

I was fortunate to have a research partner Nora'ayu Ahmad Uzir. I have no idea how hard my PhD journey would have been if I did not have her company. She made this challenging journey much easier and enjoyable. I would also like to thanks our close collaborators across the globe, Lisa Lim, Jorge Maldonado-Mahauad, Mar Pérez-Sanagustín, and Danijela Gašević. Despite the distance, the collaborations have been fruitful.

Last but not least, I would love to thank my family Surasak, Soimah, Maisaro, Adundet, and Chutima Matcha. These few years had been the hardest journey for us all. Despite the difficulties, you all supported and encouraged me to be here. A "thank you" would not be enough to express how much I appreciated them.

Declaration of authorship

I declare that this thesis has been composed solely by myself and that it has not been submitted, either in whole or in part, in any previous application for a degree. Except where otherwise acknowledged, the work presented is entirely my own. This thesis also includes five peer-reviewed publications produced under the joint authorship:

(1) Matcha, W., Ahmad Uzir, N., Gasevic, D., & Pardo, A. (2020). A Systematic Review of Empirical Studies on Learning Analytics Dashboards: A Self-Regulated Learning Perspective. *IEEE Transactions on Learning Technologies*, *13*(2), 226–245. https://doi.org/10.1109/TLT.2019.2916802

(2) Matcha, W., Gašević, D., Ahmad Uzir, N., Jovanović, J., & Pardo, A. (2019). Analytics of Learning Strategies: Associations with Academic Performance and Feedback, In *Proceedings of the 9th international conference on learning analytics & knowledge*. https://doi.org/10.1145/3303772.3303787

(3) Matcha, W., Gašević, D., Ahmad Uzir, N., Jovanović, J., Pardo, A., Maldonado-Mahauad, J., & Pérez-Sanagustín, M. (2019). Detection of Learning Strategies: A Comparison of Process, Sequence and Network Analytic Approaches, In *European conference on technology enhanced learning*, Springer. https://link.springer.com/chapter/10.1007/978-3-030-29736-7_39

(4) Matcha, W., Gašević, D., Ahmad Uzir, N., Jovanovic, J., Pardo, A., Lim, L., Maldonado-Mahauad, J., Gentili, S., Perez-Sanagustín, M., & Tsai, Y.-S. (2020). Analytics of Learning Strategies: Role of Course Design and Delivery Modality. *Journal of Learning Analytics*

(5) Matcha, W., Gašević, D., Jovanović, J., Ahmad Uzir, N., Oliver, C. W., Murray, A., & Gasevic, D. (2020). Analytics of Learning Strategies: the Association with the Personality Traits, In *Proceedings of the 10th international conference on learning analytics and knowledge (lak '20)*, Frankfurt, Germany, ACM. https://doi.org/10.1145/3375462.3375534

I declare that I substantially contributed to all five publications (i.e., over 50% of the work done) and was involved in all phases of the research process, including study conceptualization, data collection, data analysis and interpretation, as well as the writing of the final publications.

Wannisa Matcha

Table of contents

1 Introduction

All of life is a constant education.

— Eleanor Roosevelt, *The Wisdom Of Eleanor Roosevelt*

TECHNOLOGY has reformed the way we approach learning. Conventional face-to-face learning environments are now embedded with the use of technology, which brings forward new pedagogical models such as flipped learning and blended learning. Not only that, but technology has also enabled asynchronous distance education and non-credited, free online courses such as Massive Open Online Courses (MOOCs). By the end of 2019, there were more than 13,000 MOOCs offered by over 900 universities worldwide. There were more than 110 million learners in total registered in these MOOCs (Shah, 2019). Various cases have also shown that MOOC contents are utilised to transfer traditional classrooms to blended learning and/or flipped classroom (Ghadiri et al., 2013; Pérez-Sanagustín et al., 2017; Rodríguez et al., 2017). Despite the increasing number of courses and learners, research reports that learners are continually facing challenges in regulating their own learning (Carpenter et al., 2020; Kizilcec et al., 2017). For instance, Reich and Ruipérez-Valiente (2019) found that based on six years of data collected from MITx and HarvardX (MOOC platforms offered by MIT and Harvard University), the dropout rates have not changed. That is, less than 10% of registered learners were able to finish these courses. This low number of completion rates highlights the challenge of studying in online learning environments.

Self-regulated learning (SRL) has been identified as an essential set of skills for successful learning in different learning environments (Bjork et al., 2013; DiFrancesca et al., 2016; Winne, 2013). High SRL skills are particularly needed in online learning environments (Abrami et al., 2011). Winne and Hadwin (1998) state that the SRL process involves four cyclical recursive phases, including, phase 1: task definition, phase 2: goal setting and planning, phase 3: enactment of strategies and tactics, and phase 4: adaptation. They propose a model that illustrates key learning constructs in each learning phase including, Conditions, Operations, Products, Evaluate, and Standards (COPES) as presented in Figure 1. In the early phases of learning, students define tasks and set learning goals by considering internal and external conditions. Examples of internal (cognitive) conditions can be knowledge of tasks, domain topic, learning tactics, and different facets of motivation. Ex-

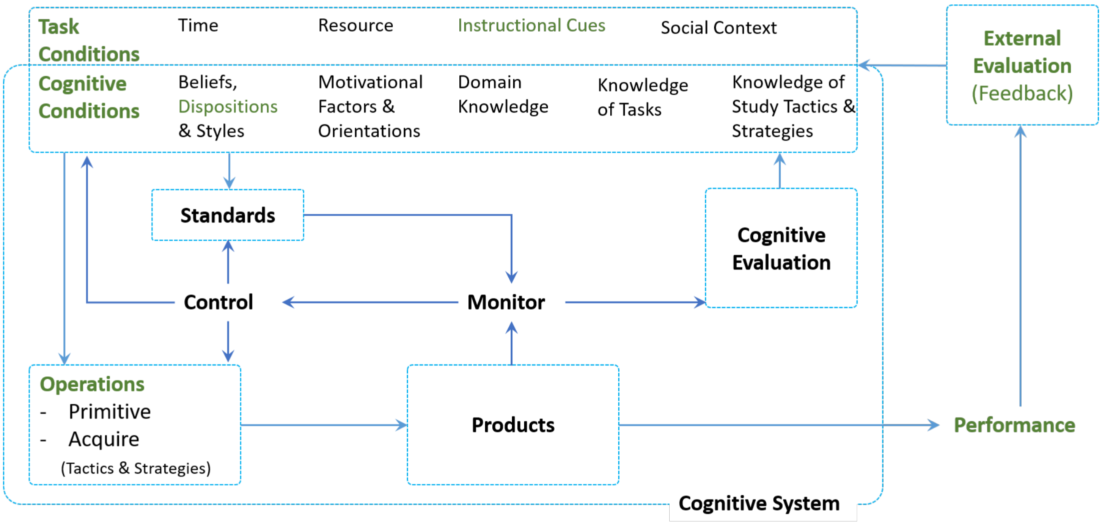


Figure 1. Self-Regulated Learning Process (adopted from Winne & Hadwin, 1998)

ternal (task) conditions can be resources available, learning environment, instructional cues, and time constraints to complete a given task. Based on these conditions, students make judgement to set learning goals, plan their schedule, and set some expectation or 'standards'. Students 'operate' their learning by applying the selected learning strategies and tactics. At the end, the 'products' of learning are created. Students 'evaluate' these products and the choices of learning strategies against the standards that had earlier set. This evaluation can result in the change of learning strategies or updating standards (Winne & Hadwin, 1998). Therefore, the substantial skill in the SRL process is realising and utilising tactics and strategies that can assist an individual in the learning process (Zimmerman, 2011).

Learning strategy is denoted as the factors that predict success in learning (Winne, 2006). Learning strategies are used to facilitate knowledge construction which involves the use of specific study tactics or techniques that may help students perform some specific tasks in their learning process to achieve their learning goals (Derry, 1989; Malmberg et al., 2014; Rachal et al., 2007). Learning tactic and strategy are often used interchangeably. However, these two terms are different. Learning tactic refers to a sequence of actions performed by a student to complete a learning task (Hadwin et al., 2007), whereas, learning strategy is defined as *"a coordinated set of study tactics that are directed by learning goal, and that are aimed at acquiring a new skill or gaining understanding"* (Malmberg et al., 2014, p.116). The literature suggests that not all students apply effective learning methods (Dunlosky et al., 2013; Malmberg et al., 2014; Rachal et al., 2007). Moreover, there are differences in learning strategy application between low and high-performance students (DiFrancesca et al., 2016; Proctor et al., 2006). Students also adopt different learning tactics and strategies when participating in different learning contexts. For instance, a recent study reveals that note-taking is one of the most frequently used learning tactics in the traditional classroom (Dunlosky, 2013). However, current research reports that students rarely use the note-taking tactic when participated in online learning courses (Morehead et al., 2019). Broadbent (2017) found that students reported lower levels of peer learning and help-seeking when participated in online learning as compared to blended learning. These simple cases demonstrate that learners adopted different learning tactics when participating in different learning contexts.

Simply leaving students to manage their own learning does not guarantee effective learning (Kirschner et al., 2006). Students require extensive support and guidance to select and adapt effective learning strategies (Dunlosky et al., 2013; Winne, 2013). However, existing research has found that students rarely receive feedback to guide them in the selection of learning strategies (see Chapter two for a systematic literature review). One of the potential barriers to providing feedback on learning tactics and strategies application is the difficulty in accurately detecting the types of learning tactics and strategies used by students (Matcha, Ahmad Uzir, et al., 2020).

Capturing learning strategy is a difficult task because learning strategy is a “latent construct”, which is sometimes invisible and difficult to observe (Jovanovic et al., 2017). Much research into student learning strategies has relied on self-reports that are collected through questionnaires or think-aloud protocols (Pardo et al., 2017; Winne, 2013, 2014). However, students’ memories about the choices of learning strategies are often biased and incomplete (Broadbent, 2017; Winne, 2013), whereas using think-aloud protocols could impede students’ learning due to cognitive overload (Winne, 2014). Data collected from the digital environment, such as log data, can better show actual behaviour with less discrepancy between perception and actual learning behaviours. For instance, by comparing actual learning activities collected in trace data to self-reports, Hadwin et al. (2007) found that self-reports did not reflect students’ actual behaviour. Moreover, self-reports usually fail to capture the development of learning strategies. The use of learning trace data allows for understanding the students’ actual learning behaviour without intervening in their learning or inadvertently increasing their cognitive overload. Zhou and Winne (2012) found that trace data were better correlated to the students’ learning achievement than self-reports. Still, self-reports are successful in capturing students’ perceptions and intentions and can assist in understanding the choices of actions students make.

Extracting learning tactics and strategies used by students from trace data can empower the instructors to provide timely feedback on students’ learning strategies. In this way, students can have a better idea on the effectiveness of their learning tactics and strategies. Finding the right mechanism to capture learning strategies, communicate the discovered strategies, and recommend the use of effective learning strategies are among the most important steps that should be taken to optimise learning. *This PhD thesis aims to fulfil this gap by exploring data mining techniques and learning analytics-based approaches that can be used to provide insights into how students learn.*

Learning analytics is a growing field of research focusing on utilising the collected data in educational environment to gain insights that may inform and enhance learning practice (Gašević, Kovanović, et al., 2017; Siemens & Baker, 2012). By employing data mining methods, learning tactics and strategies can be identified from trace data (Fincham et al., 2018; Jovanović et al., 2019; Maldonado et al., 2018) that are digital footprints about students' interaction with online resources. However, there is less understanding of what actual learning tactics and strategies are used as captured with the trace data, and how they related to the existing educational theory.

In this PhD thesis, we first explore how learning analytics-based tools are currently used to support learning process. Then, a novel approach for detecting learning tactics and strategies used by learners when they interact with the online learning materials is proposed. The thesis then explores how the learning tactics and strategies discovered with the proposed approach are informed by existing education theories such as approaches to learning and SRL model. In this thesis, the connections of the SRL theory and the adoption of learning tactics and strategies are explored by focusing on the associations of the learning tactics and strategies detected from trace data with the SRL constructs proposed in Winne and Hadwin's (1998) COPES model, including, products, feedback (external evaluation), cognitive conditions (i.e., personality traits), and task conditions (i.e., course designs and delivery modalities).

1.1 Research goals and questions

The work presented in this thesis was conducted with three main research goals. The *first goal* was to examine how learning analytics tools have been used to guide students in their learning process, which is under-explored in the literature (Bodily & Verbert, 2017; Corrin & de Barba, 2015; Gašević et al., 2015; Gayane Sedrakyan et al., 2016). To address this gap, the thesis systematically surveyed how existing learning analytics-based tools had been used to support SRL according to Winne and Hadwin's (1998) COPES model. Moreover, the literature often claims that feedback generated from the learning analytics-based tools supports the SRL process; however, there had been insufficient evidence in the literature to validate this claim (Matcha, Ahmad Uzir, et al., 2020). Hence, to bridge the gap, this thesis presents a systematic literature review of learning analytics-based feedback tools (i.e. learning analytics dashboards–LADs). Thus, the first research question is

Research Question 1:

Based on the SRL perspective, how have learning analytics-based tools been used to support students' learning process?

By addressing the first research goal, we found that students rarely received feedback on their choices of learning tactics and strategies. We posited that a possible reason for this could be attributed to the difficulty of detecting learning tactics and strategies from trace data (Jovanovic et al., 2017). Therefore, the *second goal* of the study was to design and develop a learning analytics-based approach that can automate the detection of theoretically-meaningful learning tactics and strategies from trace data. As such, our second research question is

Research Question 2:

Can we develop a learning analytics-based approach to detect meaningful learning tactics and strategies? To what extent can the proposed approach detect (dis)similar learning tactics and strategies when compared to alternative approaches and across different learning contexts?

The *third goal* of the presented study was to validate the proposed analytics approach for the detection of learning tactics and strategies. According to Joksimovic et al. (2019), the validation of a learning analytics-based approach can be regarded as *"the degree to which theory and evidence support the interpretation of the measurement"* (Joksimovic et al., 2019, p. 48). Hence, the validity of the proposed method could be explored in terms of how the detected learning strategies can be explained by well-accepted learning theories (Gašević, Kovanović, et al., 2017). Approaches to learning (Biggs, 1987; Entwistle, 1991; Marton & Säljö, 1976) is among the frequently referred theoretical models used for explaining learning strategies (Kovanović et al., 2019; Trigwell & Prosser, 1991; Zeegers, 2001). Approaches to learning define three types of learning strategies, namely, deep, strategic (achieving), and surface learning approaches. Association with academic performance is usually examined to further explain learning strategies. As such, our third research question is

Research Question 3:

To what extent can we explain learning tactics and strategies detected by using the analytics-based approach with existing educational theories?

In relation to the SRL process, according to Winne and Hadwin (1998) several learning constructs impact the adoption of learning tactics and strategies, especially during phase 3 (enactment of learning tactics and strategies) of the SRL process, which is the central focus of this thesis. Four learning constructs of the COPES model are examined in this thesis, including products, external evaluation, tasks (i.e., external) conditions, and cognitive (internal) conditions.

- Products: Products refer to the outputs of learning in each phase of Winne and Hadwin's (1998) model of SRL. The most observable proxy of products is the academic performance obtained as a result of how students operate their learning.

- External evaluation (Feedback): Students evaluate their learning products against 'standards' or expectations and learning goals. However, existing research shows that internal evaluation or feedback generated by students' own perception is not accurate (Bjork et al., 2013; Carpenter et al., 2020). Therefore, providing students with external feedback can improve their learning. External feedback can also influence students' choices of learning strategies.

- Task conditions: Task conditions refer to the resources available, time and the constraints inherent to the learning tasks and learning environment (Greene & Azevedo, 2007; Winne & Hadwin, 1998). Tasks are shaped by instructional designs and the mode of delivery (Elen

& Clarebout, 2005), both of which are examined in this thesis regarding their effects on the choices of learning tactics and strategies.

- Cognitive conditions: Cognitive conditions refer to the constraints that are internal to students' cognitive processes such as beliefs, dispositions, and previous knowledge of a learning subject (Winne & Hadwin, 1998). In this study, we explore the effects of internal conditions on the choices of learning strategies through the use of self-reported measures of personality traits.

With this in mind, our fourth research question is

RESEARCH QUESTION 4:

How are the SRL constructs associated with the adoption of learning tactics and strategies?

SRL constructs in this PhD thesis refer to the learning facets that influence the SRL process as conceptualised in the COPES model (Winne & Hadwin, 1998). Specifically, this thesis investigates the relationship of products (academic performance), external evaluation (feedback), and two types of conditions (i.e. instructional design and modalities; and dispositions) with learning strategies.

1.2 Methodology

The studies presented in this thesis were conducted by using the quantitative learning analytics-based approaches and empirical data collected from real-world across three different learning modalities, namely, flipped classroom, blended learning, and MOOC.

The research began by exploring the state of the art of learning analytics-based tools and their applications to support SRL process. The systematic literature review was conducted by focusing on the research papers that reported on the used of learning analytics-based tools to provide feedback to students, i.e. LADs (RQ1). The analysis of the collected research studies was done by investigating which of the SRL elements based on the well-known SRL theory developed by Winne and Hadwin (1998) was present in the design and implementation of these LADs. The results of this systematic literature review show that students hardly ever receive any feedback on their selection of learning tactics and strategies.

One of the associated challenges of this problem is the difficulty in detecting learning tactics and strategies used by students (Jovanovic et al., 2017). This research, therefore, aimed to fill in this gap by proposing a learning analytics-based approach that can be used to capture learning tactics and strategies applied by students from trace data (RQ2). We proposed the learning analytics-based approach for the detection of learning tactics based on a combination of (i) a process mining technique that is rooted in first-order Markov models (FOMM), and (ii) the Expectation-Maximization (EM) algorithm. The detected tactics were then explored by using both process mining and sequence mining techniques. Learning strategies were detected based on the pattern of how students applied learning tactics over time by using a clustering technique, i.e. agglomerative hierarchical clustering.

Since the proposed learning analytics-based approach relied on an unsupervised machine learning algorithm, the validity of the analytics-based approach is examined in terms of how the detected learning tactics and strategies are corresponded to the well-established learning theories (Gašević, Kovanović, et al., 2017; Joksimovic et al., 2019; Jovanovic et al., 2017). Therefore, we explored how the detected learning tactics and strategies can be supported by relevant learning theories (RQ3 and RQ4). The results of applying the analytics-based approach, i.e. detected learning tactics and strategies are interpreted according to the theory of approaches to learning (RQ3). We framed our research under the SRL model proposed by Winne and Hadwin (1998). Hence, the association with the learning constructs were explored (RQ4). Inferential statistics were used to investigate the relationship between the detected learning strategies and SRL constructs.

1.3 Thesis structure and overview

In order to address the four research questions, the overall thesis is structured based on the consolidated model of learning analytics as proposed by Gašević, Kovanović, et al. (2017). The model highlights the three core dimensions of learning analytics, including theory, design, and data science. The thesis is made of seven chapters, and the middle five chapters correspond to one or more dimensions of the learning analytics model as shown in Figure 2. Each of the five chapters addresses one or more research questions (Table 1), and includes one peer-reviewed publication which constitutes the core of the chapter. A short preface and summary are included in each chapter to explain how the included publication fits into the overall structure and topic of the thesis. The final chapter provides a summary of the findings, highlights the implications of the research and suggests the direction for future work.

In the remainder of this section, a brief overview of each chapter and how they contribute to the research goals of the thesis as defined in Section 1.1 are described.

Table 1. Overview of the thesis research questions by individual chapters.

		Research questions			
Chapter	**Title**	RQ 1	RQ 2	RQ 3	RQ 4
Chapter 2	Support for Learning Strategies by Learning Analytics-based Tools	✓			
Chapter 3	Analytics of Learning Strategies: Automatic Detection of Learning Tactics and Strategies		✓	✓	
Chapter 4	Analytics of Learning Strategy: Associations with Academic Performance and Feedback			✓	✓
Chapter 5	Analytics of Learning Strategy: Role of Course Design and Delivery Modalities		✓	✓	✓
Chapter 6	Analytics of Learning Strategy: Associations with Personality Traits			✓	✓

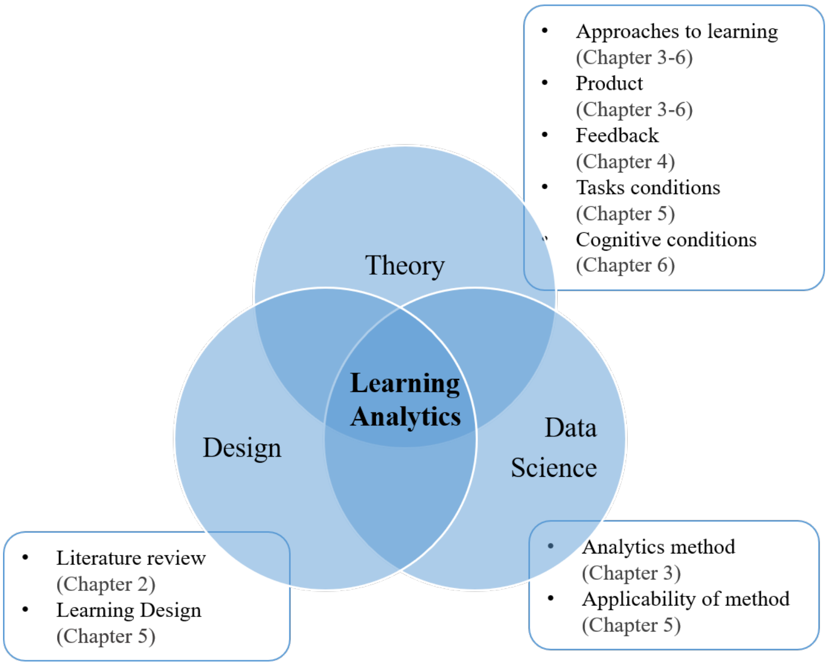


Figure 2. Graphical structure of the thesis

1.3.1 Overview of chapter two: "Support for Learning Strategies by Learning Analytics-based Tools" (RQ 1)

Learning analytics offers insights into the learning process. Research employs learning analytics to enhance learning by offering the results of the analysed data about students' activities as feedback. Communication of learning analytics-based feedback is typically done by using LADs. Nonetheless, how LADs support SRL is unclear. In order to address this research problem, we examine how LADs have been used to support the self-regulation process.

Research contributions:

- We conducted a systematic literature review of studies that reported on empirical uses of LADs to support SRL. Our investigation was based on the model of SRL developed by Winne and Hadwin (1998).
- The systematic literature review provides a great level of details on the research on LADs. It also reports on the limitations of existing research and identifies future research directions.
- The results of this review highlight the oversight of using LADs to support the selection of learning strategies. Moreover, existing LADs do not support metacognition, and are not grounded in learning theories.

Research output:

1. Matcha, W., Ahmad Uzir, N., Gasevic, D. Pardo, A.: *"A Systematic Review of Empirical Studies on Learning Analytics Dashboards: A Self-Regulated Learning Perspective"*: A journal article reporting important issues in research, design and development of LADs. This article has been accepted for publication in the IEEE Transaction in Learning Technologies.

1.3.2 Overview of chapter three: "Analytics of Learning Strategies: Automatic Detection of Learning Tactics and Strategies" (RQs 2 and 3)

Feedback is an important process in supporting learners to adjust and select effective learning tactics and strategies (Greene & Azevedo, 2007; Hattie & Timperley, 2007). As such, capturing learning tactics and strategies used during the learning process is an important step to enable such feedback. Traditional research relies on self-report instruments to detect learning tactics and strategies used by students. However, students are not always accurate in reporting on how they learn (Bjork et al., 2013; Broadbent, 2017; Carpenter et al., 2020). Simply relying on self-reports is not sufficient in terms of presenting the "dynamic" application of learning tactics and strategies over time.

Recent research shows that trace data collected from a digital learning environment is more suitable than self-reports to describe learning constructs (Zhou & Winne, 2012). However, research into the detection of learning tactics and strategies from trace data is scarce. Moreover, most of the learning analytics-based approaches are context dependent (Jovanovic et al., 2017; Maldonado-Mahauad et al., 2018), and their validity and generalisability have not been sufficiently studied. It is also understudied how learning tactics and strategies detected with learning analytics-based approaches correspond to educational theories. To fulfil this gap, this chapter reports the findings of the study that explored (dis)similarity in detecting learning tactics and strategies with different learning analytics-based approaches when applied to the same dataset.

Research contributions:

- We proposed and evaluated the use of learning analytics-based approaches for the identification of different learning tactics and strategies by considering students' learning actions when interacting with online learning environments. Three main approaches, which were based on sequential, network, and process analytics, were explored.
- Using the data collected for a fully-online credited course, we applied the three analytics-based approaches to identify corresponding learning tactics and strategies used by the students and examined their association with academic performance.
- Our results revealed that the three learning analytics-based approaches were able to detect similar learning tactics and strategies to some extent. The choice of data analytics approaches has a direct implication on the analysed learning tactics and consequently impact on the detection of the learning strategies.

Research output:

1. Matcha, W., Gašević, D., Ahmad Uzir, N., Jovanović, J., Pardo, A., Maldonado-Mahauad, J. Pérez-Sanagustín, M.: *"Detection of Learning Strategies: A Comparison of Process, Sequence and Network Analytic Approaches"*: A full conference paper reporting on the comparison of three analytics based methods for the detection of students' learning tactics and strategies. The paper was presented at the 14th European Conference on Technology Enhanced Learning *(EC-TEL 2019)*.

1.3.3 Overview of chapter four: "Analytics of Learning Strategy: Associations with Academic Performance and Feedback" (RQs 2, 3 and 4)

According to Winne and Hadwin's (1998) COPES model, external evaluation, i.e. feedback provided by instructors is an important learning construct. However, providing feedback to a large cohort class is a challenging task. Learning analytics has recently been used to provide an approach for the generation of semi-automated and personalised messages that communicate feedback from the instructor to students (Pardo et al., 2017). In this chapter, we explore the role of feedback and its association with learning strategies detected from trace data using the process-based approach explained in Chapter three. The corresponding association between learning strategies and academic performance is also investigated.

Research contributions:

- We evaluated the approach based on the process mining approach for the detection of learning tactics and strategies. We examined the detected tactics and strategies in terms of their temporal and sequential characteristics.
- We also investigated the relationship between the detected learning strategies, academic performance and external evaluation (feedback). The detected strategies were found to correspond to those conceptualised in the theory of approaches to learning.
- We demonstrated that data-driven learning tactics and strategies could provide insights into how students actually learned and the actual learning strategies that they applied.
- We showed that the use of personalised learning analytics-based feedback was associated with the increase in the use of effective learning strategies and academic performance.

Research output:

1. Matcha, W., Gašević, D., Ahmad Uzir, N., Jovanović, J. Pardo, A.: *"Analytics of Learning Strategies: Associations with Academic Performance and Feedback"*: A full conference paper describing the proposed method used to detect learning tactics and strategies and exploring the association of the detected strategies with academic performance and feedback. The paper was presented at the Ninth International Learning Analytics and Knowledge Conference *(LAK'19)* and was shortlisted for the best full research paper award.

1.3.4 Overview of chapter five: "Analytics of Learning Strategy: Role of Course Design and Delivery Modalities" (RQs 2, 3 and 4)

In order to validate the proposed process analytics-based approach, we examined how the approach can be applied across different learning contexts (i.e. flipped classroom, blended learning, and MOOC). The learning contexts are shaped by the course instructional designs and delivery modalities, which affect task conditions as theorised in Winne and Hadwin's (1998) COPES model. Task conditions refer to the external factors that may influence the selection of learning tactics and strategies. Task conditions are shaped by the time constraints, resource available, instructional cues and social context. These constructs are driven by the design of a course and course delivery modalities. In this chapter, we examined how the detected learning tactics and strategies reflect the course instructional design and delivery modalities.

Research contributions:

- We replicated the use of the proposed process analytics-based approach across three different learning contexts, in order to examine the role of course design and delivery modality on learning tactics and strategies. The three learning contexts were different in terms of course instructional design (problem-solving and practice-based learning) and delivery modalities (flipped classroom, blended learning, and MOOC).
- The results revealed that the choices of learning tactics were influenced by course instructional designs, and learning strategies were sensitive to delivery modalities.
- This study provided robust evidence that the proposed analytics approach developed based on process mining is applicable across different learning contexts. The analysed learning strategies are corroborated with the findings of learning strategies detected by using traditional methods such as self-report instruments (Byrne et al., 2010; Chonkar et al., 2018; Mattick et al., 2004).

Research output:

1. Matcha, W., Gašević, D., Ahmad Uzir, N., Jovanovic, J., Pardo, A., Lim, L., Maldonado-Mahauad, J., Gentili, S., Perez-Sanagustín, M. Tsai, Y.-S.: *"Analytics of Learning Strategies: Role of Course Design and Delivery Modality"*: An invited journal article describing the role of course designs and delivery modalities in the detection of learning tactics and strategies with the proposed learning analytics approach. The paper is accepted to be published in the Journal of Learning Analytics.

1.3.5 Overview of chapter six: "Analytics of Learning Strategy: Associations with Personality Traits" (RQs 2, 3 and 4)

According to Winne and Hadwin (1998), cognitive conditions include beliefs, dispositions and styles; motivational factors and orientations; domain knowledge; knowledge of task; and knowledge of study tactics and strategies. Personality is an individual difference that can explain the disposition of one's learning behaviours (Bidjerano & Dai, 2007). Personality is defined as *"relatively enduring patterns of thoughts, feelings, and behaviours that distinguish individuals from one another"* (Roberts & Mroczek, 2008, p. 31). Numerous research studies have been carried out to explore personality and its relationships with academic performance, learning strategies, and learning processes (Bidjerano & Dai, 2007; Chamorro-Premuzic & Furnham, 2008; Conard, 2006). Research in psychology (John & Srivastava, 1999) identifies five personality traits, namely, extraversion, agreeableness, conscientiousness, emotional instability (or neuroticism), and openness to experience (or intellect) (Bidjerano & Dai, 2007; John & Srivastava, 1999; Roberts & Mroczek, 2008).

Extraversion is characterised as being a sociable and talkative person. Agreeableness refers to the generosity. Conscientiousness represents a dependable and organised individual. Emotional instability tends to be anxiety. Openness is described as being creative and willing to expose oneself to a new experience. Using self-reported measures of personality and learning strategies, existing research has found that conscientiousness and openness are predictive of academic performance and the choice of learning strategies (Chamorro-Premuzic & Furnham, 2008; Conard, 2006). However, up to now, there has been no research study that explored the association of personality traits and the learning strategies detected by using an data analytics-based approach.

Research contributions:

- We investigated the relationship between learning strategies detected from trace data and personality traits as reported by students who participated in a MOOC.
- Our results confirmed the association between learning strategies and personality traits as it has previously been observed by studies using self-reported methods. In particular, we observed that conscientiousness and agreeableness were associated with effective learning strategies (i.e., those that promote deep approaches to learning) and were predictive of academic performance. Emotional instability is associated with the adoption of the surface approaches to learning and was predictive of lower academic performance.
- The results of this chapter offer evidence about the reliability and validity of the proposed learning analytics-based approach for the detection of learning strategies by confirming similar findings already reported elsewhere in the literature that used conventional self-reported research methods.

Research output:

1. Matcha, W., Gašević, D., Jovanović, J., Ahmad Uzir, N., Oliver, C. W., Murray, A. Gasevic, D.: *"Analytics of Learning Strategies: the Association with the Personality Traits"*: A full conference paper reporting the association between learning strategies and personality traits variables. The paper was presented at the 10th International Conference on Learning Analytics and Knowledge *(LAK'20)* and shortlisted for the best paper award.

1.3.6 Overview of chapter seven: "Conclusions and future directions"

In the final chapter, the impact of the thesis with respect to the research questions identified in Chapter one is discussed. The implications for research and practice are highlighted as well as potential directions for future research. The thesis is concluded with a summary of key contributions.

2 Support for Learning Strategies by Learning Analytics-based Tools

Sometimes all it takes is a tiny shift of perspective to see something familiar in a totally new light.

— Dan Brown, *The Lost Symbol*

2.1 Introduction

LEARNING Analytics is a growing field of research that exploits the data collected about learners to optimise the learning process. Learning analytics is defined as *"the measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimising learning and the environments in which it occurs"* (Siemens & Baker, 2012, p. 252-253). Gašević, Kovanović, et al. (2017) posit that in order for the learning analytics research to achieve its promising goal, research in this area needs to reinforce three core components (referred to as the 'consolidated model of learning analytics'), including theory, design and data science when conducting the research (Figure 3). Learning analytics research highly emphasises on the application of data science methods to extract insights from data, thereby generating feedback for relevant stakeholders. However, research in data-driven fields such as learning analytics often comes with challenges. The literature reveals that learning analytics research suffers from a lack of educational theories to inform practices (Gašević, Kovanović, et al., 2017; Reimann, 2016; Wise & Shaffer, 2015). The critical absence of theory may have hidden the promising benefits gained by using learning analytics-based methods and tools.

Feedback is recognised as a key mechanism to facilitate learning (Hattie & Timperley, 2007; Van der Kleij et al., 2015). Current research in learning analytics has widely used LADs as a means to communicate feedback to stakeholders (Schwendimann et al., 2016; Sedrakyan et al., 2016). LADs visualise the information obtained from the analysis of digital trace data and present the results to stakeholders based on an identified set of "indicators" (Schwendimann et al., 2016). One of the primary aims of LADs is to support the SRL process (Bodily & Verbert, 2017; Jivet et al., 2017).

SRL is referred to a learning process that involves the use of cognition, metacognition, affect,

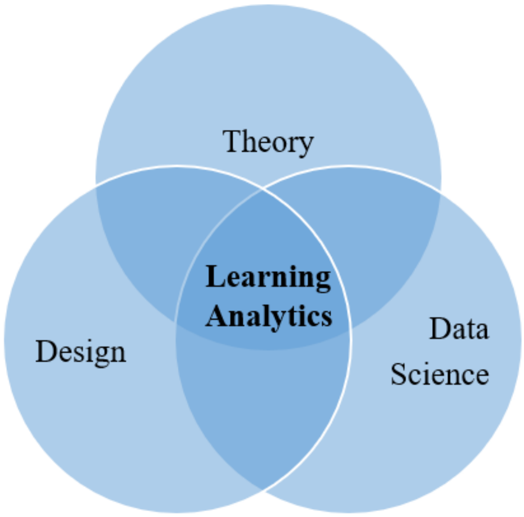


Figure 3. Consolidated model of learning analytics proposed by Gašević, Kovanović, et al. (2017)

and motivation to direct and manage learning to achieve a set of learning goals (Panadero, 2017; Schunk, 2008; Winne & Hadwin, 1998; Zimmerman, 2011). In Winne and Hadwin's (1998) COPES model, five key learning constructs recursively influence the SRL process, including:

- Conditions: refer to the constraints that could restrict how students engage with learning. The conditions could be internal or external to the cognitive process. External (task) conditions include the time available, learning resources provided to students, instructional cues, and social context. The internal (cognitive) conditions involve
  - Beliefs, disposition, and styles: They refer to students' perceptions of their ability to accomplish the learning tasks. The students' beliefs, disposition, and styles closely link to the motivation and consequently influence the goals and overall learning process.
  - Motivational factors and orientation: They are the internal factors that drive intentions and goals of students, thus, influencing how students choose to proceed with their learning. For example, a student with intrinsic motivation usually aims to broaden the understanding and mastery of the skills. Therefore, during phase 3: enactment of strategies, the student tends to be highly engaged in learning and applies effective learning tactics such as cognitive elaboration (Greene & Azevedo, 2007).
  - Domain knowledge: This refers to the long-term memory, or knowledge students previously had acquired relevant to the topic at hand.
  - Knowledge of tasks: This refers to the understanding of the steps students need to take in order to accomplish the tasks they are asked to work on.
  - Knowledge of study tactics and strategies: This refers to the understanding and knowledge of how to apply study tactics and strategies. Even though learning tactics and strategies are often used interchangeably, the literature clearly differentiate the mean-

ings of these two constructs. That is, a learning tactic is a sequence of actions operated by students to accomplish a task (Hadwin et al., 2007; Malmberg et al., 2014). A learning strategy is a purposeful use of one or more learning tactics (Derry, 1989; Malmberg et al., 2014; Rachal et al., 2007). In short, a learning tactic is defined at the task level, whereas, the learning strategy is broader (Malmberg et al., 2014), and is defined at the level of an overall learning goal. A learning strategy develops over time and becomes the characteristic of one's learning (Winne et al., 2002).

- Operations: refer to the actual actions performed by a student to process information. Operations can be defined as
  - Primitive: refer to the primitive actions performed by a student to recall the prior knowledge, searching for information, connect new information and translate it into knowledge (Winne & Hadwin, 1998).
  - Acquired (tactics and strategies): students develop a set of skills when and how to apply learning tactics and strategies. These skills influence the knowledge of learning tactics and strategies, as well as the self-efficacy and beliefs of the students about tactics and strategies.
- Products: refer to the results or artefacts created by students during each phase of the SRL process (Greene & Azevedo, 2007; Winne & Hadwin, 1998). Each phase of the SRL results in different type of products. For instance, the results of applying learning tactics and strategies in phase 3 can produce higher memory recall.
- Evaluation: students evaluate products against learning goals and/or standards (i.e., expectations) (Winne & Hadwin, 1998). Evaluation can take place internally within or externally to a cognitive system (Pardo et al., 2017). Internal evaluation refers to the process where a student makes a judgement on their own learning. The external evaluation feedback on student's learning by external agents such as instructor, teacher, or automatically generated by a learning system (Pardo et al., 2017). Based on the evaluation, students update the conditions and standards, which in turn influences the adoption of specific types of learning tactics and strategies.
- Standards: are beliefs and the criteria of success set by a student in each phase of SRL (Greene & Azevedo, 2007). For example, during phase 3 of Winne and Hadwin's model (enactment of learning tactics and strategies), the standards could be i) the time that a student expects to be engaged in learning and ii) the knowledge and understanding that the student intends to gain in each learning session.

This theoretical model of SRL highlights the important role of feedback as a driving force that influences cognitive conditions, operations, goals, and standards (refer to Figure 1 in Chapter one) (Winne

& Hadwin, 1998) . Effective feedback should, therefore, help students 'operate' their learning in effective manners and adopt suitable learning tactics and strategies to maximise the quality of their learning products. However, how LADs facilitate and impact self-regulation is under-explored (Gašević et al., 2015; Jivet et al., 2017). Moreover, it is unclear how existing LADs have been developed and to what extent they are grounded by educational theories (Gašević et al., 2015; Jivet et al., 2017). The insufficiency of theoretical support can lead to ineffective feedback and consequently unsuccessful learning (Wise & Shaffer, 2015).

The main objective of this chapter is to review the state of the art of research in LADs in terms of their use in feedback provision to support the learning process. Principally, based on the consolidated model of learning analytics research (Gašević, Kovanović, et al., 2017), this chapter explores:

- Theory – centred around the well-known SRL model by Winne and Hadwin (1998), the chapter investigates how current learning analytics tools have been used to provide support on the learning process based on the COPES model.
- Design – several dimensions of the design and development of LADs are explored along with the study designs used for evaluations of the LADs.
- Data science – the role of data science is explored in terms of data sources used to generate feedback.

2.1.1 Chapter overview

A systematic literature review is a rigorous research method, used to provide an overview of a research topic (Sutherland, 2004). Focusing on clear research questions, a systematic literature review relies on a pre-defined and clear protocol of how to conduct the review (Armstrong et al., 2011; Moher et al., 2016). The following steps of the systematic literature review need to be defined, documented, and planned before carrying out the review, including, systematic search of relevant research studies, selection of the relevant studies, critical evaluation of the studies, and a synthesis of the findings to provide robust evidence to research questions set (Armstrong et al., 2011; Sutherland, 2004).

- Systematic search strategy: during this initial step, explicit keywords relevant to the topic of the review need to be clearly stated (Sutherland, 2004). These keywords are then transparently used to search for relevant studies across the pre-defined databases.
- Selection: the criteria of inclusion and exclusion of the literature need to be clearly defined. Specifying the inclusion and exclusion criteria allows the researchers to obtain the literature on the same basis (Jahan et al., 2016). It is suggested that the selection of the literature should be done by at least two independent reviewers and discussed to resolve any conflicts if existed (Jahan et al., 2016).

- Critical evaluation: by considering the research questions, the information obtained from the relevant literature needs to be specified. The information is structured in form of a table to ensure the consistency with respect to the review questions (Kitchenham & Charters, 2007). Moreover, the quality of the literature should also be evaluated.

- Synthesis: the findings should be synthesised by providing the qualitative and/or quantitative evidence to address the research questions.

As such, a systematic literature review was chosen to investigate the research question one (RQ1) in Section 1.1, that is, how learning analytics tools such as LADs have been used to support the SRL process.

The systematic literature review presented in this chapter began with keyword searching to obtain relevant research studies that report the use of LADs in five main academic databases, including ACM Digital library, IEEE, Springerlink, ScienceDirect, and Wiley. The first 100 results from Google scholar were also included as additional sources of information. The second step was the selection of the studies by excluding irrelevant papers, short papers, and those that did not report the results in English. Finally, the coders manually read through the papers. At the end, only those papers that reported on empirical studies of LADs and were presented in English were included in this study. As a result, 29 papers were included in the final analysis. Two independent coders manually coded the papers according to three main research questions, including:

- What is the support of LADs for the elements of SRL as established in learning sciences?

- What type of information is offered as feedback in LADs?

- What is the quality of study designs and reporting that discussed the empirical evaluations of the LADs?

A significant contribution of the present work is that it provides a comprehensive overview of contemporary research on LADs which potentially can be used to inform the design and development of LADs. It provides insights by synthesizing the general practice of the design and development of the LADs including the selection of the indicators to present the results, the reference frames used, the main purposes the LADs were designed for, the types of visualisation used, target users, and the quality of the study designs. The review also i) discusses the oversights and limitations in LADs research that require attention from educators and researchers and ii) suggests future research directions to address the highlighted limitations.

2.2 Publication: A Systematic Review of Empirical Studies on Learning Analytics Dashboards: A Self-Regulated Learning Perspective

The following section includes the verbatim copy of the following publication:

Matcha, W., Ahmad Uzir, N., Gasevic, D., & Pardo, A. (2020). A Systematic Review of Empirical Studies on Learning Analytics Dashboards: A Self-Regulated Learning Perspective. *IEEE Transactions on Learning Technologies*, *13*(2), 226–245. https://doi.org/10.1109/TLT.2019.2916802

A Systematic Review of Empirical Studies on Learning Analytics Dashboards: A Self-Regulated Learning Perspective

Wannisa Matcha, Nora'ayu Ahmad Uzir, Dragan Gašević, and Abelardo Pardo

***Abstract*—This paper presents a systematic literature review of learning analytics dashboards (LADs) research that reports empirical findings to assess the impact on learning and teaching. Several previous literature reviews identified self-regulated learning as a primary focus of LADs. However, there has been much less understanding how learning analytics are grounded in the literature on self-regulated learning and how self-regulated learning is supported. To address this limitation, this review analyzed the existing empirical studies on LADs based on the well-known model of self-regulated learning proposed by Winne and Hadwin. The results show that existing LADs are rarely grounded in learning theory, cannot be suggested to support metacognition, do not offer any information about effective learning tactics and strategies, and have significant limitations in how their evaluation is conducted and reported. Based on the findings of the study and through the synthesis of the literature, the paper proposes that future research and development should not make any *a priori* design decisions about representation of data and analytic results in learning analytics systems such as LADs. To formalize this proposal, the paper defines the model for user-centered learning analytics systems (MULAS). The MULAS consists of the four dimensions that are cyclically and recursively interconnected including: theory, design, feedback, and evaluation.**

***Index Terms*—Dashboards, empirical research, feedback, information visualization, learning analytics, self-regulated learning.**

I. Introduction

OVER last several years, the role of technology increased significantly in different educational settings from the widespread use of learning management systems to social media, interactive simulations, and serious games to name a few. The growth in the use of technology propelled the development of the capacity for capturing data about various aspects of learning experiences. This is done through the collection of digital footprints learners leave behind whenever they interact with technology. These digital footprints have been recognized as a promising source of data (also known as trace of log data) that can be leveraged to inform and optimize decision making of a wide range of stakeholders such as learners, teachers, and administrators.

To harness the potential of digital footprints, the field of learning analytics focuses on the collection, analysis, and reporting of data about learners and contexts in which learning occurs [1]. Learning analytics make use of data science methods to analyze data and report the results of the analysis with different visual and textual approaches [2]. Within learning analytics, dashboards have received much attention as tools that can provide users with relevant insight, prompt user reflection, and potentially inform interventions that are aimed at optimizing learning and the quality of the student experience [3]. Schwendimann and colleagues [4] define (LAD) as "a single display that aggregate different indicators about learner(s), learning process(es) and/or learning context(s) into one or multiple visualization" (p. 8).

In spite of some promising results, critical limitations in the existing research and design of LADs have been identified [5]. Some studies suggest that learners find it hard to interpret the data presented in dashboards and to make use of feedback presented in dashboards to inform the choices of their learning strategies [6]. The actual impact of learning dashboards and recommendation systems is found to be relatively low [7]. Moreover, some authors question whether feedback presented in LADs could be translated into a meaningful actionable recommendation to guide students in their learning [5], [8], [9].

Given the growing interest in LADs, several systematic literature reviews have recently been published (see Section II.A for a summary). These reviews identified key themes that emerge in the literature including the focus on metacognitive, cognitive, affective, and behavioral aspects of learning [10], [11]. The focus on metacognition and self-regulated learning has particularly been emphasized. The conclusions of the existing literature reviews suggest that current LADs should i) have theoretical grounding to overcome some of the limitations in the existing LADs [10], [11], ii) support all phases of self-regulated learning

Manuscript received August 12, 2018; revised April 22, 2019; accepted April 26, 2019. Date of publication May 13, 2019; date of current version June 18, 2020.. This work was supported in part by the LALA Project under Grant 586120-EPP-1-2017-1-ES-EPPKA2-CBHE-JP from the European Commission (EC), in part by EMBED Project under Grant 2017-1-NL01-KA203-035288 from the EC, in part by Smart-Data Schools Discovery Project under Grant DP190102286 from the Australian Research Council (ARC). This paper reflects the views only of the authors, and the EC and ARC are not responsible for any use which may be made of the information contained therein. *(Wannisa Matcha and Nora'ayu Ahmad Uzir contributed equally to this work.) (Corresponding author: Dragan Gašević.)*

W. Matcha and N. Ahmad Uzir are with the School of Informatics, University of Edinburgh, Edinburgh EH8 9AB, U.K (e-mail: w.matcha@ed.ac.uk; n.uzir@ed.ac.uk).

D. Gašević is with the Faculty of Information Technology, Monash University, Clayton, VIC 3800, Australia and also with the School of Informatics, University of Edinburgh, Edinburgh EH8 9AB, U.K (e-mail: dragan.gasevic@monash.edu).

A. Pardo is with the Division of Information Technology, Engineering and Environment, University of South Australia, Adelaide, SA 5000, Australia. (e-mail: abelardo.pardo@unisa.edu.au).

Digital Object Identifier 10.1109/TLT.2019.2916802

[10], iii) significantly improve how evaluation is conducted [4], [10]–[12], and iv) establish closer connections with the literature on open learner modeling [13].

Although the limitations are reported in the current literature reviews, there have been a shortage of systematic analyses of the existing LADs based on an established theoretical model of (self-regulated) learning. This is a significant limitation, as some of the findings in the existing reviews may not be completely aligned with the contemporary understanding of self-regulated learning in the learning sciences. For example, metacognition in the reviews of LADs is only linked to the reflection phase on self-regulated learning [10]. This exclusive link is not consistent with the Winne and Hadwin [14] model of self-regulated learning. The key reason for the misalignment is the fact that metacognition is not a phase of self-regulated learning. Instead, metacognition is exercised by two key processes – monitoring and control. These two processes underline the engagement with all four phases of the Winne and Hadwin model (see Section II.B for details). Therefore, to be able to draw conclusions about the extent to which LADs support metacognition and self-regulated learning, there is a need to perform a systematic literature review of LADs against in a well-established theoretical model. This motivated our first research question:

1. What is the support of LADs for the elements of self-regulated learning based on a well-established theoretical model?

In this review, the analysis of the LADs reported in the literature was performed based on the Winne and Hadwin [14] model of self-regulated learning. The Winne and Hadwin model is particularly suitable for the analysis of LADs, as the model is derived from the well-known synthesis of the literature on learning feedback [15]. Our results of the literature show that existing LADs i) are rarely grounded in learning theory; ii) cannot be suggested to support metacognition; iii) do not offer any information about effective learning tactics and strategies; and iv) have significant limitations in their evaluation. Note that, in this systematic literature review, the effectiveness of individual SRL element on the learning outcome were not discussed.

In addition to the theoretical analysis of LADs published to date, the current study also aimed to replicate and complement some of the findings published in the previous literature reviews through two additional research questions.

2. What types of information is offered as feedback in LADs?

This research question corroborated the findings of the previous literature reviews [4] about the techniques used for presentation of data and results of analysis in LADs. The findings showed that individual references frames were most prevalent in the existing dashboards followed by comparisons with group average scores. The use of chart bars is the most relevant.

3. What is the quality of study designs and reporting the literature that reports on empirical evaluations of LADs?

The final research question expanded the existing literature reviews by critically appraising the quality of design and reporting of empirical research on LADs. This appraisal was performed by using an instrument adapted from medical research. The findings suggest that no studies discuss generalizability of their findings, while limitations of study results are rarely reported. Causal effects of the use of LADs can hardly be made due to the limited used of experimental designs and mixed-methods.

To guide future work of developers, researchers, and adopters, a model for LADs has been proposed in this paper (Section V). The model integrates the findings of these three research questions.

II. Background

This section summarizes the findings of the previous reviews of the literature on LADs. The section also introduces the Winne and Hadwin model of self-regulated learning.

A. *Reviews of LAD Literature*

The initial review of the 15 LADs was carried out by Verbert *et al.* with the aim to illustrate the conceptual framework proposed by the authors of that review. The review analyzed target users of dashboards, data that were tracked, and evaluations performed. This review was further extended by Verbert and her colleagues [16] who ranked existing papers based on categories of LADs that had been deployed in face-to-face lectures, face-to-face group work, and blended learning settings. Then, they analyzed the dashboards in terms of the data sources, data tracking, target users, devices used and evaluation to support the four elements of the conceptual model (awareness, reflection, sense making, and behavioral change) originally proposed by Verbert *et al.* [17]. Although these two reviews provided useful categorizations of the literature, the two reviews did not perform a systematic search of the literature as a guarantee for a comprehensive representation of the state-of-the-art.

The first systematic literature review of learning dashboard research is reported by Schwendimann *et al.* [4] and included a total number of 55 papers. The review presents the result based on four categories: types of contribution (e.g., theoretical proposal or framework), learning context (e.g., target users and learning scenarios), learning dashboard solution (e.g., purpose and data sources), and evaluation. Although it has some similarities with the review reported by Verbert *et al.* [16], the Schwendimann *et al.* review additionally scrutinized the types of indicators presented in individual dashboards into six broad groups and categorized the types of visualization used in LADs. The main finding of the study was that existing papers on LADs rarely reported on results of empirical evaluations, because dashboards were mainly developed as part of exploratory work and built as proof-of-concepts.

The systematic literature review conducted by Bodily & Verbert [12] focuses on student facing learning analytics reporting systems. Such reporting systems include LADs but can also include recommender systems and textual messages with feedback generated based on learning analytics. Building on the previous three literature reviews, 94 papers were included into the review and coded according to five dimensions: functionality, data sources, design analysis, perceived effects and actual effects. The review concluded that further research is needed on the process of design of LADs and recommender systems and not only on the final products of design. The review also suggested that more rigorous experimental studies are needed to determine effects of LADs and recommender systems. Although these suggestions are of critical importance, the review did not analyze the relevance of learning theory and the role of learning sciences in the design and evaluation of LADs.

The systematic reviews conducted by Jivet and her colleagues [10], [11] provide important steps towards bridging the gap between the literature on LADs and the learning sciences. Specifically, Jivet *et al.* analyzed how theories and models that have been integrated at learner-facing LADs. They found six clusters of papers based on their general theoretical tendencies, including, cognitivism, constructivism, humanism, descriptive models, instructional design, and psychology. Of these, the cognitivist cluster had the highest number of papers with the sub-cluster of self-regulated learning being the largest. From the perspective of the types of competences promoted, the reviews found the following categories of papers – metacognitive, cognitive, behavioral, emotional, and self-regulation. Given the growing recognition that LADs can be a helpful tool for providing reference frames [18], the Jivet *et al.* reviews classified papers into those providing social, achievement, and progress reference frameworks.

Although the existing literature reviews offer valuable contributions toward incorporation of the learning sciences into the design and evaluation of LADs, there is a need for more rigorous examination of the existing dashboards from the perspective of specific theoretical models of learning. Given the overwhelming recognition of the role of LADs to support self-regulated learning, the current study was set out to scrutinize systematically the existing literature based on a well-known model of self-regulated learning.

B. *Self-Regulated Learning and Feedback*

Self-Regulated Learning (SRL) research aims to optimize learning skills by exploring cognitive and metacognitive processes that encompass several internal and external factors. Zimmerman defines self-regulated learning as "the process whereby students activate and sustain cognitions, behaviors, and affects, which are systematically oriented toward attainment of their goal" as cited by [19] (p. 465.). Several SRL models were proposed. Zimmerman developed several versions of SRL models based on socio-cognitive theory. Boekaerts' model was developed based on the role of goal and emotion [20]. Winne and Hadwin developed SRL model based on the Information Processing Theory [14]. Regardless of the difference foundation that built up the model, the proposed SRL models involve cognitive, metacognitive, motivational factors and goal that drive the learning process. In this study, we based our systematic literature review around Winne and Hadwin model, as the model has been broadly adopted in the computer supported learning [21], [22], the cognitive and metacognitive components were described in more details as compared to other models [21] and the role of external feedback was clearly

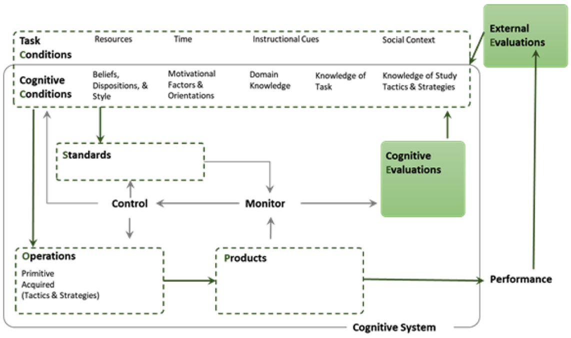


Fig. 1. Representation of the COPES model [14].

highlighted in the model [15]. Winne and Hadwin [14] state that a self-regulated learning process involves four cyclical recursive phases including *task definition, goal setting and planning*, *enactment of tactics and strategies*, and *adaptation*. They highlight that within these four phases, five components are running recursively. These five components are conditions, operations, products, evaluation, and standards, to form the COPES model. Figure 1 illustrates the COPES model developed by [14].

Learning tasks are designed to guide the learner to achieve a specific goal. Understanding the task definition is important as it will affect the selected learning strategies [23]. In the early phase of learning, learners define the tasks and set learning goals by considering several constraints. These constraints are referred as conditions in the COPES model. Condition can be internal ones such as knowledge of tasks, domain topics, learning tactics as well as motivational factors. External conditions can be resources available, learning environment, instructional cues and time constraints to complete a given task. Based on these conditions, students make judgements that drive setting goals, planning their learning, and setting expectations or standards. That is, according to Winne and Hadwin, goals can be considered as a vector of different standards learners will use for metacognitive monitoring (e.g., how long they expect to study to recall some information or how they will self-assess the coherence of the argument in their learning activities). Learners operate their learning by the learning strategies and tactics. As they progress with their learning, products of learning will be created (e.g., memory recall or essay). As the learning process unfolds, learners evaluate the learning products and learning processes with the standards they had set earlier. The evaluation can lead to the choice of new learning strategies, maintenance of the current learning strategies, or updating their standards and thus revising existing and setting new goals [14].

Supporting learners to regulate their learning requires understanding and incorporating the self-regulated learning process into the supporting system. Feedback is one of the crucial elements in any SRL processes. In the COPES model developed by [14] feedback occurs internally when learners evaluate their learning against the standards that define their goals. Whereas, external feedback could also be implemented and provided by instructors or other external agents. External feedback can confirm, add to, or alter the internal feedback perceived by students, which subsequently effects on the learning process [15]. However, [24] states that students are inaccurate in judging their performance. Students with good performance tend to underestimate their learning process whereas students with lower performance tend to overestimate it [24]. Moreover, Winne and Jamieson-Noel [25], [26] found that what students' self-report about own learning was not in accordance with their action. These misperceptions of the learning progress and performance can lead to selecting an ineffective or inadequate learning strategy [23]. Hence, external feedback, especially from teachers or learning technologies, could enhance the accuracy of judgments made by students regarding their progress and performance.

This systematic literature review is based on the assumption that LADs are a form of feedback that aims to equip learners to take control over their learning and thus better self-regulate their learning [27]. LADs can play a role of feedback for students and teachers. A conceptual model has recently been proposed to show how LADs can be used to "provide cognitive and behavioral process-oriented feedback to learners and teachers to support regulation of learning" [28] (p.1). LADs are also suggested to provide cues to support the evaluation on students' current state of self-regulated learning and progression towards their goals [29]. Other authors propose that the LADs have a potential to reduce negative affect, motivate students, and assist them to reflect on their self-regulated learning process [30]. Specifically, this review aims to examine the role of learning analytics as a form of feedback by using the COPES model proposed in [14].

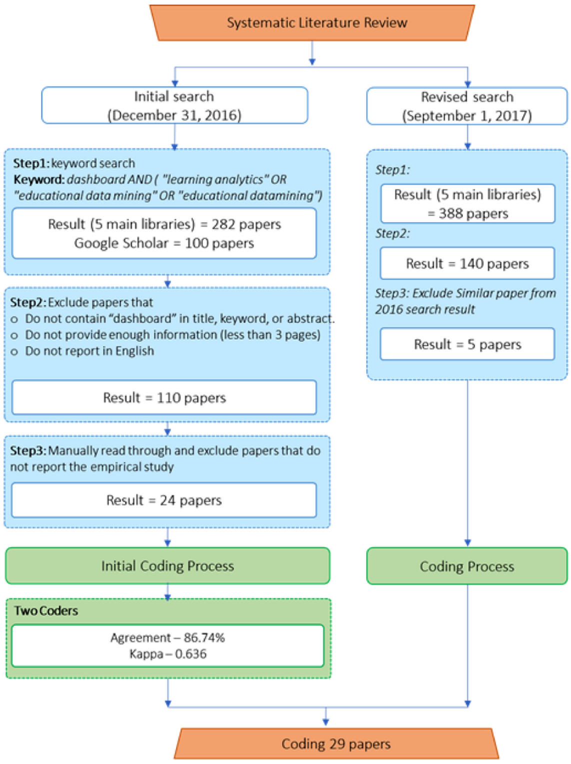


Fig. 2. Methodology used in this systematic literature review.

III. Method

The process of the systematic review followed in this study is summarized in Figure 2. The review focused on the existing literature published between 2010 and 2017 until the time the search was completed (September 1st, 2017). By following guidelines proposed by Kitchenham & Charters [31], three steps were taken in order to conduct the systematic literature review and these steps were repeated twice during 2016 and 2017 to ensure that the relevant papers were included.

The first step in the systematic review was keyword search. Five main academic databases were selected: ACM Digital Library, IEEE, SpringerLink, Science Direct, and Wiley. Google Scholar was included as an additional database to detect other research resources. A total of 488 papers were obtained by running the query: dashboard AND ("learning analytics" OR "educational data mining" OR "educational data mining") (initial search in 2016 = 382 papers and revised search in 2017 = 488 papers). A total of 488 also includes top 100 papers from Google Scholar.

The second step was carried out to filter insufficient and irrelevant papers by screening paper titles, keywords, and abstracts to identify those that were describing LADs. Only papers related to LADs were included. Papers that were not written in English and papers containing less than 3 pages (e.g., posters) were excluded. As a result, 140 papers were selected (initial search in 2016 = 110 papers; revised search in 2017 = 140 papers).

Finally, papers that solely shared opinions, provided reviews or designs of dashboards, offered proposals or conceptual frameworks and the duplicate papers or preliminary versions of papers (conference papers compared to extended journal papers) were removed (unless they presented new and different aspects). Lastly, a total of 29 papers successfully passed all the inclusion criteria and were included in the final analysis.

In the analysis, we also referred to the screenshots of dashboards presented in the papers to scrutinize the indicators presented in the dashboard. Then, we coded the papers according to several dimensions based on the three research questions as summarized in Table I. The papers were manually coded by two independent coders (the first two authors of the paper) and the coding process is detailed in the remainder of this section. Out of 822 coding items, the percent agreement between two coders was 86.74% and Cohen's Kappa = 0.64. All initial differences in coding were then discussed and reconciled. The results presented in the paper are based on the reconciled coding.

To answer the first research question, the study focused on whether the papers mentioned the theoretical or relevant models that were used to guide the choices of indicators that should be presented to the users. Furthermore, we also examined if the papers mentioned any educational theory used to guide the design and development of LAD. Then, each indicator described

TABLE I
CODING DIMENSION USED FOR ANALYSIS OF LAD PAPERS FOR THE THREE RESEARCH QUESTIONS

Research Questions	Elements in Dashboard	Coding Category
1. What is the support of LADs for the elements of self-regulated learning as established in the learning sciences?	Indicators Selection Methods	Method used in selection of the indicators in dashboards (based on previous work/Users centered design)
	Theory Reference	Mention of theory on the selection of indicators in dashboards (Yes/No)
	Indicators in the dashboards	SRL constructs and Items (Conditions, Operations, Products, Evaluations, Standard)
	Learning Phases	4 Learning phases according to COPES model (Refer to Table III)
	Theme of Dashboard	Theme of dashboard (Refer to Table IV)
2. What type of information is offered as feedback in LADs?	Reference frames	Social Norm (individual viewing and average class comparison)
	Types of visualization	Type of chart used
3. What is the quality of study designs and reporting the literature that reports on empirical evaluations of LADs?	Target users	Stakeholders (student, teacher, administrator, designer)
	Participants demographics	Participants (Expert, Student, and number of participants)
	Evaluation method described	Evaluation instruments used
	Quality of study reporting	Assessment based on a 12 questionnaire for research quality appraisal

in the papers and presented in the figures/screenshots of the LADs was mapped to its corresponding elements of the COPES model. Table II presents the description of SRL constructs used in our analysis of the selected papers.

The indicators presented in the dashboards were retrieved from the results of each learning phases. Therefore, these indicators can be considered products of learning in each phase. Hence, products in the COPES model were analyzed according the four phases of self-regulated learning as shown in Table III [14]. LADs aim at providing feedback for different purposes. We identified eight themes of feedback support (Table IV).

The analysis of information presented in LADs as asked in research question 2 looked at two key dimensions. First, it examined the extent to which different reference frames were supported in LADs [18]. Similar to the approach followed by Jivet and her colleagues, references frames included individual (self) and social (average comparison and course-wide including the provision all information available about other users and environment) as follows: a) individual – students can only see their own individuals' activities; b) average comparison (social) – students are provided with the average comparison against their peers, classmates or course mates; c) course-wide information (social) – all information is available to all target users.

Quality of study design and reporting was assessed by using the instrument for appraisal of empirical research specifically developed for this study, given that there is no generally accepted instrument for assessment of empirical research in education or educational technology. The instrument used in this study was informed by the MERSQI (Medical Education

TABLE II
DESCRIPTION OF THE SRL CONSTRUCT AND ITEMS BASED ON THE COPES MODEL

SRL Construct	Items	Description	Example of the indicator(s)
Task Condition	Resource	Indicators that show the available resources or learning materials	List of course and relevant competency
	Instructional Cues	Indicators that represent the structure of the courses or learning materials	Student's weekly workload; Structure of learning environment (e-book); Learning Concept
	Time	Time available; ability of student to manage the time	Students' time divided among learning space (face-to-face & online); Time utilization
	Social Context	Indicators that are related to group participation, collaboration, connections, students' interaction with teacher, or role of students in learning environment	Roles adapted by learner from discussion (using Classification); Collaboration
Cognitive Condition	Domain Knowledge	Indicators that represent the knowledge of student in certain topic or learning performance	Quality of post in discussion board; Scores
	Beliefs, Dispositions, & Style	Indicators that represent the student's beliefs, disposition and learning style	Deep/surface learning; Result of writing habit by a single student
	Motivational Factors & Orientations	Indicators that aim to motivate learning; such as by using of badges and awards	Achievement badges and details; Total stars earned
	Knowledge of Task	Indicators that represent the task related knowledge	Task Value
	Knowledge of Study Tactics & Strategies	Indicators that represent the tactics of study	Remediation
Operation	Primitive	Indicators that represent actions of students in learning environment e.g. numbers of video watched	Group progress in their task; Assignment Progress; Total no. of comments; Total no. of tweet
	Acquired (tactics & Strategies)	Indicators that indicate if students had achieved the specific learning skills	Average skill mastery plotted against average amount of practice
Product	-	The result of the operations; coded based on learning phases	-
Evaluation	-	Indicators that indicate if student had achieved their target outcome	-
Standard	-	Indicators that specify the student's target outcome	-

TABLE III
PHASES OF SELF-REGULATED LEARNING BASED ON THE COPES MODEL

Learning phases	Description	Example of Indicator(s)
Phase 1: Task identification	Indicators that present information to guide students on their learning tasks	Grading criteria; Student's weekly workload;
Phase 2: Goal setting and planning	Indicators that represent the goal of learning activities and conditions that are relevant to learning strategies selections	Goal on table
Phase 3: Enactment of learning strategies and study tactics	Indicators that represent the actual behaviors and interactions of students	Group progress in their task; Level of engagement with videos;
Phase 4: Adaptation	Indicators that represent the continuation of certain learning strategies (maintain of action) or changing of learning strategies over time	Most spent timing website; Result of writing habit by single student;

TABLE IV
THEMES FOR WHICH LADS ARE USED

Theme	Description
Competency	Any dashboards that aim to track competency development progress (set of knowledge or skills that students are required to achieve)
Emotions	Any dashboards that track students' emotions while learning
Game-based learning	Any dashboards that track learning activities retrieved from informal game-based learning system
Learning progress	Any dashboards that track learning activities retrieved from formal face-to-face, blended or online learning context
Learning design	Any dashboards that track the activities in the course design process
Learning difficulty detection	Any dashboards that were developed to detect students who diagnose with learning difficulty
Study plan	Any dashboards that were developed to assist the students and/or study advisers to plan their learning for a given period (semester)
Teamwork progress	Any dashboards that track students' learning activities in group-based learning

Research Study Quality Instrument) instrument [32] and the coding scheme used in [33]. The MERSQI instrument is a 10-item coding instrument often used to evaluate the quality of the research studies in the medical field, focusing primarily on the quality of double-blind clinical trials.

Similarly, the instrument from [33] focuses on assessing the methodological quality of the studies that compared distance education (DE) and face-to-face (F2F) education. The [33] instrument includes 13 items that focus on the quality of reporting of statistical results and equality of comparison groups (e.g., did DE and F2F conditions have the same instructional approach, same instructors, and assessment instruments). The final instrument used in this study included 13 questions, which are shown in Figure 11. The responses to each of the questions were "Yes" indicating some presence of the dimensions of the study design and report in the papers, and "No" indicating the total of the discussion or indicators pertaining to the dimensions analyzed.

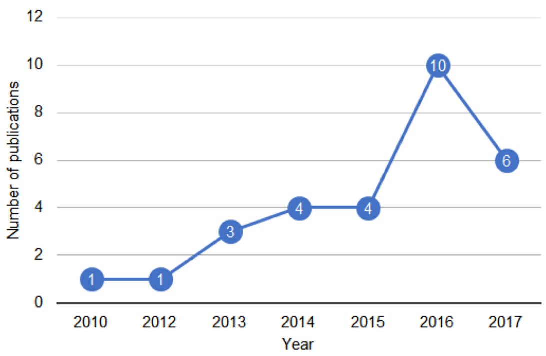


Fig. 3. Publications included in this review according to years of publication.

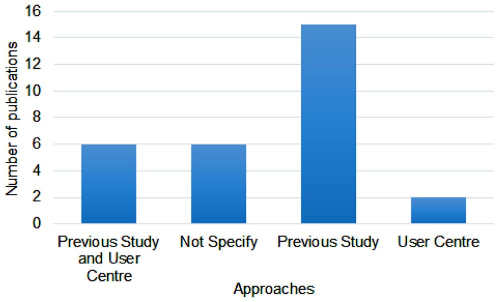


Fig. 4. Indicators selection approaches used in the design of LADs.

IV. RESULTS

In this section, we present the results of the article review according to the three research questions. Figure 3 shows that there has been a steady increase in the number of empirical evaluations on LADs over past few years. The amount of empirical publications on the LADs research peaked during 2016. Meanwhile, the number slightly declined in 2017. Note that, the search conducted in September 2017 and the amount of empirical publications on the LADs research is likely to increase during the final quarter of year 2017. Appendix provides summary of the research studies included in this systematic literature review.

A. RQ1: Support of LADs for the Elements of Self-Regulated Learning

1) Indicator: An indicator represented the state or level of a student's actions and performance. Selection of indicators is recognised as one of the most important steps in the design of LADs as dashboard provides users with the feedback via the selected indicators. Figure 4 presents the sources and the approaches used for the selection of indicators as reported in papers included in this study.

Two main approaches were followed: referring to the previously published papers and/or following user centered design. A total of 52% of the included papers mentioned that they

TABLE V
EDUCATIONAL THEORIES AND MODELS USED TO DERIVE THE DEVELOPMENT OF THE LADS

Related learning model used in the design and development of dashboard		Number of papers
No	Not specified	20
Open Learner Model (OLM)	Open Learner Model (OLM)	1 [35]
	Open Learner Model; Activity-based Learner-Model (Engeström's Activity Theory and Actuator-Indicator Model)	2 [36], [37]
Emotion	The TEA Model (TEAM)	1 [39]
	Discrete and dimensional emotions theories by using Automatic Emotion Recognition process (AER)	1 [42]
Motivation and Goal Orientation	Goal Orientation	1 [43]
	Learning and Study Strategies Inventory (LASSI)	1 [44]
Competency	Fully-Embedded Assessment Model (FEAM)	1 [46]
Teamwork	Dickinson and McIntryre's teamwork model	1 [47]

solely selected the indicators based on the previously published works. A total of 20% of the included papers referred to previously published papers and in combination with user-centered design to select which indicators should be presented. A small number of the papers (7%) relied on the users' input alone in the selection of the indicators. Whereas, 21% did not provide any justification on how the indicators were selected.

2) Theory and Learning Model Application: A main criticism of the literature on LADs is that many of them do not have grounding in a sound theoretical foundation [5], [34]. Thus, this review looked at the theory or related learning model referenced in the papers as sources that informed the design of LADs. Table V provides details of educational or pedagogical models used in each category. Despite the importance of educational theory, most of the LADs were not grounded in any established educational theories (69 percent). Open Learner Models (OLMs) were used by three research papers [31]–[33]. OLMs present to students the information such as knowledge, preferences, and skills which previously were not available to the student. This information is often extracted and computed by an analytic algorithm [35]. OLMs have been used by many systems with the proposition that by unveiling this information through OLMs to students, students can improve their awareness and reflection on their learning [36]. Florian-Gaviria *et al.* [37] and Mejia *et al.* [36] incorporated OLMs with activity theory [38] to develop a framework that was used for the design of LADs.

Two dashboards were developed to detect the emotion of learners with the assumption that emotion impacts learning. The TEA model was proposed by [39] which was informed by the work of Pekrun, Goetz, Frenzel, Barchfeld, & Perry [40] and Arroyo *et al.* [41]. The TEA model defines six positive emotions and six negative emotions that affect learning. Based on the TEA model, a self-report instrument was developed to track students' emotion according to teachers' defined events [39]. The data collected in this way were then presented to learners in the dashboard. Another study that emphasized emotion tracking was carried out by Ez-zaouia & Lavou [42]. They stated that two most commonly used theoretical viewpoints for consideration of emotions are discrete and dimensional. To capture discrete and dimensional emotions, the Automatic Emotion Recognition process (AER) based on facial expression and voice recognition was used in combination with self-reported data and trace data to automate detection of students' emotion and shown the results in a LAD.

Theories of achievement goal orientation and motivation are also used in the papers included in this literature review. Beheshitha, Hatala, Gašević, & Joksimović [43] developed three LADs that promoted participation in online discussions by considering achievement goal orientations (AGO) of students. Self-reported measures of AGOs were used to explore the association among three different LADs, individual differences, and learning behavior. Broos, Peeters, Verbert, & Soom [44] employed an extensively and widely used Learning and Study Strategies Inventory (LASSI) questionnaire [45] to capture students' learning skills and motivation in order to present feedback and recommendations to support the development of learning skills. The dashboard was focused on five out of ten scales that were detected from the LASSI instrument including performance anxiety, concentration, motivation, the use of test strategies, and time management.

The fully-embedded assessment model (FEAM) was developed by Capella University to measure students' competency development [46]. The FEAM identified a set of learning goals and grading criteria. The dashboard was then developed according to FEAM to offer feedback to students. Tarmazdi, Vivian, Szabo, Falkner, & Falkner [47] relied on Dickinson and McIntryre's [48] teamwork model to capture the role and emotion of students working in a group. Dickinson and McIntryre's [48] identified seven core components of teamwork based on practical teamwork including team leadership, team orientation, monitoring, coordination, communication, feedback and backup behavior. The measurement of these core components was based on how the team responded to the critical events according to the seven core components.

3) Indicators and SRL Constructs: Since no paper explicitly considered SRL theories in the design of LADs, the current study coded the indicators presented in the papers based on the descriptions of the dashboards provided in the papers (including screenshots) according to the elements of the COPES model (Tables II and III). There was a total of 266 individual indicators from the 29 dashboards presented in the 29 papers. A majority of indicators represented Cognitive Conditions and Operations as shown in Table VI. The indicators related to Operations – Primitive (66 indicators) indicated actions such as a number of log-ins, number of message posts, and number of questions answered. However, there was no indicator representing the skills or tactics acquired by students. A total of 74 indicators represented domain knowledge as part of learners' cognitive conditions. These indicators were test or exam scores, correctly completed exercises, or quality of message posts. A total of 80 indicators represented motivational factors. Most of these indicators were based on the use of badges to create game-based learning such as a number of badges earned or

TABLE VI
INDICATORS REPRESENT BASED ON THE CORRESPONDING LEARNING PHASES AND SRL CONSTRUCT

COPES Model	Item in COPES	Learning Phase					Total Number of indicators
		Phase 1	Phase 2	Phase 3	Phase 4	Not Related	
Cognitive Conditions	Beliefs, Dispositions, & Styles	2		1	3		6
	Domain Knowledge	19		53	2		74
	Knowledge of Study Tactics & Strategies	1					1
	Knowledge of Task	1					1
	Motivational Factors & Orientations	24	2	54			80
Task Conditions	Instructional Cues	11		5			16
	Resource	1					1
	Social Context			5			5
	Time	2		8			10
Operations	Primitive			63	3		66
	Acquire (Tactics & Strategies)						0
Others	Demographic Information					6	6

level of participants in games. Other types of SRL constructs were relatively low. Six indicators represented beliefs, dispositions, and styles, one indicator represented the knowledge of study tactics, and one displayed the knowledge of tasks.

Task conditions in the Winne and Hadwin model of SRL indicate the constraints in external conditions (e.g., group task or open book exam). Only 16 indicators were about task conditions, of which 10 indicators represented time utilization, five indicators reported about social context, and one indicator addressed the resources available. Based on our review, two elements of the COPES model were completely missing in the papers describing LADs, namely standards and evaluation.

4) Indicators and Self-Regulation Phases: We coded the indicators present in the LADs based on the four phases of self-regulated learning of the Hadwin and Winne model (Table III). Table VI presents the cross-tabulation of the numbers of indicators according to the COPES elements and the phases of self-regulated learning.

Phase 1 – Task identification: In this stage, students develop their perceptions and understanding of what tasks need to be done [14]. The indicators in this phase represent the constraints and conditions of their learning such as grading criteria, previous learning difficulty, previous grades, and successful rate of previous year's students. Based on this literature review, the indicators at this stage focused on the task conditions and cognitive conditions. Other elements in COPES model were absent. Cognitive condition primarily included domain knowledge along with motivational factors and goal orientations (19 and 24 indicators, respectively). Knowledge of study tactics and strategies [49] and knowledge of task [44] each had only one indicator in the reviewed dashboards. Knowledge of study tactics and strategies was represented by the indicator that was called the 'level of test strategies usage (respective norm scales) of a student' and knowledge of task was presented by indicator called 'task value'. Instructional cues were the most represented indicators in task conditions (11 indicators). Task conditions were represented by two indicators of time: i) the level of time management (respective norm scales) of a student and prediction of academic year based on historical data; and ii) resources had one indicator, which offered information about dynamic question text editing and indexing.

Phase 2 – Goal setting and planning: In this stage, students set goals for their learning, select learning strategies, and create a plan to achieve their target. There were only two indicators that represented the COPES elements in phase 2. These two indicators were related to goals which fall under cognitive conditions - motivational factors and goal orientation – in the COPES model [49], [50].

Phase 3 – Enactment of learning strategies: Most of the learning activities occur during this phase where students enact the learning strategies that they have selected in phase 2 [14]. The most frequently presented indicators were operations – primitive (63 indicators); cognitive conditions – domain knowledge (53 indicators) and motivational factors and orientation (54 indicators). Whereas, knowledge of tasks, resources, and tactics and strategies that students acquired were completely absent in the dashboards presented in the papers included in this study.

Phase 4 – Adaptation: Adaptation refers to a large-scale adjustment of students' learning behavior based on the performance of metacognitive monitoring and control. There are generally two types of adaptation – i) adaptation of learning strategies, and ii) updating standards that constitute goals set in phase 2. The changes of standards can also affect students' motivation, and self-efficacy, belief, or disposition. The indicators of operations from the COPES model related to phase 4 were focused on primitive operations. Examples of such indicators included the most frequently visited sites and most frequently active documents (three indicators). Cognitive conditions were represented by indicators of domain knowledge including 'reflection on mastery level' and 'competency achieved'. No other elements of the COPES model were present in the dashboards to support phase 4 of the Winne and Hadwin model.

5) Theme and SRL Constructs: We categorized 29 dashboards according to the themes of the dashboards as described earlier in Table IV. As shown in Table VII, a large number of the dashboards tracked students' learning progress in face-to-face, online and blended learning environments (18 papers).

Other themes of dashboards that were captured in this study were competency tracking [35], [46], game-based learning [51], [52], emotion tracking [39], [42], and study planning [44], [53] (two dashboards in each theme) while learning difficulty detection [36], teamwork progress tracking [47] and learning design tracking [54] were supported by a single dashboard in each case.

Competency: Two dashboards were developed to track the competency development of students [35], [46]. The indicators presented in the competency tracking dashboards focused on cognitive conditions – domain knowledge and motivational factors and orientations (five and seven indicators, respectively). For cognitive conditions, domain knowledge related indicators focused on grade and performance. Similarly, as part of cognitive conditions, motivational factors and orientations related indicators represented the ranking (based on the course

TABLE VII
INDICATORS BASED ON THE THEME OF DASHBOARDS AND THE CORRESPONDING COPES ELEMENTS

OPES Iodel	Item in COPES	Theme of the dashboard							
		Competency (2 LADs)	Emotion (2 LADs)	Game-Based (2 LADs)	Learning Design (1 LAD)	Learning Difficulty Detection (1 LAD)	Learning Progress (18 LADs)	Study Plan (2 LADs)	Teamwork Progress (1 LAD)
nitive ditions	Beliefs, Dispositions, & Styles					3	1	2	
	Domain Knowledge	5		6		10	35	18	
	Knowledge of Study Tactics & Strategies							1	
	Knowledge of Task						1		
	Motivational Factors & Orientations	7	23	14			13	21	2
ditions	Instructional Cues	2			4		2	8	
	Resource				1				
	Social Context						3		2
	Time						8	2	
rations	Primitive	1	3	9			53		
	Acquire (Tactics & Strategies)								
ers	Demographic Information					5	1		

completion and performance) of individual students against the peer and course. That is, they were primarily focused on social reference frames [11]. Task conditions highlighted instructional cues by using grading criteria and lists of relevant competencies. Only one primitive operation was tracked in this type of the dashboards and it was about the progress on assignments.

Emotion: The dashboards that tracked emotions [39], [42] mostly used the indicators fell under the category of cognitive conditions – motivational factors and orientations (23 indicators). Some of the primitive operation were also presented in the dashboards such as learner's audio record or facial expression during a certain learning period (three indicators). Other types of indicators from the Winne and Hadwin model were not used in the dashboards included in the review.

Game-based learning: Most of the game-based dashboards [51], [52] used badges with the aim to motivate students (14 indicators). Cognitive conditions focused on domain knowledge through the use of indicators like the level of achievement in certain rounds of games (six indicators). Operation were primitive and included indicators such as total number of activities and total number of badges awarded to the class (nine indicators).

Learning Progress: There were 18 dashboards which fell under the learning progress tracking theme. A variety of feedback types were provided to users. The most common indicators represented information about primitive operations such as number of logins, posts, or tweets (53 indicators). The second most frequently provided information to students was about domain knowledge (35 indicators). Cognitive condition was represented by motivational factors and goal orientations through 13 indicators. Task conditions were highlighted by a small number of indicators, including eight indicators about time – time spent on course site, time spent on different kinds of activities, time before the first attempt, and time of submission of before a deadline [57]–[59] whereas three indicators showed social context – collaboration and communication [49] and two indicators represented instructional cues – feedback on overall performance and recent behavior and behavioral activities [30], [58]**.**

Learning Design: A single dashboard that captured the activities of students to inform teachers' design of courses included the information such as concept connections [54]. Four indicators about instructional cues (i.e., concept count, concept connections, correct answer indication, and question marks assignment) and one about resources (i.e., dynamic question text editing and indexing) were presented in the dashboard.

Learning Difficulty Detection: The learning difficulty detection dashboard [36] only displayed information on domain knowledge (i.e., the previous diagnosis of learning disabilities, the number of associated difficulties with reading) and belief, disposition and styles (the result of writing habit by a single student and the result for a single student with learning style: active, sensory, visual and sequential) with ten and three indicators respectively. Beliefs, dispositions, and styles were represented by using indicators that reflect learning styles and were collected with self-reporting instruments.

Study Plan: This type of dashboard aims to help the study advisers and students to plan learning and course enrolments [44], [53]. Of cognitive conditions, motivational factors and goal orientation as well as domain knowledge were the most frequently presented indicators for this type of dashboard (21 and 18 indicators, respectively). Other types of cognitive conditions were observed such as two indicators of beliefs, dispositions and styles and one indicator represented knowledge of study tactics and strategies [44]. Task conditions were observed through the indicators that represented instructional cues and time condition (eight and two indicators, respectively).

Teamwork Progress: The teamwork dashboard [47] provided information about the role played by a student in a group learning activity (i.e., two indicators represented social context within task conditions) and one emotions of team members (i.e., two indicators represented motivational factors and orientations within cognitive conditions).

B. RQ2: Types of Information Offered as Feedback

1) Types of Reference Frames: Types of reference frames used in LADs are presented in Figure 5. Some dashboards were designed for several target user groups. Therefore, the retrieved indicators were counted independently for each target user group.

Most of the indicators represented the individual-related learning activities (80 indicators from students' dashboards, 81 indicators from teachers' dashboards, and seven from others' type of users' dashboard). The dashboards that were designed for students provided with the average comparison (so called social reference frames) against their peers and class (31 and 32 indicators, respectively) while there was only one indicator that presented the students against the top performers.

Course-wide information indicators allow users to view all the information available in a learning analytics system. Most of the dashboards in this dashboard type were designed for

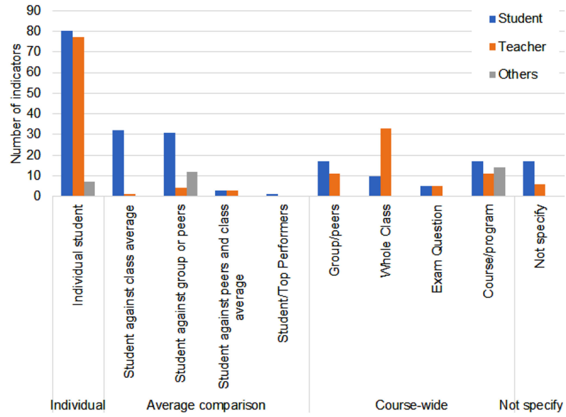


Fig. 5. Type of reference frames used in LADs for each target user group.

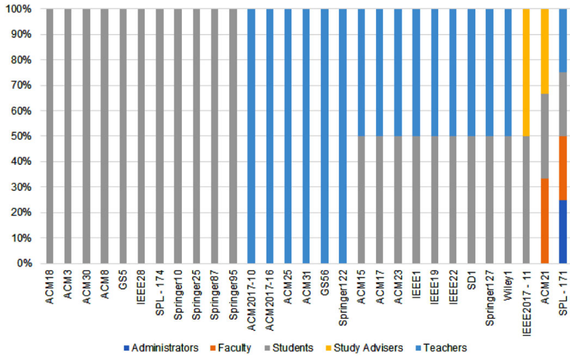


Fig. 7. Target user of LADs.

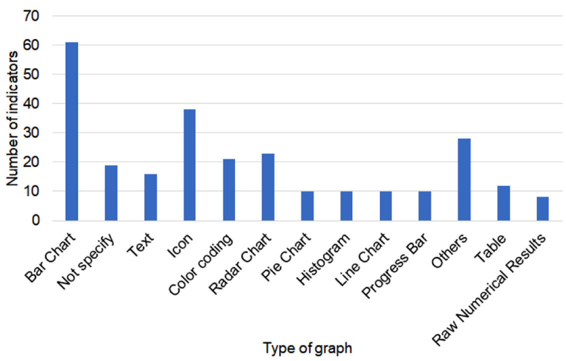


Fig. 6. Visualization type.

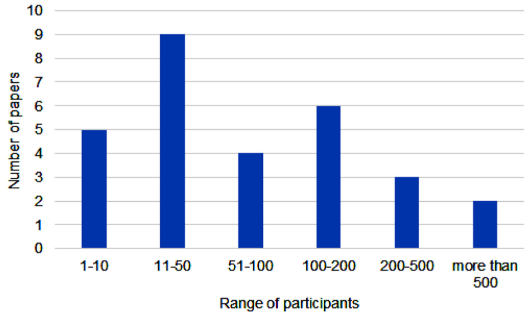


Fig. 8. Range of participants.

teachers who can view a whole class's related information (33 indicators) [52], [55], [57], [59], [60]. The dashboards that were developed to track team progress also presented all information to all students in a given team as well as some of the indicators that aimed to display levels of achievement to motivate the students and help them plan their learning (17 indicators) [50]–[52]. A dashboard that was developed to track the design of exam questions displayed the structure of each exam questions to teachers (five indicators) [54]. Two study plan tracking dashboards also offered some course-wide information such as the success rate of those who previously took the course (17 indicators from students' dashboards; 11 from teachers' dashboards, and 14 from other users' dashboards) [44], [53].

Data representation is another debated issues in LADs. Based on our research, common graphs used to represent data was bar chart as presented in Figure 6. The use of simple visualizations aims to aid students in interpreting the result. However, some research reports that students face a problem in interpreting graphs available in contemporary dashboards [61], [62]. Similarly, Corrin, Kennedy, & Mulder [63] stress the concern of teachers regarding the ability of students to interpret the feedback presented in dashboard. Yet, there have been few empirical studies to inform on the selection of visual display to represent the identified indicators in the dashboard and the influence of these representation on students' level of understanding and motivation (if any) of feedback [4].

C. *RQ3: Quality of Study Design and Reporting*

1) Participants: Figure 7 illustrates the dashboards' target users. Based on the 29 dashboards described in the papers, the main target users were students (22 dashboards in total) and teachers (18 dashboards in total). Among these, some dashboards were designed to be viewed by both teachers and student (10 dashboards). Other audiences of dashboards were administrators, other faculty members, study advisers, and designers (three dashboards in total).

Based on 29 papers, the number of participants involved in each study varied, range from one to more than 500 participants. Figure 8 presents the number of participants involved during the empirical studies.

Five papers involved between one to ten participants. Nine papers included participants between 11 to 50 persons. Four papers reported the number of participants from 51 to 100 and 11 papers included more than 100 participants. Based on the different numbers of participants, we further explored if the participants referred to students or experts. Experts in this context refer to both teachers and other audiences such as learning/ course designers and study advisers.

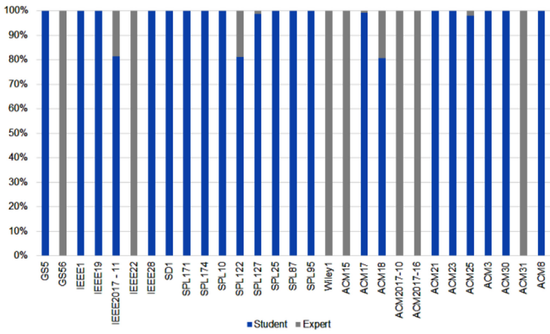


Fig. 9. Participants included in each study.

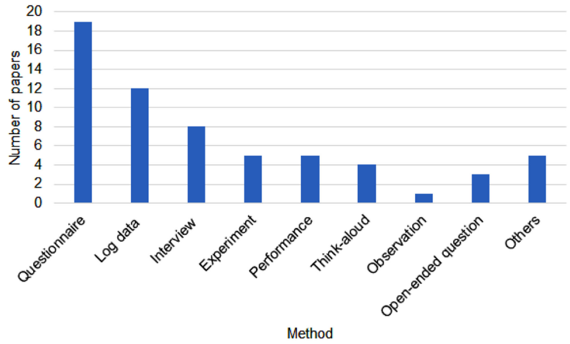


Fig. 10. Research approaches used for data collection.

Figure 9 shows the percentage of students and experts participated in LAD evaluation. Most of the papers included students in their studies (20 studies). There were nine papers that involved experts alone in their studies. Four papers included both experts and students in their studies.

2) Course Demographics: Most of the papers reported that students were from a single course in their studies (12 papers). Another group of papers (11) reported that students were from two to five courses. Four papers reported that their participants were from diverse background courses and two papers did not report the academic context in which their evaluations took place. The courses were diverse and a majority of them were computer science and engineering related courses such as C programming, human-computer interaction, and object-oriented programming. Other courses included Japanese, law, psychology, and Chinese literature.

3) Data Collection Approaches: The approaches used to gather data varied. Figure 10 shows the methodology used by each paper. Most of the papers used more than one instrument to gather the required data. The main approaches that had been used to gather data were questionnaires, trace data, and interviews (19, 12 and 8 papers, respectively). The evaluation of LADs was based on usability studies mostly and self-report instruments. There was a lack of solid evidence in terms of observational data in real class setting on how LADs, as a form of feedback provision, influenced behavioral change, learning strategy, and learning performance.

4) Quality of Study Designs and Reporting: Figure 11 shows the results of the quality of study design and reporting as presented in the papers to evaluate the impact of LADs. The analysis showed that less than half of the studies (11 papers) reported limitations. We also observed that there was no discussion on the generalizability derived from any of the studies. However, slightly more than half of the studies interpreted the results with respect to the current literature (16 papers). Just under a half of the studies (14) studies discussed implications for research and 20 studies addressed some implication for practice. Only five studies reported the use of experimental study designs that involved the use of control and treatment groups. The use of quantitative research based were reported by 17 studies. The main methods of data collection were questionnaires as highlighted earlier in the paper. Most of the research studies (19 studies) provided some description of the courses in which the evaluations were conducted and clearly stated the number of participants involve in the studies. However, the demographics of participants were not stated by many research studies (only 13 studies reported the demographics of participants). Besides quantitative methods, there were four papers that solely employed qualitative methods in their studies which involved collecting data from interview (2 studies), interview and think-aloud protocols (1 studies) and case studies (1 study). In addition, there were only eight research papers conducting and reporting on mixed methods studies (combination of qualitative and quantitative research methods). The most common mixed method research designs were a combination of surveys and interviews (4 studies) followed by survey and think-aloud protocol (1 study), survey and focus group (1 study), survey and log data (1 study) and log data and interview (1 study).

V. Discussion

Based on the insights obtained from this systematic literature review, we highlight four dimensions that should be considered when researching and developing user-centered learning analytics systems. These dimensions include theory, design, feedback, and evaluation. These dimensions are included in the model for user-centered learning analytics systems (MULAS) (Figure 12). The model assumes the cyclical and recursive nature of the four dimensions. Each of the four dimensions is discussed in the remainder of the section.

At the core of MULAS is the recommendation that future research and development should make any a priori design decisions about representation of data and analytics results in learning analytics systems such as LADs. Instead, the focus should be on developing user-centered learning analytics systems with the emphasis to support users with learning analytics to accomplish set tasks in the most effect way. Therefore, the reminder of the discussion section draws recommendations for

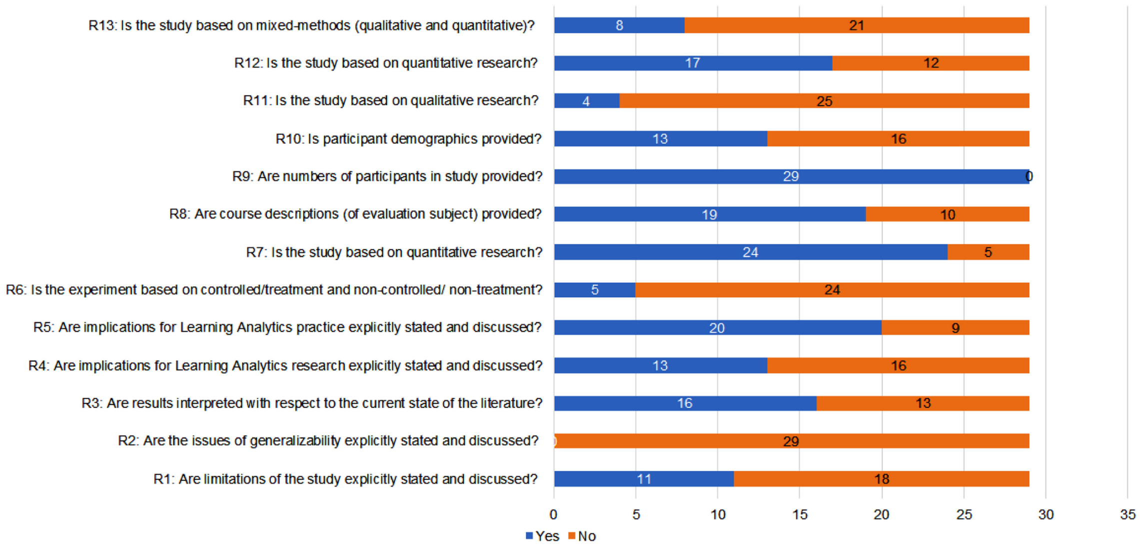


Fig. 11. Quality of reporting of the empirical studies with LADs.

user-centered learning analytics systems as a class of systems that subsumes LADs.

A. *Theory*

The most striking finding of this study is that a great majority of the existing dashboards (68%) are atheoretical in their choices of indicators and content to be shown in Table VI. The arguments for the importance of theory and the use of learning sciences for informing learning analytics have already been made by several researchers [2], [5], [28], [34], [64]. In a nutshell, without building on what is already known about learning and teaching and instead, using design- or data-driven approaches will likely result in ineffective or even deteriorating effects on learning. Unsurprisingly, several authors have already reported negative effects of existing LADs on learning [6], [65], [66]. Possible reasons for this can be derived from the results of this study that analyzed the existing learning dashboards against a well-known theoretical model of self-regulated learning [14]. The analysis against the COPES models revealed that the existing LADs did not support knowledge of learning strategies and tactics, and knowledge of tasks under cognitive conditions, acquisition of learning tactics and strategies under operations, and standards and evaluation. Failing to capture knowledge of and operations associated with learning tactics and strategies makes it hard if not impossible to understand and optimize learning; understanding and optimization of learning are the ultimate purposes of learning analytics [67].

Current educational psychology research indicates that learners use suboptimal learning tactics and strategies and many are even unaware of the effective study tactics (e.g., self-testing and spaced learning significantly enhance memory retention in comparison to reading and rereading) [24], [68]. Just informing learners of the benefits of some of the effective tactics and strategies significantly increases the chances of learners to adopt them in their learning [69], [70]. Therefore, LADs are likely to be less effective if they are ignorant of what learning tactics and strategies learners follow, and if they do not increase their awareness and offer recommendations for more optimal approaches. The recent approaches to mining tactic and strategy from trace data [73]–[75] are highly promising for the future work on user-centered learning analytics systems. The important next step for research is to find the right mechanisms to communicate the learning tactics and strategies discovered with data mining to learners and educators along with some recommendations on how to optimize the learning process.

Capturing knowledge of tasks is as important as knowledge of learning strategies and tactics. Awareness of the benefits of a tactic or strategy is not sufficient for the learners to adopt the tactic or strategy for a given task. Winne [74] suggests this steams from the fact that learners need to be aware that the tactic or strategy is beneficial in a context different from the original context in which the learners experienced the use of the tactic or strategy [75]. Therefore, for dashboards to unlock the full potential of learning analytics, they need to incorporate indicators of knowledge of tasks the learners are working on. This clearly indicates a need for building long-term learner profiles that gradually capture knowledge of tasks over time. Ideally, learner profiles that underlie user-centered learning analytics systems will go beyond a single learning context (e.g., course) and capture longitudinal changes in knowledge of tasks and knowledge of learning tactics and strategies. Even if a potential cold start problem may exist (e.g., enrolment into a university), learner profiles can still gradually be built as the students are progressing in their learning. Given the wealth of experience in building and using learning profiles, developers of user-centered learning analytics systems should pay close attention to the research done on open learner modelling [13].

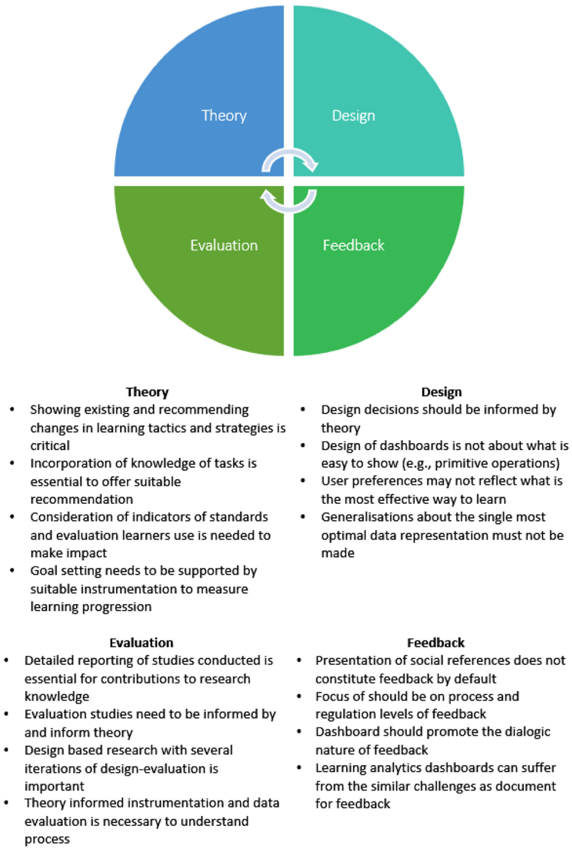


Fig. 12. Model for development, research, and evaluation of LADs (MLAD).

The total absence of support for standards and evaluation from the COPES model in contemporary LADs has several negative consequences. First, standards are used by learners to evaluate the products of their learning and the choices of their study strategies. If standards and evaluation are not supported, it is very unlikely that metacognitive monitoring and control support can effectively be supported too [14], [76]. Therefore, we question whether the choice of terminology used in the previous reviews of LADs [10], [11] can be considered adequate considering the lack of evidence found to support the references to metacognition. Second, standards underlie the definition of goals learners set according to the Winne and Hadwin model of SRL. Unsurprisingly, no support for standards in LADs was also associated with no support for phase 2 (goal setting) of the Winne and Hadwin [14] model of SRL. Likewise, no support for standards and evaluation led to rather limited support for phase 4 (adaptation) of the Winne and Hadwin model of SRL. Phase 4 is where metacognition is fully exercised in terms of the choices of learning tactics and strategies upon evaluating learning progression against standards encoded in learning goals. Given that our findings identified no support for acquisition of learning tactics and strategies, the potential of support for phase 4 is additionally limited. Therefore, future research and practice on user-centered learning analytics systems should prioritize the incorporation of indicators of standards and evaluation from the COPES model if the intention is to inform and enhance decisions of self-regulated learners, especially if the goal is to scaffold and enhance goal setting and adaptation.

It is difficult to capture indicators of standards and evaluation from the COPES models given that learning is driven by many internal and dynamic feedback loops [15]. Providing students with the features in learning environments and/or LADs to set their learning goals is a promising direction that can be used to capture standards of the student. However, goal setting functionality needs to be connected with a data collection mechanism that can track whether learners are working effectively towards their goals and how effectively they evaluate the products of their learning and choices of learning tactics and strategies. For example, Santos, Govaerts, Verbert, & Duval [50] presented a goal-oriented dashboard which allowed students to keep track if the learning goals had been achieved. Yet, their dashboard did not provide any mechanism for detailed collection of data that are related to standards based on which students evaluated their learning products and learning processes. That is, the presence of a goal setting functionality is not sufficient without the presence of a detailed instrumentation of the learning environments to collect data that can be used as feedback for a) enhancing students' evaluation against standards they set in the goals and b) informing the choices of study tactics and strategies. A promising example how user-centered learning analytics systems can benefit from more granular data that can support metacognitive monitoring and control is proposed by Marzouk *et al.* [77]. Marzouk *et al.* discussed the six scenarios in which user-centered learning analytics systems are informed by the learning sciences. One of their scenarios suggested goals should be presented in terms of indicators that promote the use of proven study tactics – note taking, summarization, and highlighting [24], [75]. Moreover, the Marzouk *et al.* approach suggests the importance of setting so-called SMART (specific, measurable, achievable, relevant, and time-bound) goals that can increase student motivation while maintaining their autonomy [28], [78], [79].

The weaknesses identified in theoretical underpinnings have implications for design, feedback, and evaluation dimensions of the MULAS.

B. Design

MULAS posits the importance of theory-informed design of user-centered learning analytics systems to maximize their effects on learning. Theory-informed design has implications on specific themes across which user-centered learning analytics systems, including LADs, are implemented, choices of indicators of learning that are incorporated to support the entire learning process, and the ways how information is presented to users.

The thematic analysis of the LADs showed significant limitations in the support of different elements of the learning process as theorized in the COPES model (see Table VI and Table VII). For example, dashboards that aim at informing the learning design cannot provide a complete insight to the stakeholders without taking into consideration aspects of task conditions. Currently, task conditions are only minimally presented at the level of instructional cues and resources, but they do not offer any insights into social context and time. For example, time defined for task completion of learning tasks and the order in which learners complete the tasks is essential to be considered if effectiveness of learning designs is to be assessed [80]. Learning difficulty in the current dashboards is considered without accounting for task conditions, knowledge of study tactics and strategies, knowledge of tasks, and motivational factors and orientations. Learning difficulty cannot adequately be assessed if these factors are not considered. Dashboards that considered affective states did so only in consideration of a single dimension of cognitive conditions (motivational elements and orientations) but without any dimension of cognitive conditions. Our entire sample of the dashboards across all the themes showed a complete shortage of attention paid to tactics and strategies students have acquired or performed. There was some limited attention paid to primitive operations mostly around the theme of learning progression but with little information across other themes.

Possible reasons for the weaknesses in the current LADs can be found in the design approaches followed in the selection of the indicators and formats for presentation. Most of the designs were based on the references to some of the previous studies, usually in learning analytics, and user-centered design methodologies (Figure 4). Although the references to the previously published papers and user-centered design for the indicator selection are important, there are some potential challenges that need to be considered. Previous research indicates that much of the existing literature on LADs has not been evaluated robustly [4], [7], [11]. Very few of those studies indeed used theory-informed decisions in the choices of their indicators, which triggered questions of validity [2], [5]. Relying on the input of the users only might not be the most robust approach either. The use of learner preferences only for adaptation of learning resources is commonly criticized in the literature given the insufficient evidence to support such approaches [81], [82]. Sometimes, users are not clear on what they can expected from the system, especially if they are not aware of a) possibilities learning analytics can afford or b) mechanisms how data can effectively be collected. Instead, we suggest that the choices of indicators should be theory-driven, while input of the users should be sought to understand the extent to which those indicators are practically useful to optimize learning and teaching. The discussion provided in Section V.A offers a comprehensive overview of the elements that need to be considered if the optimization of self-regulated learning, metacognition, and the overall learning process and outcomes is the target.

The major challenge for the designers of user-centered learning analytics systems is to address some of the potential tensions that may come from a combined use of user-centered approaches and theory. For example, while the use of effective study tactics and strategies has consistently been proven to promote effective learning, research has also shown that learners may not prefer to adopt them due to perceived difficulty in their use. This is the reason why they are commonly referred to as desirable difficulties [83]. Design of user-centered learning analytics systems, including LADs, should not only be informed by theories centered on cognition and metacognition but also by those that consider motivation dimensions that are also recognized to play a significant role in adoption of new learning tactics, strategies, and tools [74], [84], [85]. A promising direction is the use of self-determination theory that can inform the creation of conditions to motivate learners to engage with uninteresting tasks [86]. Self-determination is relevant to maintain and grow the sense of agency of learners by moving the decision making power to the learner while offering the reason why something can be beneficial for their learning [77]. This can also be a promising direction for the entire field of learning analytics to mitigate concerns suggesting that learning analytics may increase external control over and suppress the agency of learners.

The design of user-centered learning analytics systems, including LADs, should not make assumptions that only one representation of data universally works for all tasks as argued by Gašević *et al.* [2] by referencing the theory of cognitive fit [87]. Recent studies suggest that strong positive effects on learning outcomes and satisfaction with feedback can be achieved if analytics-based feedback is provided in form of weekly emails [88]. We recommend that future work on the design of user-centered learning analytics systems, including LADs, should consider multiple forms of information presentation in order to maximize the value of the insight provided by analytics.

C. *Feedback*

The results of this study suggest that the existing generation of dashboards is unlikely to meet any recommendations for effective feedback provided in the literature. Feedback can be defined in different ways and we highlight two recent definitions that effectively summarize the present body of research knowledge. According to Boud and Malloy [89] (p. 6), feedback is "a process whereby learners obtain information about their work in order to appreciate the similarities and differences between the appropriate standards for any given work, and the qualities of the work itself, in order to generate improved work". Carless [90] (p. 190) further extends this definition and suggest that feedback is "a dialogic process in which learners make sense of information from varied sources and use it to enhance the quality of their work or learning strategies". These two definitions clearly suggest that learners are agents, feedback is dialogical not unidirectional from educators to learners, and the use of standards and learning strategies is essential. As already established in the previous subsections, all these elements are either highly underrepresented or non-existent in the existing dashboards. As long as the design does not incorporate these elements, it is unlikely to expect user-centered learning analytics systems, including LADs, to provide potent and actionable feedback. For example, presentation of the number of logins or the number of posts can hardly offer sufficient guidance on the quality and strategy of learning [91]. It should also be stressed that the presence of reference frameworks (social, normative, or individual as shown in Figure 5 and [10]) is insufficient to consider LADs as feedback if reference frames are not grounded in and comprehensively capture relevant elements of the COPES model. That is, rather than focusing on what information is easily available, the provision of feedback should focus on what information is required in order to provide meaningful feedback to the students.

The recommendation for the designers of user-centered learning analytics systems, including LADs, is to use established frameworks for feedback such as the one proposed by Hattie and Timperley [92] who distinguish four levels of feedback: task, process, regulation, and self. While the literature indicates little support for the value of the self-level, the process and regulation levels of feedback are the most effective while the task level feedback is beneficial when combined with process and regulation levels of feedback. However, effective provision of feedback on the process and regulation levels can only be enabled if information about standards, evaluation, and learning tactics and strategies is considered [15]. As well, the design of particular features of a dashboard should address questions on all four levels of feedback: where am I going, how am I going, and where to next. Future work on user-centered learning analytics systems, including LADs, should consider recommendations provided by Pardo and colleagues [93], [94] on provision of feedback in data-rich environments. Finally, for user-centered learning analytics systems, including LADs, to exhibit the dialogic nature of feedback, some lessons learned from the literature in open learner modelling should be considered [13] by allowing users to update their user models when they potentially disagree or find discrepancies in data or results of data analysis. Not only will such an approach promote the agentic behavior of learners, promote reflection, and open the dialogue between learners and educators, but it can also increase the validity of learning analytics as an important side effect.

Research and the development of user-centered learning analytics systems can potentially face common issues as reported in the literature on feedback. The key challenge for future research is to study the extent to which learners understand and can act on feedback received through a dashboard or other representation of data/analytics. The literature suggests that good quality feedback reflects students' performance correctly, provides information about the task, and offers suggestions how to proceed or enhance learning [95], [96]. Feedback can achieve its potential benefits when students understand it and take some actions based on it. Very frequently, students struggle to make a clear interpretation from feedback externally provided [97]. As highlighted by many researchers, students who receive their feedback sometimes do not understand it, are not able to make use of it, or do not recognize benefits steaming from it [96]. The work by Corrin and de Barba [6] precisely highlights this lack of understanding of feedback communicated through dashboards that may affect even top performing learners.

D. Evaluation

Evaluation of the impact of user-centered learning analytics systems, including LADs, is an area that requires immediate attention. The studies on the impact of LADs as feedback are limited [17], [54]. Moreover, the evaluation on how LADs act as a mediator of feedback is under-explored. Most of existing research evaluates dashboards in terms of perceived usefulness. There were only a few studies that observed the real impact of implementing dashboards in field studies. As observed by Pardo and Khan [98], no significant association between the number of dashboard views and mid-term scores of students was found in a large enrolment computer engineering course. Kim *et al.*, [29] found that students who received feedback through a dashboard showed significantly higher final scores than those who did not receive dashboard feedback. However, the frequency of dashboard views had no significant association with performance. Similarly, Brouwer, Bredeweg, Latour, Berg, & Huizen [56] compared a group of students who received a dashboard intervention with another group that did not receive it. They found a significantly higher performance on the group of students who were provided with the dashboard, but there was no statistically significant association between the frequency of usage of the dashboard and performance.

The analysis of the quality of study design and reporting has revealed some significant limitations in the current studies. The most striking limitation is the total absence of discussion about generalizability, which should serve as a key source for informing future studies and inviting other researchers to pay attention in their future research endeavors. Even more importantly, the discussion of the study limitations is essential to inform practitioners who need to understand the extent to which the results reported in the studies are applicable to practice and be translated to policy. When this is coupled with very few studies that reported on their generalizability, the extent to which empirical studies on LADs can inform practice and policy is questionable. Therefore, there should be an expectation made that each study reporting empirical findings on the effects of the use of user-centered learning analytics systems, including LADs, provide a detailed discussion about study limitations and the extent to which study results can be generalized. We particularly refer to some technology-related fields such as empirical software engineering from which empirical research on user-centered learning analytics systems can learn and which have guidelines how results should be reported and threats to validity discussed [99], [100].

The analysis of empirical research revealed the shortage of the studies that provide some discussion on the implications for learning analytics research and interpretation of the results with respect to the existing literature and most importantly theory. This finding is consistent with the predominant atheoretical nature of the designs of the LADs. However, if research on user-centered learning analytics systems, including LADs, aims to contribute to the body of research knowledge that advances understanding of human learning and learning environments, not only do studies need to discuss its findings against the related literature, but the study objectives and research questions need to be informed by and advance relevant theories and previous research [2].

The shortage of experimental and mixed-method studies suggests that causal inferences cannot be made from the existing literature. Experimental studies or correlational studies complemented with qualitative methods to form mixed-methods afford for opportunities to identify specific effects that the user-centered learning analytics systems may have on particular learning processes or outcomes. Conducting experimental studies with user-centered learning analytics systems can be challenging due to ethical and practical implications. This is the space where we see a need for more attention to be paid to design-based research when developing and evaluating user-centered learning analytics systems [2], [101]. The papers included in this review did not offered any representative examples of design-based research on LADs. Design-based research would involve several iterations of designs where each iteration introduced a new intervention that was tested in practice [102]. Although most of the papers indeed were motivated to address problems in practice, their weak grounding in theory and ambition to advance the body of research on human learning would likely reduce the potential of existing studies to fold under the design-based research umbrella. Therefore, a strong integration of existing theory to inform the work on solving practical problems while conducting several design-evaluation cycles is a key recommendation for future empirical studies on user-centered learning analytics systems, including LADs.

The effects of LADs on different learning processes and outcomes have been discussed in previous learning analytics reviews [7], [11]. Our study corroborates their findings in terms of the general focus on quantitative measures of perceived usability and association with learning outcomes. Few studies attempted to assess the relationships between the use of LADs and learning outcomes or relevant cognitive conditions (e.g., motivation and approaches to learning) [43], [98]. However, even the findings of these studies cannot offer sufficient understanding about learning processes. The instrumentation and analysis did not extract indicators or proxies that were theoretically grounded to identify which exact learning processes and how were affected by certain components of LADs. The work presented by Siadaty and colleagues [103] offers a promising direction how technological interventions, including LADs, can be evaluated to assess impact on learning processes. Siadaty *et al.* proposed a theory-informed pre-analysis of digital traces in order to identify indicators of relevant process. In the case of the work proposed by Siadaty *et al.*, processes based on Zimmerman's [104] model of self-regulated learning were extracted from trace data. Once extracted from trace data, temporal relationships between such processes and interaction with technological interventions including LADs are studied with techniques from areas such as social network analysis, process mining, or sequence mining to reflect on the temporal nature of learning.

VI. Limitation

The primary limitation of this study is related to the searching process in which we restricted our search to papers that only contain term "dashboard". This could possibly exclude papers that did not explicitly use that particular term although could be considered related to LADs.

Second, we faced several challenges during the coding process because some of the papers provided insufficient information to describe LADs adequately. For example, there is a paper which indicated that the dashboard presented targeted both teachers and learners; however, the paper only provided description for the teacher's viewpoint. Furthermore, we often relied on the figures or screenshots of the dashboards to extract relevant indicators. Moreover, some papers may have not included complete information about the dashboards under study in the figures or screenshots included in the papers. Thus, we might have missed some important indicators in the review process.

Finally, there is a possible limitation related to the analysis of the reporting of the evaluation results. Some papers did mention the use of mixed method in conducting their study (e.g., survey and interview) but they only presented results of interviews and did not report the results of the survey analysis. Therefore, our information about the quality of study designs and reporting might not include complete information about the papers.

VII. Conclusion

The review provided in this paper highlighted significant limitations in the existing literature on LADs. The model of user-centered learning analytics systems (MULAS) is proposed to guide developers, researchers, evaluators, and practitioners in their endeavors that aim to understand and optimize learning and environments in which learning occurs. The model reinforces the need for strong grounding of user-centered learning analytics systems, including LADs, in the literature on learning processes, effective study methods, and feedback. Only when those aspects are systematically combined with user-centered design approaches, user-centered learning analytics systems are posited to provide effective support for learning. The review also emphasizes the need to grow rigor in the empirical evaluation of user-centered learning analytics systems, including LADs, especially in authentic learning contexts through several iterations where the use of design-based research offers a solid methodological foundation. It should be also acknowledged that the research on learning analytics requires strong interdisciplinary teams that can come with expertise in learning sciences, human-information interaction, design, and research methods. Although forming and coordinating such teams can often be a complex task, the proposed MULAS model offer some guidance for team competencies necessary to develop and evaluate user-centered learning analytics systems, including LADs.

Acknowledgment

This publication reflects the views only of the authors, and the Commission, the Agency, and ARC are not responsible for any use which may be made of the information contained therein.

APPENDIX

Table VIII
RESEARCH PAPERS INCLUDED IN THIS STUDY

Paper ID	Theme	Theory	Target User	Context
ACM15 [54]	Learning Design	No	Teacher	Blended Learning
ACM17 [39]	Emotion	TEA Model	Teacher, Student	Traditional Classroom
ACM18 [50]	Learning Progress	No	Teacher, Student	Computer Supported Collaborative Learning
ACM2017-10 [42]	Emotion	Discrete and Dimensional emotion theory (Automatic Emotion Recognition process)	Teacher	Online Learning
ACM2017-16 [57]	Learning Progress	No	Teacher	Online Learning and Blended Learning and Flipped Learning
ACM21 [46]	Competency	Fully-Embedded Assessment Model (FEAM)	Student, Faculty, Study Advisor	Online Learning
ACM23 [55]	Learning Progress	Open Student Model (OSM)	Teacher, Student	Blended Learning
ACM25 [47]	Teamwork Progress	Dickinson and McIntryre's teamwork model	Teacher	Blended Learning
ACM3 [98]	Learning Progress	No	Student	Flipped Learning
ACM30 [49]	Learning Progress	No	Student	Online Learning
ACM31 [59]	Learning Progress	No	Teacher	Blended Learning
ACM8 [43]	Learning Progress	Achievement Goal Orientation	Student	Blended Learning
GS5 [6]	Learning Progress	No	Student	Blended Learning
GS56 [60]	Learning Progress	No	Teacher	Intelligent Tutoring Systems (ITS)
IEEE1 [35]	Competency	Open Student Model (OSM)	Teacher, Student	Online Learning
IEEE19 [36]	Learning Difficulty Detection	Open Student Model; Activity-based Learner-Model (Engeström's Activity Theory and Actuator-Indicator Model)	Teacher, Student	Academic Supporting Tool
IEEE2017 – 11 [53]	Study Plan	No	Student; Study Adviser	Academic Supporting Tool
IEEE22 [37]	Learning Progress	Open Student Model; Activity-based Learner-Model (Engeström's Activity Theory and Actuator-Indicator Model)	Teacher, Student	Online Learning and Blended Learning
IEEE28 [105]	Learning Progress	No	Student	Blended Learning
SD1 [51]	Game-Based	No - Flow Theory; Analogical Reasoning Theory; Pragmatic Constraint Theory in the game development	Teacher, Student	Online Game-based Learning
SPL – 171 [106]	Learning Progress	No	Student, Teacher, Administrators, Faculty	MOOC
SPL – 174 [44]	Study Plan	Learning and Study Strategies Inventory (LASSI)	Student	Academic Supporting Tool
Springer 10 [29]	Learning Progress	No	Student	Online Learning
Springer 122 [52]	Game-Based	No	Teacher	Blended Learning
Springer 127 [107]	Learning Progress	No	Teacher, Student	Blended Learning
Springer 25 [56]	Learning Progress	No	Student	Blended Learning
Springer 87 [30]	Learning Progress	No	Student	Intelligent Tutoring Systems (ITS)
Springer 95 [61]	Learning Progress	No	Teacher, Student	Online Learning
Wiley1 [108]	Learning Progress	No	Teacher, Student	Blended Learning

REFERENCES

[1] G. Siemens and D. Gašević, “Guest editorial-learning and knowledge analytics,” *Educational Technol. Soc.*, vol. 15, no. 3, pp. 1–2, 2012.

[2] D. Gašević, V. Kovanović, and S. Joksimović, “Piecing the learning analytics puzzle: A consolidated model of a field of research and practice,” *Learn. Res. Pract.*, vol. 3, no. 1, pp. 63–78, 2017.

[3] W. Greller and H. Drachsler, “Translating learning into numbers: A generic framework for learning analytics author contact details,” *Educational Technol. Soc.*, vol. 15, no. 3, pp. 42–57, 2012.

[4] B. A. Schwendimann *et al.*, “Perceiving learning at a glance: A systematic literature review of learning dashboard research,” *IEEE Trans. Learn. Technol.*, vol. 10, no. 1, pp. 30–41, Jan./Mar. 2016.

[5] D. Gašević, S. Dawson, and G. Siemens, “Let’s not forget: Learning analytics are about learning,” *TechTrends*, vol. 59, no. 1, pp. 64–71, 2015.

[6] L. Corrin and P. De Barba, “Exploring students ’ interpretation of feedback delivered through learning analytics dashboards,” in *Proc. Austrlas. Soc. Comput. Learn. Tertiary Educ.*, 2014, pp. 629–633.

[7] R. Bodily and K. Verbert, “Trends and issues in student-facing learning analytics reporting systems research,” in *Proc. 7th Int. Learn. Analytics Knowl. Conf.*, 2017, vol. 1, no. 212, pp. 309–318.

[8] G. Sedrakyan, S. Järvelä, and P. Kirschner, “Conceptual framework for feedback automation and personalization for designing learning analytics dashboards,” in *Proc. 2016 Conf. EARLI SIG 27 Processes matter – Measures, Anal. Appl. Tracing Learn.*, Oulu, Finland, 2016, pp. 1–3.

[9] L. Corrin and P. de Barba, “How do students interpret feedback delivered via dashboards?,” in *Proc. Int. Conf. Learn. Anal. Knowl.*, 2015, pp. 430–431.

[10] I. Jivet, M. Scheffel, H. Drachsler, and M. Specht, “Awareness is not enough. pitfalls of learning analytics dashboards in the educational practice,” in *Proc. Eur. Conf. Technol. Enhanced Learn.*, 2017, pp. 82–96.

[11] I. Jivet, M. Scheffel, M. Specht, and H. Drachsler, “License to evaluate: Preparing learning analytics dashboards for educational practice,” in *Proc. 8th Int. Conf. Learn. Analytics Knowl.*, 2018, pp. 31–40.

[12] R. Bodily and K. Verbert, “Review of research on student-facing learning analytics dashboards and educational recommender systems,” *IEEE Trans. Learn. Technol.*, vol. 10, no. 4, pp. 405–418, Oct. 2017.

[13] R. Bodily J. Kay, V. Aleven, I. Jivet, D. Davis, F. Xhakaj, and K. Verbert, “Open learner models and learning analytics dashboards: A systematic review,” in *Proc. 8th Int. Conf. Learn. Analytics Knowl.*, 2018, pp. 41–50.

[14] P. H. Winne and A. F. Hadwin, "Studying as self-regulated learning," *Metacognition Educational Theory Pract.*, vol. 93, pp. 277–304, 1998.

[15] D. L. Butler and P. H. Winne, "Feedback and self-regulated learning: A theoretical synthesis," *Rev. Educational Res.*, vol. 65, no. 3, pp. 245–281, 1995.

[16] K. Verbert, S. Govaerts, E. Duval, G. Parra, and J. Klerkx, "Learning dashboards: An overview and future research opportunities," *Pers. Ubiquitous Comput.*, vol. 18, pp. 1499–1514, 2014.

[17] K. Verbert, E. Duval, J. Klerkx, S. Govaerts, and J. L. Santos, "Learning analytics dashboard applications," *Amer. Behavioral Sci.*, vol. 10, pp. 1500–1509, 2013.

[18] A. F. Wise, "Designing pedagogical interventions to support student use of learning analytics," in *Proc. 4th ACM Int. Conf. Learn. Analytic Knowl.*, 2014, pp. 203–211.

[19] D. H. Schunk, "Metacognition, self-regulation, and self-regulated learning: Research recommendations," *Educational Psychol. Rev.*, vol. 20, no. 4, pp. 463–467, 2008.

[20] M. Boekaerts, "Self-regulated learning at the junction of cognition and motivation," *Eur. Psychol.*, vol. 1, no. 2, pp. 100–112, 1996.

[21] E. Panadero, "A review of self-regulated learning: Six models and four directions for research," *Frontiers Psychol.*, vol. 8, pp. 1–28, 2017.

[22] E. Panadero, J. Klug, and S. Järvelä, "Third wave of measurement in the self- regulated learning field: When measurement and intervention come hand in hand," *Scandinavian J. Educational Res.*, vol. 60, no. 6, pp. 723–735, 2016.

[23] J. Malmberg, J. Sanna, and P. A. Kirschner, "Elementary school students' strategic learning: Does task-type matter?," *Metacognition Learn.*, vol. 9, no. 2, pp. 113–136, 2014.

[24] R. A. Bjork, J. Dunlosky, and N. Kornell, "Self-regulated learning: Beliefs, techniques, and illusions," *Annu. Rev. Psychol.*, vol. 64, no. 1, pp. 417–444, 2013.

[25] P. H. Winne and D. Jamieson-Noel, "Exploring students' calibration of self reports about study tactics and achievement," *Contemporary Educational Psychol.*, vol. 27, no. 4, pp. 551–572, 2002.

[26] P. H. Winne and D. Jamieson-Noel, "Self-regulating studying by objectives for learning: Students' reports compared to a model," *Contemporary Educational Psychol.*, vol. 28, no. 3, pp. 259–276, 2003.

[27] D. J. Nicol and D. Macfarlane-Dick, "Formative assessment and self-regulated learning: A model and seven principles of good feedback practice," *Studies Higher Educ.*, vol. 31, no. 2, pp. 199–218, 2006.

[28] G. Sedrakyan, J. Malmberg, K. Verbert, S. Järvelä, and P. A. Kirschner, "Linking learning behavior analytics and learning science concepts: Designing a learning analytics dashboard for feedback to support learning regulation," *Comput. Human Behav.*, May 2018. https://doi.org/10.1016/j.chb.2018.05.004

[29] J. Kim, I. Jo, and Y. Park, "Effects of learning analytics dashboard: Analyzing the relations among dashboard utilization, satisfaction, and learning achievement," *Asia Pacific Educ. Rev.*, vol. 17, no. 1, pp. 13–24, 2016.

[30] K. Muldner, M. Wixon, D. Rai, W. Burleson, B. Woolf, and I. Arroyo, "Exploring the impact of a learning dashboard on student affect," in *Proc. Int. Conf. Artif. Intell. Educ.*, 2015, pp. 307–317.

[31] B. Kitchenham and S. Charters, "Guidelines for performing systematic literature reviews in software engineering," Keele University and University of Durham, EBSE Technical Report EBSE-2007-01, 2007.

[32] D. A. Reed, D. A. Cook, T. J. Beckman, D. E. Kern, and S. M. Wright, "Association between funding and quality of published medical education research," *J. Amer. Med. Assoc.*, vol. 298, no. 9, pp. 1002–1009, Sep. 2007.

[33] R. M. Bernard, P. Abrami, Y. Lou, and E. Borokhovski, "A methodological morass? How we can improve quantitative research in distance education," *Distance Educ.*, vol. 25, no. 2, pp. 175–198, Oct. 2004.

[34] A. Wise and D. W. Shaffer, "Why theory matters more than ever in the age of big data," *J. Learn. Anal.*, vol. 2, no. 2, pp. 5–13, Dec. 2015.

[35] C. Chou S.F. Tseng *et al.*, "Open student models of core competencies at the curriculum level: Using learning analytics for student reflection," *IEEE Trans. Emerg. Topics Comput.*, vol. 5, no. 1, pp. 32–44, Jan./Mar. 2017.

[36] C. Mejia, B. Florian, R. Vatrapu, S. Bull, S. Gomez, and R. Fabregat, "A novel web-based approach for visualization and inspection of reading difficulties on university students," *IEEE Trans. Learn. Technol.*, vol. 10, no. 1, pp. 53–67, Jan./Mar. 2017.

[37] B. Florian-Gaviria, C. Glahn, and R. Fabregat Gesa, "A software suite for efficient use of the European qualifications framework in online and blended courses," *IEEE Trans. Learn. Technol.*, vol. 6, no. 3, pp. 283–296, Jul./Sep. 2013.

[38] Y. Engeström, *Learning by Expanding: An Activity-Theoretical Approach to Developmental Research*. Cambridge, UK: Cambridge Univ. Press, 2014.

[39] S. Ruiz, S. Charleer, I. Fernández-castro, and E. Duval, "Supporting learning by considering emotions: Tracking and visualization. A case study," in *Proc. 6th Int. Conf. Learn. Analytical. Knowl.* , 2016, pp. 254–263.

[40] R. Pekrun, T. Goetz, A. C. Frenzel, P. Barchfeld, and R. P. Perry, "Measuring emotions in students' learning and performance: The achievement emotions questionnaire (AEQ)," *Contemporary Educational Psychol.*, vol. 36, no. 1, pp. 36–48, 2011.

[41] I. Arroyo, D. G. Cooper, W. Burleson, B. P. Woolf, K. Muldner, and R. Christopherson, "Emotion sensors go to school," in *Proc. Conf. Artif. Intell. Educ.*, 2009,vol. 200, pp. 17–24.

[42] M. Ez-zaouia and E. Lavou, "EMODA: A tutor oriented multimodal and contextual emotional dashboard," in *Proc. 7th Int. Learn. Analytical Knowl. Conf.*, 2017, pp. 429–438.

[43] S. S. Beheshitha, M. Hatala, D. Gašević, and S. Joksimović, "The role of achievement goal orientations when studying effect of learning analytics visualizations," in *Proc. 6th Int. Conf. Learn. Analytic Knowl.*, 2016, pp. 54–63.

[44] T. Broos, L. Peeters, K. Verbert, C. Van Soom, G. Langie, and T. De Laet, "Dashboard for actionable feedback on learning skills: Scalability and usefulness," in *Proc. Int. Conf. Learn. Collaboration Technol.*, 2017, pp. 229–241.

[45] C. E. Weinstein, D. Palmer, and A. C. Schulte, *Learning and Study Strategies Inventory*. Clearwater, FL, USA: H & H Publishing, 1987.

[46] J. Grann and D. Bushway, "Competency map: Visualizing student learning to promote student success," in *Proc. 4th Int. Conf. Learn. Analytic Knowl.*, 2014, pp. 168–172.

[47] H. Tarmazdi, R. Vivian, C. Szabo, K. Falkner, and N. Falkner, "Using learning analytics to visualize computer science teamwork," in *Proc. ACM Conf. Innov. Technol. Comput. Sci. Educ.*, 2015, pp. 165–170.

[48] T. L. Dickinson and R. M. McIntyre, "A conceptual framework for teamwork measurement," in *Team Performance Assessment and Measurement: Theory, Methods, and Applications*, Mahwah, NJ, USA: Lawrence Erlbaum Associates Publishers, 1997, pp. 19–43.

[49] J. P.-L. Tan, S. Yang, E. Koh, and C. Jonathan, "Fostering 21st century literacies through a collaborative critical reading and learning analytics environment," in *Proc. 6th Int. Conf. Learn. Analytic Knowl.*, 2016, pp. 430–434.

[50] J. L. Santos, S. Govaerts, K. Verbert, and E. Duval, "Goal-oriented visualizations of activity tracking: A case study with engineering students," in *Proc. 2nd Int. Conf. Learn. Analytics Knowl.*, 2012, pp. 143–152.

[51] D. D. Reese, "Chapter 11 – Dashboard effects challenge flow-learning assumption in digital instructional games," in *Emotions, Technology, and Digital Games*, S. Y. Tettegah and W. D. Huang, Eds., San Diego: Academic Press, 2016, pp. 231–287.

[52] S. Charleer, J. L. Santos, J. Klerkx, and E. Duval, "Improving teacher awareness through activity, badge and content visualizations," in *Proc. Int. Conf. Web-Based Learn.*, 2014, pp. 143–152.

[53] S. Charleer, A. Vande Moere, J. Klerkx, K. Verbert, and T. De Laet, "Learning analytics dashboards to support adviser-student dialogue," *IEEE Trans. Learn. Technol.*, vol. 11, no. 3, pp. 389–399, Jul./Sep. 2018.

[54] I.-H. Hsiao, S. K. Pandhalkudi Govindarajan, and Y.-L. Lin, "Semantic visual analytics for today's programming courses," in *Proc. 6th Int. Conf. Learn. Analytics Knowl.*, 2016, pp. 48–53.

[55] J. L. Santos, K. Verbert, S. Govaerts, and E. Duval, "Addressing learner issues with StepUp!," in *Proc. 3rd Int. Conf. Learn. Analytics Knowl.*, 2013, pp. 14–22.

[56] N. Brouwer, B. Bredeweg, S. Latour, A. Berg, and G. Van Der Huizen, "Learning analytics pilot with coach2 - searching for effective mirroring," in *Proc. Eur. Conf. Technol. Enhanced Learn.*, 2016, vol. 2, pp. 363–369.

[57] S. L. Dazo, N. R. Stepanek, A. Chauhan, and B. Dorn, "Examining instructor use of learning analytics," in *Proc. CHI Conf. Extended Abstr. Human Factors Comput. Syst.*, 2017, pp. 2504–2510.

[58] S. Park, "Examining learning experience in two online courses using web logs and experience sampling method (ESM)," in *The Des. Learn. Experience: Creating the Future of Educational Technol.*, B. Hokanson, G. Clinton, and M. W. Tracey, Eds., Cham: Springer International Publishing, 2015, pp. 269–287.

[59] E. Van Alphen and S. Bakker, "Lernanto: Using an ambient display during differentiated instruction," in *Proc. CHI Conf. Extended Abstr. Human Factors Comput. Syst.*, 2016, pp. 2334–2340.

[60] V. Aleven, F. Xhakaj, K. Holstein, and B. M. McLaren, "Developing a teacher dashboard for use with intelligent tutoring systems," *Technology*, vol. 34, pp. 44–50, 2010.

[61] Y. Park and I. Jo, "Development of the learning analytics dashboard to support students ' learning performance learning analytics dashboards (LADs)," *J. Universal Comput. Sci.*, vol. 21, no. 1, pp. 110–133, 2015.

[62] A. Ramos-Soto, M. Lama, B. Vazquez-Barreiros, A. Bugarin, M. Mucientes, and S. Barro, "Towards textual reporting in learning analytics dashboards," in *Proc. IEEE 15th Int. Conf. Adv. Learn. Technol.*, 2015, pp. 260–264.

[63] L. Corrin, G. Kennedy, and R. Mulder, "Enhancing learning analytics by understanding the needs of teachers," in *Proc. 30th Annu. Conf. Australian Soc. Comput. Learn. Tertiary Educ.*, 2013, pp. 201–205.

[64] M. Neelen and P. A. Kirschner, "Where Are the Learning Sciences in Learning Analytics Research?," 3-Star learning experiences, Oct. 2017. https://3starlearningexperiences.wordpress.com/2017/10/17/where-are-the-learning-sciences-in-learning-analytics-research/

[65] S. Chaturapruek, T. Dee, R. Johari, R. F. Kizilcec, and M. L. Stevens, "How a data-driven course planning tool affects college students' GPA: Evidence from two field experiments," in *Proc. 5th ACM Conf. Learn. Scale*, 2018, pp. 63:1–63:10.

[66] S. Lonn, S. J. Aguilar, and S. D. Teasley, "Investigating student motivation in the context of a learning analytics intervention during a summer bridge program," *Comput. Human Behav.*, vol. 47, pp. 90–97, 2015.

[67] P. D. Long, G. Siemens, G. Conole, and D. Gašević, Eds., *Proceedings of the 1st International Conference on Learning Analytics and Knowledge*. New York, NY, USA: ACM, 2011.

[68] J. Dunlosky, K. A. Rawson, E. J. Marsh, M. J. Nathan, and D. T. Willingham, "Improving students' learning with effective learning techniques promising directions from cognitive and educational psychology," *Psychol. Sci. Public Interest*, vol. 14, no. 1, pp. 4–58, Jan. 2013.

[69] G. Clarebout, J. Elen, N. A. J. Collazo, G. Lust, and L. Jiang, "Metacognition and the use of tools," in *International Handbook of Metacognition and Learning Technologies*, R. Azevedo and V. Aleven , Eds. Berlin, Germany: Springer, 2013, pp. 187–195.

[70] J. McCabe, "Metacognitive awareness of learning strategies in undergraduates," *Memory Cognition*, vol. 39, no. 3, pp. 462–476, Apr. 2011.

[71] M. S. Boroujeni and P. Dillenbourg, "Discovery and temporal analysis of latent study patterns in MOOC interaction sequences," *Proc. 8th Int. Conf. Learn. Analytics Knowl.*, 2018, pp. 206–215.

[72] O. E. Fincham, D. V. Gasevic, J. M. Jovanovic, and A. Pardo, "From study tactics to learning strategies: An analytical method for extracting interpretable representations," *IEEE Trans. Learn. Technol.*, vol. 12, no. 1, pp. 59–72, Jan./Mar. 2019.

[73] J. Jovanovic, D. Gasevic, S. Dawson, A. Pardo, and N. Mirriahi, "Learning analytics to unveil learning strategies in a flipped classroom," *Internet Higher Educ.*, vol. 33, pp. 74–85, 2017.

[74] P. H. Winne, "How software technologies can improve research on learning and bolster school reform," *Educational Psychol.*, vol. 41, no. 1, pp. 5–17, 2006.

[75] P. H. Winne, "Leveraging big data to help each learner and accelerate learning science," *Teacher College Rec.*, vol. 119, no. 3, pp. 1–24, 2017.

[76] P. H. Winne, "A metacognitive view of individual differences in self-regulated learning," *Learn. Individual Differences*, vol. 8, no. 4,pp. 327–353, 1996.

[77] Z. Marzouk et al., "What if learning analytics were based on learning science?," *Australas. J. Educational Technol.*, vol. 32, no. 6, pp. 1–18, Dec. 2016.

[78] A. E. Black and E. L. Deci, "The effects of instructors' autonomy support and students' autonomous motivation on learning organic chemistry: A self-determination theory perspective," *Sci. Educ.*, vol. 84, no. 6, pp. 740–756, 2000.

[79] E. A. Locke and G. P. Latham, "New directions in goal-setting theory," *Current Directions Psychol. Sci.*, vol. 15, no. 5, pp. 265–268, Oct. 2006.

[80] R. F. Kizilcec, M. Pérez-Sanagustín, and J. J. Maldonado, "Self-regulated learning strategies predict learner behavior and goal attainment in massive open online courses," *Comput. Educ.*, vol. 104, pp. 18–33, 2017.

[81] P. A. Kirschner and J. J. G. van Merriënboer, "Do learners really know best? urban legends in education," *Educational Psychol.*, vol. 48, no. 3, pp. 169–183, Jul. 2013.

[82] H. Pashler, M. McDaniel, D. Rohrer, and R. A. Bjork, "Learning styles: Concepts and evidence," *Psychol. Sci. Public Interest*, vol. 9, no. 3, pp. 105–119, Dec. 2008.

[83] E. L. Bjork and R. A. Bjork, "Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning," in *Psychology and the Real World: Essays Illustrating Fundamental Contributions to Society*, M. A. Gernsbacher, R. W. Pew, L. M. Hough, and J. R. Pomerantz, Eds., New York, NY, US: Worth Publisher, pp. 56–64, 2011.

[84] D. Gašević, N. Mirriahi, S. Dawson, and S. Joksimović, "Effects of instructional conditions and experience on the adoption of a learning tool," *Comput. Human Behav.*, vol. 67, pp. 207–220, 2017.

[85] M. Y. Yi and Y. Hwang, "Predicting the use of web-based information systems: Self-efficacy, enjoyment, learning goal orientation, and the technology acceptance model," *Int. J. Human Comput. Studies*, vol. 59, no. 4, pp. 431–449, Oct. 2003.

[86] H. Jang, "Supporting students' motivation, engagement, and learning during an uninteresting activity," *J. Educational Psychol.*, vol. 100, no. 4, pp. 798–811, 2008.

[87] I. Vessey, "Cognitive fit: A theory-based analysis of the graphs versus tables literature*," *J. Decis. Sci. Inst.*, vol. 22, no. 2, pp. 219–240, Mar. 1991.

[88] A. Pardo, J. Jovanovic, S. Dawson, D. Gašević, and N. Mirriahi, "Using learning analytics to scale the provision of personalised feedback," *Brit. J. Educational Technol.*, vol. 50, no. 1, pp. 128–138, 2019.

[89] D. Boud and E. Molloy, "Rethinking models of feedback for learning: the challenge of design," *Assessment Eval. Higher Educ.*, vol. 38, no. 6, pp. 698–712, Sep. 2013.

[90] D. Carless, *Excellence in University Assessment: Learning From Award-Winning Practice*. Evanston, IL, USA: Routledge, 2015.

[91] V. Kovanović, D. Gašević, S. Dawson, S. Joksimović, R. S. Baker, and M. Hatala, "Does time-on-task estimation matter? Implications for the validity of learning analytics findings," *J. Learn. Analytics*, vol. 2, no. 3, pp. 81–110, 2015.

[92] J. Hattie and H. Timperley, "The power of feedback," *Rev. Educational Res.*, vol. 77, no. 1, pp. 81–112, 2007.

[93] A. Pardo, "A feedback model for data-rich learning experiences," *Assessment Eval. Higher Educ.*, vol. 43, no. 3, pp. 428–438, 2018.

[94] A. Pardo, O. Poquet, R. Martinez-Maldonado, and S. Dawson, "Provision of data-driven student feedback in LA & EDM," in *Handbook Learn. Anal., 1st ed.*, C. Lang, G. Siemens, A. Wise, and D. Gašević, Eds., Beaumont, AB, Canada: Society for Learning Analytics Research, 2017, pp. 163–174.

[95] A. Forsythe and S. Johnson, "Thanks, but no-thanks for the feedback," *Assessment Eval. Higher Educ.*, vol. 42, no. 6, pp. 850–859, 2017.

[96] C. Evans, "Making sense of assessment feedback in higher education," *Rev. Educational Res.*, vol. 83, no. 1, pp. 70–120, 2013.

[97] M. Price, K. Handley, J. Millar, and B. O'Donovan, "Feedback: All that effort, but what is the effect?," *Assessment Eval. Higher Educ.*, vol. 35, no. 3, pp. 277–289, 2010.

[98] I. Khan and A. Pardo, "Data2U: Scalable real time student feedback in active learning environments," in *Proc. 6th Int. Conf. Learn. Analytics Knowl.*, 2016, pp. 249–253.

[99] A. Jedlitschka and D. Pfahl, "Reporting guidelines for controlled experiments in software engineering," in *Proc. Int. Symp. Empirical Softw. Eng.*, 2005, pp. 95–104.

[100] P. Runeson and M. Höst, "Guidelines for conducting and reporting case study research in software engineering," *Empirical Softw. Eng.*, vol. 14, no. 2, Apr. 2009, Art. no. 131.

[101] P. Reimann, "Connecting learning analytics with learning research: the role of design-based research," *Learn. Res. Pract.*, vol. 2, no. 2, pp. 130–142, Jul. 2016.

[102] T. Anderson and J. Shattuck, "Design-based research a decade of progress in education research?," *Educational Res.*, vol. 41, no. 1, pp. 16–25, Jan. 2012.

[103] M. Siadaty, D. Gašević, and M. Hatala, "Associations between technological scaffolding and micro-level processes of self-regulated learning: A workplace study," *Comput. Human Behav.*, vol. 55, no. Part B, pp. 1007–1019, Feb. 2016.

[104] B. J. Zimmerman, "Theories of self-regulated learning and academic achievement: An overview and analysis," in *Self-Regulated Learning and Academic Achievement,* 2nd ed., B. J. Zimmerman and D. H. Schunk, Eds. Mahwah, NJ, USA: Lawrence Erlbaum Associates, 2001, pp. 1–37.

[105] N. R. Aljohani and H. C. Davis, "Learning analytics and formative assessment to provide immediate detailed feedback using a student centered mobile dashboard," in *Proc. 7th Int. Conf. Next Gener. Mobile Apps, Services Technol.*, 2013, pp. 262–267.

[106] L. Vigentini, A. Clayphan, X. Zhang, and M. Chitsaz, "Overcoming the MOOC data deluge with learning analytic dashboards," in *Learning Analytics: Fundaments, Applications, and Trends: A View of the Current State of the Art to Enhance e-Learning*, A. Peña-Ayala, Ed. Cham: Springer International Publishing, 2017, pp. 171–198.

[107] V. Kumar and D. Boulanger, "Competence analytics," *J. Comput. Educ.* vol. 1, pp. 251–270, 2014.

[108] A. Ramos-Soto, B. Vazquez-Barreiros, A. Bugarin, A. Gewerc, and S. Barro , "Evaluation of a data-to-text system for verbalizing a learning analytics dashboard," *Int. J. Intell. Syst.*, vol. 32, no. 2, pp. 177–193, 2014.

Wannisa Matcha is currently working toward the Ph. D. degree at School of Informatics, University of Edinburgh, Edinburgh, Scotland. Her research interests center on learning analytics, learning strategy, self-regulated learning, and technology enhanced learning to promote the behavioral changes and the adoption of effective learning strategies.

Dragan Gašević is a Professor of learning analytics with the Faculty of Information Technology, Monash University, Melbourne, VIC, Australia, and an Honorary Professor with the School of Informatics, University of Edinburgh, Edinburgh, Scotland. His research in learning analytics involves developing computational methods that can shape next-generation learning technologies and advance our understanding of information seeking, sense-making, and self-regulated and collaborative learning.

Nora'ayu Ahmad Uzir is currently working toward the Ph.D. degree at the School of Informatics, University of Edinburgh, Edinburgh, Scotland, and is a Lecturer with the Faculty of Information Management, Universiti Teknologi MARA, Shah Alam, Malaysia. Her research interests revolve around learning analytics and self-regulated learning to promote effective time management behaviors among online learners and to provide insight towards the provision of learning analytic based feedback at scale.

Abelardo Pardo is a Professor and a Academic Dean with the Division of Information Technology, Engineering and the Environment, University of South Australia, Adelaide, SA, Australia. He specializes in technology enhanced learning with an emphasis on learning and behavioral analytics, computer supported collaborative learning, and personalization of learning experiences.

2.3 Summary

In this chapter, the systematic literature review is presented with the main objective of investigating how learning analytics-based tools have been used to support the SRL. To ensure the quality of the analysis of the systematic literature review, the COPES model developed by Winne and Hadwin (1998) is used as a focal point of the analysis. The main findings of this chapter are summarised below, and thus, outlines how research question one (RQ1) of the thesis has been addressed accordingly.

The important contributions of the work presented in this chapter are multifold. First, the analysis of the empirical research studies provides an overview of existing LADs research. The results show that the selection of indicators in these LADs is primarily based on previous works and users centre approaches. The majority of these studies failed to incorporate theories to support the design and development (68% of the included studies did not refer to any theory). In general, LAD feedback aims to provide support during the task identification (phase 1), and enactment of learning tactics and strategies (phase 3) of the SRL process but it fails to provide necessary information regarding choices of learning tactics and strategies and how to improve them. The type of reference frames used include the individual reference frames (i.e. by presenting individual feedback to individual students), and social reference frame (i.e. by offering a comparison with group average and/or peers). Existing LADs focus on employing simple and basic visualisations such as bar charts to ease the interpretation of the results. The main sources of data collection are questionnaires, trace data, and interviews.

Second, this systematic literature review discusses current limitations of LADs that require close attention from researchers and practitioners. In particular, there is a lacuna in the use of theory to support the design and development of LADs. Hence, it is unsurprising that much research reported on ineffective use of LADs. Moreover, several of the learning constructs related to SRL have been neglected. For instance, standards and evaluations from the COPES model have not been considered. This is due to the iteration process of feedback provision and the difficulty of capturing the standards and expectation set by students. The knowledge of tasks is among one of the important cognitive conditions that current LADs fail to provide necessary feedback. Most importantly, based on the results obtained from the systematic literature review, we find that students have not received due guidance on how to select and employ effective learning tactics and strategies. Learning tactics and strategies are central to SRL processes (Zimmerman, 2011). Failing to offer feedback on learning tactics and strategies makes it difficult or impossible to enhance self-regulation. The quality of existing research thus requires improvement. Particularly, the empirical evaluations of LADs were not discussed by considering the current research literature and theory. Moreover, the generalisability and implications of research are generally not explicitly discussed.

Finally, the literature review proposes several directions that can be employed to bridge the highlighted research gaps. For instance, LADs should provide students with functionalities to set learning goals and track their performance against those goals. Data mining techniques can be used to capture relevant information about learning tactics and strategies. Hence, necessary feedback can be provided to guide the selection of learning tactics and strategies. In terms of the design of LADs, research should incorporate theories to inform choices of indicators. In relations to the use of LADs to deliver feedback, the design of dashboards should ensure that learners understand what insights are provided. Moreover, the indicators should suggest "actionable" feedback. Frameworks of good practice for feedback such as the one proposed by Hattie and Timperley (2007) can be used to guide the design of the dashboards. In terms of the quality of evaluations, current research primarily focuses on perceived usefulness of and satisfaction with LADs. Evaluation of the impact on behavioural change when offered with feedback through LADs is needed. Discussion of generalisability and implications of research on LADs is needed. This is fundamentally important for the maturity of research on LADs (Baker, 2019).

The findings presented in this chapter highlight the gap in feedback provision on the application of learning tactics and strategies which are denoted as essential components of SRL (Winne, 2013; Zimmerman, 2011), especially, during phase 3: enactment of learning tactics and strategies (Winne & Hadwin, 1998). One of the barriers that prevent timely feedback provision on learning tactics and strategies is the difficulty of automatic detection based on data about learners' actions (Jovanovic et al., 2017). Hence, the next step for this thesis focuses on the data science perspective and address research questions two to four (RQ2-RQ4) of the thesis. Specifically, the remainder of the thesis aims to develop a novel approach to analyse trace data to detect the learning tactics and strategies by using data mining and machine learning algorithms.

3 Analytics of Learning Strategies: Automatic Detection of Learning Tactics and Strategies

The key to success is action
— Brian Tracy, *Eat That Frog!: 21 Great Ways to Stop Procrastinating and Get More Done in Less Time*

3.1 Introduction

ADOPTION of effective learning tactics and strategies has been identified as a core process of learning (Zimmerman, 2011). However, current research reports that students lack the skills to select and adapt learning strategies that best suit a learning situation (Lust et al., 2013; Rachal et al., 2007). Moreover, students frequently use sub-optimal learning tactics and strategies (Dunlosky et al., 2013; Winne & Jamieson-Noel, 2003). Hence, guiding students to employ learning tactics and strategies identified as effective in the literature (Dunlosky et al., 2013) is crucial for the success of learning. Nonetheless, providing timely feedback to guide the selection of learning tactics and strategies is challenging (Pardo et al., 2017). Rarely do students receive feedback on their application of learning tactics and strategies (Matcha, Ahmad Uzir, et al., 2020). One of the barriers preventing feedback provision on learning tactics and strategies is the detection of learning tactics and strategies used by students in learning activities (Jovanovic et al., 2017).

Traditional research employs self-report instruments to capture learning tactics and strategies. Questionnaires are among the most frequently used instruments. However, the drawbacks of questionnaires are widely acknowledged. That is, learners are not always accurate in reporting and judging how they learn (Bjork et al., 2013; Broadbent, 2017; Carpenter et al., 2020). Self-reports inherently capture learners' perceptions, reflect intentions of using learning tactics and strategies rather than the actual use during their learning (Zhou & Winne, 2012). Moreover, self-reports are usually applied to specific points in time (e.g., before or after learning), hence, are not effective at capturing the "dynamics" of what happens during a learning process.

Trace data record the actual interaction of learners with learning resources together with timestamps. Using trace data, therefore, enables us to observe the actual behaviour and the dynamic nature of learning. However, without adequate data analytics approaches, it is hard to uncover potential patterns hidden in trace data.

Learning analytics employs data mining and machine learning techniques to process large amounts of educational data to discover knowledge and patterns (Gašević et al., 2015; Siemens & Baker, 2012). Even though research on the analysis of learning behaviour from trace data is growing, research focused on the detection of learning tactics and strategies from the trace data is scarce. Moreover, recent studies often used learning tactics and strategies interchangeably (Derry, 1989; Malmberg et al., 2014). However, these two terms are different in the literature as discussed in the introduction of the thesis (see Section 1). In this PhD thesis, learning tactic refers to a sequence of actions performed by a learner to complete a learning task (Hadwin et al., 2007). Learning strategy involves the selection and coordination of a set of learning tactics to achieve a learning goal (Derry, 1989; Malmberg et al., 2014; Rachal et al., 2007).

Among the variety of data mining and machine learning techniques, three approaches are frequently used in the current literature that aims to analyse learning patterns, namely, sequence mining, process mining, and network analysis. This chapter examines how the three data analytics approaches can be used to capture learning patterns indicative of tactics used within learning sessions. The chapter also explores how each of the three approaches enables the detection of (dis)similar learning tactics and strategies. In accordance with the consolidated model of learning analytics as identified in Figure 3 presented in Chapter two, the results of the study presented in this chapter are discussed based on a well-known theory related to learning strategies – approaches to learning (Biggs, 1987; Entwistle, 1991; Marton & Säljö, 1976).

3.1.1 Approaches to learning

A closely related learning construct used to explain learning strategies is approaches to learning (Biggs, 1987; Kovanović et al., 2019; Marton & Säljö, 1976; Trigwell & Prosser, 1991). Approaches to learning posit that students may take different approaches to complete learning activities, depending on the learning context, learning tasks, and intention of the students (Entwistle, 2007). This notion is well-aligned with the SRL concept in that tasks and cognitive conditions influence the learning process (Winne & Hadwin, 1998). Entwistle (1991) and Biggs (1987) characterised approaches to learning into three types, including, surface, deep, and strategic learning.

- Surface: learners simply intend to complete only the compulsory tasks (Biggs, 1987; Emilia et al., 2012). They employ a shallow learning effort, rely on rote learning (Biggs, 1987), and show a low level of engagement with learning content (Fincham et al., 2018). Students who employ the surface approach to learning often focus on the assessment with little understanding of learning content (Biggs, 1987; Gašević, Jovanović, et al., 2017). Research has found that surface approach to learning is associated with low academic performance (Chonkar et al., 2018; Emilia et al., 2012; Mattick et al., 2004; Trigwell & Prosser, 1991).

- Deep: opposite to surface approach, learners seek to understand the meaning of the content studied (Biggs, 1987), hence, learners who follow the deep approach to learning apply a variety of learning tactics to complete learning tasks (Fincham et al., 2018; Gašević, Jovanović, et al., 2017). They are highly active learners who put much effort into understand learning content. Research reports positive relationships of applying the deep approach to learning and high academic performance (Chonkar et al., 2018; Emilia et al., 2012; Gašević, Jovanović, et al., 2017; Trigwell & Prosser, 1991).

- Strategic: It is also referred to as “achievement” learning (Biggs, 1987). Strategic learners apply both surface and deep approaches to learning (Diseth, 2003; Kovanović et al., 2019). Similar to the learners who follow the surface approach to learning, learners who follow the strategic approach to learning focus on achieving high scores with less effort (Emilia et al., 2012). However, they understand that a certain level of effort is required to achieve the learning goal. Hence, strategic learners often show a high positive association with high academic performance.

Although traditional research on approaches to learning relies on self-reports to capture the intended use of learning strategies, the characteristics of each approach can also be observed through the actions of students recorded in trace data. Zhou and Winne (2012) refer to the information observed within the trace data as “realised intention” of students to apply certain types of learning strategies; meanwhile, the information collected with self-report instrument reflects “perceived intention”. When studying the correlation of perceived intention and realised intention and the goal orientation, Zhou and Winne (2012) found that realised intentions observed through trace data were more predictive of learning achievement and reflective of the actual learning behaviours than self-reports. Therefore, trace data offers potentially more accurate accounts about how learners study.

3.1.2 Chapter overview

Relying on the detailed learning activities recorded in the trace data, an approach to extract insightful patterns of the learning actions is required in order to determine what types of learning tactics and strategies students used. With this regard, this chapter aims to address the research question two (RQ2) formulated in Section 1.1, to develop a novel approach to detecting learning tactics and strategies. Precisely, three learning analytics-based approaches are examined including those based on sequence, process and network analytics.

- Sequence – sequence analytics-based approach was proposed by Jovanovic et al. (2017). The sequence approach emphasises discovering the patterns of the chronological ordering of the observed items (i.e., learning actions) (Hassani et al., 2019). This chronological ordering used in the sequence approach is well-aligned with the notion of learning tactic which is considered a sequence of actions performed to complete a task (Hadwin et al., 2007). In the sequence analytics-based approach, the agglomerative hierarchical clustering (AHC) based on Ward's method technique was proposed (Jovanovic et al., 2017). AHC is suitable to detect the pattern based on the (dis)similarity of the sequences. The (dis)similarity of the sequence is commonly computed by calculating the minimal cost of insertions, deletions, and substitutions of the sequences (Gabadinho et al., 2011). This approach has successfully used in the detection of learning strategies, as demonstrated by several research studies (Boroujeni & Dillenbourg, 2018; Jovanovic et al., 2017).

- Process – the goal of process analytics-based approach is to uncover a process model that manifest the 'holistic end-to-end view' of a learning process (Hassani et al., 2019; van der Aalst, 2011). The process analytics-based approach is an approach suitable for the exploration of the dynamic nature of learning tactics and strategies (Bannert et al., 2014; Sonnenberg & Bannert, 2015). The dynamic behaviour of learning tactics and strategies can be observed through the transitions of learning actions illustrated in the process model. The transition of actions is computed by using the first-order Markov model (FOMM). According to Abbeel and Ng (2005), the first order Markov model has successfully been implemented to model the sequential data. The FOMM estimates the standard maximum likelihood based on the first-step of transition from data, resulted in the probability of transition (Abbeel & Ng, 2005). This is considered as the conditional probability; thus, it can capture the condition and dynamic of actions which are the important characteristics of learning tactics as commonly theorized in the Winne and Hadwin model of SRL (Winne, 2013; Winne & Hadwin, 1998) and proposed for analysis of learning tactics and strategies (Winne et al., 1994). While other process mining techniques (alpha algorithm, Heuristic Miner and Fuzzy Miner) are also investigated in the literature on SRL (e.g., Bannert et al. (2014)), those approaches were application to other theoretical models of SRL from the Winne and Hadwin model used in this thesis. The Expectation-Maximization (EM) algorithm was used in tandem with process mining as it works well on the estimation the maximization of log-likelihood of transitions within FOMMs (Ferreira et al., 2007). Thus, the EM algorithm is used to detect patterns of actions based on process models.

- Network – by considering the connection of learning actions, we also propose a network analytics-based approach. The approach is based on Epistemic Network Analysis (ENA). ENA builds on the notion that the structure of elements is more important than the absence or presence of the elements (Shaffer et al., 2016). As such ENA emphasises the structure of learning actions exhibited by students as they engage in learning activities. ENA generates a frequency of co-occurrence of learning actions. The ENA approach is also well aligned with the previous proposals for the use of graph theory for modelling of learning tactics and strategies from the viewpoint of the Winne and Hadwin model of SRL (Winne, 2013). The clustering technique used in network analytics-based approach is AHC. The distance is computed by using Euclidean distance measure based on Ward's method. The application of such clustering technique has been effectively applied in the detection of learning strategies based on the frequency of variables observed in the trace data, e.g. Kovanović et al. (2015), Kovanović et al. (2019), Wise et al. (2013).

The three learning analytics-based approaches are employed to analyse the log data. The sequences of learning actions with corresponding timestamps are used as the input to the computation algorithms. The main difference among the three approaches is the output. The sequence analytics-based approach produces a collection of frequent sequences (Hassani et al., 2019). The process analytics-based approach computes the matrix that represents the probability of the transitions among the learning actions (Gatta et al., 2017). Whereas, the network analytics-based approach computes the co-occurrence of the learning actions, hence, resulted in the collection of the weights (i.e frequency) of co-occurred actions (Shaffer et al., 2016).

The data used in this chapter was collected from the Introduction to the Python course. Table 2 presents the description of learning actions recorded in the trace data. This study aims to explore how the proposed analytics-based approaches enabled the detection of the learning tactics and strategies used by the students. Thus, the data used in this study mainly focused on phase 3 of the Winne and Hadwin model of SRL, where the operations to enact the learning tactics and strategies were carried out. Each learning action was recorded with the corresponding timestamp.

The main contribution of this study is the examination of the (dis)similarity of learning tactics and strategies detected using different analytics-based approaches. To ensure that the detected learning strategies are theoretically meaningful, as inquired in the research question three (RQ3) in Section 1.1, approaches to learning are used to interpret the findings.

Table 2. Description of the learning actions recorded in the trace data and the corresponding SRL phases

SRL Phases	Learning actions observed in the trace data	Description
1: Goal setting and planning		
2: Tasks Identification		
3: Enactment of learning tactics and strategies	lecture_start	Start the video lecture
	lecture_complete	Complete the video lecture
	in_video_quiz	Answer a quiz embedded in the video
	In_video_quiz_correct	Correctly answer a quiz embedded in the video lecture
	In_video_quiz_incorrect	Incorrectly answer a quiz embedded in the video lecture
	Supplement_complete	View the supplementary documents
	Quiz_start	Start a theoretical exercise
	Quiz_complete	Complete a theoretical exercise
	Quiz	Theoretical exercise progress
	Exam_start	Start a practical exercise
	Exam_complete	Complete a practical exercise
	Exam	Practical exercise progress
	Exam_correct	Correctly solved a practical exercise
	Exam_incorrect	Incorrectly solved a practical exercise
	Code_execute	Command to execute the code
	Discussion_question	Post a question to the discussion board
	Discussion_answer	Post an answer to a question in a discussion board
	Discussion_question_vote	Vote for a question
	Discussion_answer_vote	Vote for an answer to a question
	Discussion_answer_del_vote	Deleted a vote for an answer
	Discussion_follow	Flag to follow a discussion
	Discussion_unfollow	Flag to unfollow a discussion
4: Adaption		

3.2 Publication: Detection of Learning Strategies: A Comparison of Process, Sequence and Network Analytic Approaches

The following section includes the verbatim copy of the following publication:

Matcha, W., Gašević, D., Ahmad Uzir, N., Jovanović, J., Pardo, A., Maldonado-Mahauad, J., & Pérez-Sanagustín, M. (2019). Detection of Learning Strategies: A Comparison of Process, Sequence and Network Analytic Approaches, *European conference on technology enhanced learning*, Springer. https://link.springer.com/chapter/10.1007/978-3-030-29736-7_39

Detection of Learning Strategies: A Comparison of Process, Sequence and Network Analytic Approaches

Wannisa Matcha[1(✉)], Dragan Gašević[1,2(✉)], Nora'ayu Ahmad Uzir[1,3(✉)], Jelena Jovanović[4(✉)], Abelardo Pardo[5(✉)], Jorge Maldonado-Mahauad[6,7(✉)], and Mar Pérez-Sanagustín[6,8(✉)]

[1] University of Edinburgh, Edinburgh EH8 9AB, UK
{w.matcha,n.uzir}@ed.ac.uk
[2] Monash University, Clayton, VIC 3800, Australia
dragan.gasevic@monash.edu
[3] Universiti Teknologi MARA, 40150 Shah Alam, Malaysia
[4] University of Belgrade, 11000 Belgrade, Serbia
jelena.jovanovic@fon.bg.ac.rs
[5] University of South Australia, Adelaide, SA 5000, Australia
abelardo.pardo@unisa.edu.au
[6] Pontificia Universidad Católica de Chile, Santiago, Chile
jjmaldonado@uc.cl, mar.perez@ing.puc.cl
[7] Universidad de Cuenca, Cuenca, Ecuador
[8] Institute de Recherche en Informatique de Toulouse (IRIT), Université Paul Sabatier Toulouse III, 31062 Toulouse Cedex 9, France

Abstract. Research in learning analytics proposed different computational techniques to detect learning tactics and strategies adopted by learners in digital environments through the analysis of students' trace data. While many promising insights have been produced, there has been much less understanding about how and to what extent different data analytic approaches influence results. This paper presents a comparison of three analytic approaches including process, sequence, and network approaches for detection of learning tactics and strategies. The analysis was performed on a dataset collected in a massive open online course on software programming. All three approaches produced four tactics and three strategy groups. The tactics detected by using the sequence analysis approach differed from those identified by the other two methods. The process and network analytic approaches had more than 66% of similarity in the detected tactics. Learning strategies detected by the three approaches proved to be highly similar.

Keywords: Learning strategy · Learning analytics · Data analytics

1 Introduction

The objective of massive open online courses (MOOCs) is to offer learning opportunities to a wide range of learners. However, MOOCs have been associated with high dropout and failure rates [1, 2]. Research identified several factors associated with such course outcomes including motivation, intention, time management, and learning experiences, to name a few [3, 4]. Learning tactics and strategies adopted by MOOC participants have been identified as key factors of success prediction [5–7]. Much research in traditional learning environments explored students' learning strategies [6, 8]. However, students' learning strategies in MOOCs are much less understood. MOOC platforms allow for recording trace data of the actual learners' behavior. However, such data are large, diverse, and complex to analyze. As a consequence, researchers have proposed a variety of methods that go beyond traditional statistics methods to unveil students' learning strategies [9, 10]. While the applied data analytic methods led to useful findings, the diversity of the adopted methods hindered the replication and generalization of the results. Little work has been done to compare how the applied approaches differ in terms of the tactics and strategies that they identify. This study explored how three analytic approaches – drawing from sequence, process, and network analytic techniques – could influence the detection of learning tactics and strategies.

2 Background

Research has emphasised the importance of using effective learning strategies as one of the key factors of successful learning. Learning strategy can be defined as *"any thoughts, behaviors, beliefs or emotions that facilitate the acquisition, understanding, or later transfer of new knowledge and skills"* [12, p. 727]. A closely related construct is the one of learning tactic, which can be defined as a sequence of actions that a student performs in relation to a given task within a learning session [12]. Defined in terms of tactics, learning strategies can be considered the regularity in the application of learning tactics or a pattern of how each student uses certain tactics [13]. Such patterns of tactic application evolve and become the characteristics of one's learning, which may be considered as aptitudes that could further predict the future behaviors [14].

Thanks to the large dataset of trace data on students' behavior, contemporary research aims to leverage these datasets to explore learning tactics and strategies by considering how these dynamic constructs unfold. In *network analytic approaches*, learning tactics and strategies are identified from networks built based on the co-occurrence of learning states or actions. These approaches were originally proposed for studying learning strategies as learning sequences [15]. The application of graph multiplicity measures, as commonly used in network science, has been then suggested to analyze the importance of individual events that contribute to student learning. For example, Siadaty et al. [16] applied this methodology to identify how technological interventions activated different processes of self-regulated learning. More recently, approaches suggest the use of *sequence analysis* techniques combined with unsupervised learning to detect learning tactics and strategies from trace data [9]. Similarly, learning tactics and strategies can be identified by analyzing the distribution of learning sequences [17].

Process-oriented data analysis approach emphasise the timing of the events. Malmberg et al. explored self-regulated learning strategies in a collaborative learning context by using a process mining technique [18]. Similarly, Matcha et al. [10] detected learning tactics and strategies from trace data by combining temporal analysis of the trace data (first-order Markov models) and clustering (Expectation-Maximization) [10]. Maldonado-Mahauad et al. [19] used a combination of process mining and clustering techniques to identify self-regulated learning strategies that different group of learners employed when interacting with the course contents (video-lectures and assessments).

Despite the interesting insights produced by these individual approaches, there has been limited research that explored how these three analytic approaches might have influenced the results. Hence, this paper aims to answer the following research question: *How do different data analytics techniques proposed in the literature for the detection of learning tactics and strategies apply to the same dataset?* That is, the paper compares approaches that emphasize sequence, network and process dimensions.

3 Methods

3.1 Data

The data used in this study was collected from the Introduction to Python course offered by the Pontificia Universidad católica de Chile on the Coursera MOOC platform in its two different editions. A total of 4,217 students registered their interest in the course. The course was in Spanish and was offered on demand (i.e. self-pace). In 8 weeks, the course covered six programming topics with 2–3 subtopics each. For each topic, the course offered a set of short video lectures with embedded questions (to provoke a simple recall of the concepts) and a set of reading materials. The students also had several theoretical exercises (11 quizzes) and practical exercises (13 exams). Among the quizzes and exams, 22 items were graded and accumulated to calculate students final mark. At least 80% of these items had to be answered correctly to pass the course. The students were also offered the discussion board to discuss course topics. In this study, we considered only the trace data of those students who completed at least one assignment during the official course schedule between September 17th and November 4th 2018. As a result, 368 students were considered for the study. We coded the different learning actions captured in the trace data as described in Table 1.

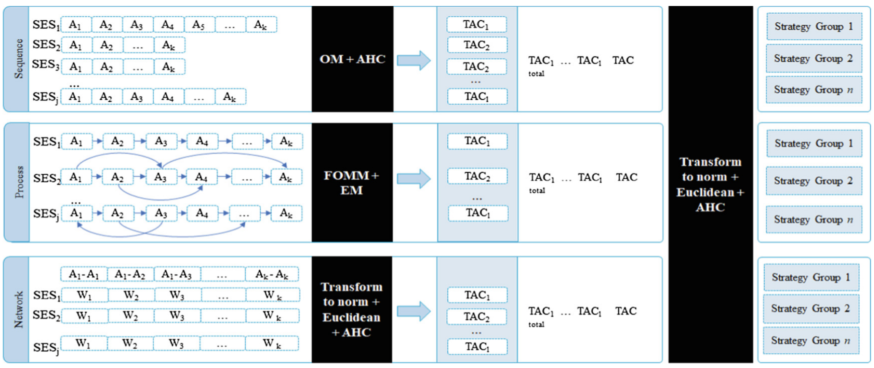
The resulting dataset for the analysis study contained the following items for each learning actions: the anonymous user ID, timestamp, type of learning action, and reference to course items. Each two consecutive learning sessions were separated by at least 30 min of inactive time [20]. Due to the requirements of analytic methods to be applied, the outliers were excluded: extremely short sessions (one action in a session) and extremely long sessions (>95th percentile of actions per session).

3.2 Methods

Figure 1 illustrates the pipeline of the analytic methods used to extract learning tactics and strategies from the trace data following the three analytic approaches discussed in Sect. 2. The data were pre-processed based on the requirement of each analytic approach.

Table 1. Coding of learning actions from the data trace

Events	Coded events	Description
Video lecture	lecture_start	Start the video lecture
	lecture_complete	Complete the video lecture
	in_video_quiz	Answer a quiz embedded in the video
	In_video_quiz_correct	Correctly answer a quiz embedded in the video lecture
	In_video_quiz_incorrect	Incorrectly answer a quiz embedded in the video lecture
Reading	Supplement_complete	View the supplementary documents
Theoretical exercises	Quiz_start	Start a theoretical exercise
	Quiz_complete	Complete a theoretical exercise
	Quiz	Theoretical exercise progress
Practical exercise	Exam_start	Start a practical exercise
	Exam_complete	Complete a practical exercise
	Exam	Practical exercise progress
	Exam_correct	Correctly solved a practical exercise
	Exam_inccorect	Incorrectly solved a practical exercise
	Code_execute	Command to execute the code
Discussion	Discussion_question	Post a question to the discussion board
	Discussion_answer	Post an answer to a question in a discussion board
	Discussion_question_vote	Vote for a question
	Discussion_answer_vote	Vote for an answer to a question
	Discussion_answer_del_vote	Deleted a vote for an answer
	Discussion_follow	Flag to follow a discussion
	Discussion_unfollow	Flag to unfollow a discussion



(*SES: Learning Session; A : Learning Action; W : Weight of co-occurrence between two actions; FOMM: First Order Markov Model; EM: Expectation-Maximization; OM: Optimal Matching Score; AHC: Agglomerative Hierarchical Clustering; TAC: Learning Tactic)

Fig. 1. The pipeline of the analytic methods used in the study

Sequential Dimension. Following the work in [9], the TraMineR R package [21] was used to explore the sequential data. Learning actions were arranged chronologically and split into learning sessions. Sessions were encoded as learning sequences using a TraMineR's sequence representation format [21]. The optimal matching technique, with substitution costs based on transition rates, was used to compute the (dis)similarity of the sequences. Agglomerative hierarchical clustering based on Ward's algorithm was used to group learning sequences based on shared patterns of learning actions.

Process Dimension. The process dimension was explored by replicating the steps proposed in [10]. The pMineR R package was used to generate a process model of learning and compute the probability of state transitions by using the first-order Markov model (FOMM) [22]. The process model was formulated using timestamped learning events in each learning session. The Expectation-Maximization (EM) algorithm was used for clustering of learning sequences as it works well with the FOMM.

Network Dimension. The rENA R package for Epistemic Network Analysis (ENA) was used to compute the co-occurrence of learning actions in each learning session [23]. By generating a network using ENA, a matrix of co-occurrences of learning actions was created. The co-occurrence values in the matrix were normalized and subsequently used as an input to the agglomerative hierarchical clustering, based on Ward's algorithm. The Euclidean method was used to calculate the (dis)similarity.

The clusters of sequences (i.e., tactics) detected by each of the three data analytic approaches were then explored in terms of sequence length and event distributions. The similarities between the three approaches were also calculated as proportions of learning sessions shared across the tactics detected by the three approaches.

To compute learning strategies, we used the results of cluster assignments of each of the three above approaches. Specifically, for each student, we computed the counts of each of the detected tactics and the total count of tactics. These counts were then normalized (i.e., reduced to the range of 0 to 1) and used as input to the agglomerative hierarchical clustering method. The computation of the (dis)similarity of students' tactic use was based on the Euclidean metric. The identified clusters were considered manifestations of the students' learning strategies (i.e., patterns of learning tactics). This was done for each of the three examined approaches. The identified learning strategies were explored based on how students applied the tactics according to the course topics. Furthermore, the association of the identified strategies and the final course marks was examined using Kruskal Wallis tests followed by pairwise Mann Whitney U tests.

4 Results

4.1 Learning Tactics

The results revealed that the three detection approaches identified four similar learning tactics. Figure 2 presents the counts of learning actions in each tactic as identified with the three analytics approaches. Further details of the tactic characteristics are provided in the supplementary document (Tables 1, 2 and 3)[1].

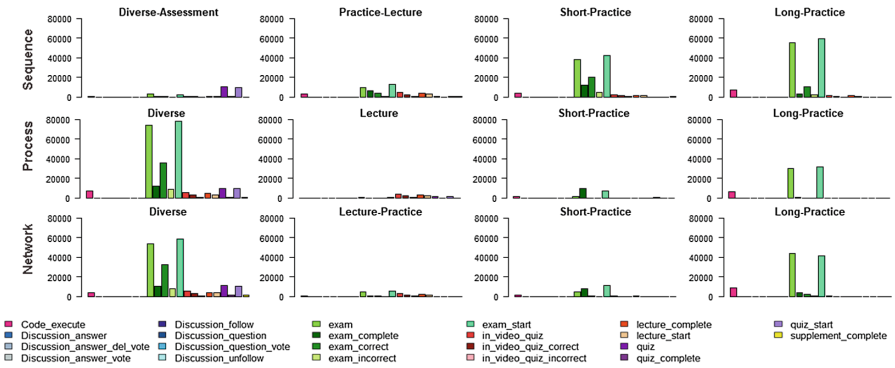


Fig. 2. The distribution of learning action counts across the tactics detected by the three analytic approaches

Sequence Approach. The dendrogram suggested four clusters as the best result. The *Practice and Lecture-oriented* cluster (N = 3134 sessions, 59.34%) was the largest and contained the shortest sequences (Mdn = 10 actions). The most dominant actions included those related to the exam activities, interaction with the video lecture, and quizzes embedded in the video. The *Diverse Assessment-oriented* (N = 208 sessions, 3.94%) cluster was very small and contained long sessions ranging from 54 to 355 actions. This tactic often began by interacting with the video lectures, followed by doing the exam and ended by interacting with the quiz items. The *Short Practice-oriented* (N = 1292 session, 24.47%) cluster included practical exams and code as the most dominant actions. Access to the video lectures was also prominent. The length of the sequences was moderate as compared to the other three tactics (Mdn = 93 actions). The *Long Practice-oriented* cluster (N = 647 sessions, 12.25%) was relatively small exhibiting a pattern similar to the previous one (Short Practice-oriented). However, learning sequences were longer, ranging from 103 to 359 events (Mdn = 214).

Process Approach. Four tactics were identified with the process analytic approach as optimal. The *Diverse* cluster (N = 2000 sessions, 37.87%) varied in the number of actions in each session in the [3–359] range (Mdn = 105). The main learning actions were related to exam activities, followed by quizzes, code execution, and interaction with lecture videos. The *Lecture-oriented* cluster (N = 1391 sessions, 26.34%) contained short sessions (Mdn = 7 actions). The most dominant actions included interaction with the video lectures and the quizzes embedded in the videos, followed by interaction with the quizzes that were part of the theoretical questionnaires. The *Short Practice-oriented* cluster (N = 772 sessions, 14.62%) consisted mostly of short

[1] Supplementary Document can be found at: https://bit.ly/2E4pFCu

sessions (Mdn = 8 actions) that were generally of two types: (i) short sessions of code executions and (ii) longer sessions of completing an exam. The *Long Practice-oriented* cluster (N = 1118 sessions, 21.17%) mostly included actions related to the exam or code execution. Unlike the *Short-practice-oriented* tactic, completed exams were rarely observed in this tactic.

Network Approach. The dendrogram inspection suggested four clusters as optimal. *Diverse-oriented* (N = 1892 sessions, 35.83%) was similar to the Diverse tactic detected by the other two approaches; this tactic included a variety of actions, dominated by those related to exam and quiz related activities. However, the number of actions within a session was much higher compared to the Diverse tactic detected by the other two methods (Mdn = 93 actions). *Lecture and Practice-oriented* (N = 929 sessions, 17.59%) was the most dominant with exam-related actions and a small proportion of actions related to the lecture videos. However, when inspecting all the sequences, this cluster contained multiple short sessions of video lecture related actions often followed by long sequences of exam related actions. Unlike the Lecture and Exam-oriented tactic detected by the process analytic approach, the frequency of interactions with exam items outnumbered lecture-related actions, while quizzes-related actions were almost invisible. *Short Practice-oriented* (N = 1776 sessions, 33.63%) was similar to the Short Practice-oriented tactic detected with the process approach. This tactic consisted of short learning sessions (Mdn = 7 actions). It was dominated by two types of sequences: (i) short session of code executes, and (ii) longer sessions of initiating and completing an exam. *Long Practice-oriented* (N = 684 sessions, 12.95%) contained longer sequences of action (Mdn = 126 actions). The most dominant learning actions were related to the exam or code execution. The proportion of initiated but not necessarily completed exams and continuing doing the exam was relatively high.

4.2 Comparison of Detected Tactics

The *Diverse* tactic detected by the process and network approaches showed similar patterns; that is, it was composed of several different learning actions and diverse length of sequences. The most frequent action was interaction with the exam, followed by the interaction with quizzes. *Diverse-assessment-oriented*, as detected by the sequence approach, showed that the interactions with the quizzes were more frequent than the exams. *Lecture and practice*-oriented included events about actions related to video lectures and exams as the most dominant. Opposite to the other two approaches, the lecture related events outnumbered the exam focused events in the case of the process approach. *Short Practice-oriented* was defined by intense interaction with the exam items and code implementation. The median length of sequences of this tactic was smaller than of that of the *Long Practice-oriented* tactic. This tactic, as identified by the sequence approach, had the highest mean length of sequences and higher frequency of video lecture interactions than the same tactic detected by the other two approaches.

The sequence approach proved to be the best in distinguishing *Long Practice-oriented* as the one characterized by long sessions of exam interaction and code execution. The process and network approaches showed inconsistency in categorising based on the length of the sequences.

Table 2. The similarity of tactics detection based on three analytic approaches

Similarity: 1861 Sessions (35.24%)		**Process Analytic Approach (100%)**			
		Diverse - Practice	Lecture	Long-Practice	Short-Practice
Sequence Analytic Approach	Diverse-Assessment	9.25	1.65	0	0
	Lecture and Exam	22.4	98.35	58.94	85.36
	Long-Practice	21.6	0	18.34	1.3
	Short-Practice	46.75	0	22.72	13.34
Similarity: 1500 Sessions (28.40 %)		**Network Analytic Approach (100%)**			
		Diverse – Practice	Lecture and Practice	Long-Practice	Short-Practice
Sequence Analytic Approach	Diverse-Assessment	10.84	0.32	0	0
	Lecture and Exam	29.49	89.56	14.62	92.57
	Long-Practice	14.64	2.26	49.71	0.51
	Short-Practice	45.03	7.86	35.67	6.93
Similarity: 3526 Sessions (66.77%)		**Network Analytic Approach (100%)**			
		Diverse – Practice	Lecture and Practice	Long-Practice	Short-Practice
Process Analytic Approach	Diverse	85.68	17.33	25.44	2.48
	Lecture	10.94	78.26	0	25.73
	Long-Practice	2.11	3.34	69.44	32.21
	Short-Practice	1.27	1.08	5.12	39.58

Table 2 compares the results of the three analytic approaches based on cluster assignments of study sessions. The similarity was computed by calculating the proportion of learning sequences that were categorized as the same tactic. The sequence approach had 35% of overlap in session assignment with that of the process analytic approach, and 28% with that of the network approach. Almost 67% of sessions were categorized as representing the same tactics by the process and network analytic approaches. The *Lecture-oriented* tactic showed a high consistency among the three methods. About 98% of sessions labelled as the lecture-oriented tactic detected with the process analytic approach were also categorised as the same tactic in the sequence analytic approach. This high consistency might be a result of the high number of short learning sessions that coincide with interaction with lecture videos. The highest inconsistency among the approaches was for the *Short Practice-oriented* tactic.

The process and network analytic approaches categorised 3,526 (out of 5,281) sessions as the same tactics. We further explored the sequences that were grouped differently to examine how the approaches differ in grouping the sequences. One of the examples is SequenceID13745 that consisted of actions shifting between practical exam_start and exam_progress. Execution of code was also observed during the exam progress, as shown in Fig. 3. This session consisted of 29 actions, which were inferred as representative of the *Long Practice-oriented* tactic by the process analytic approach. However, in case of the network analytic approach, the *Long-practice oriented* tactic had a higher median session length (Mdn = 126), so that the considered sequence (SequenceID13745) was not qualified as an instance of the *Long Practice-oriented* tactic, but rather fitted in the *Short-Practice-oriented* tactic.

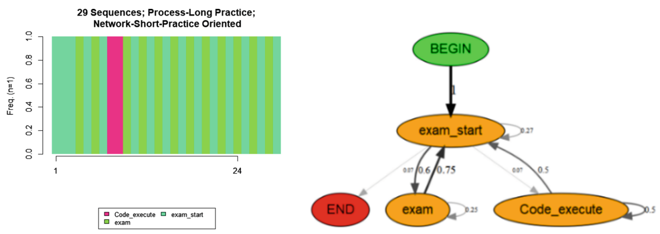


Fig. 3. The visualisation of sequenceID13745 and its first order Markov model

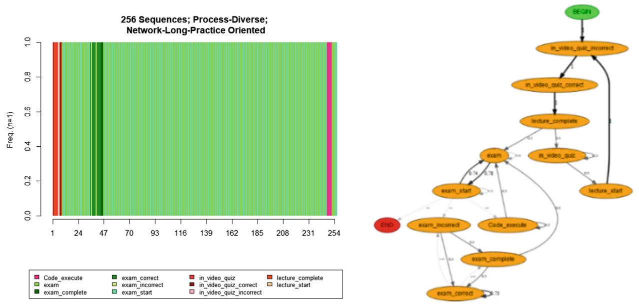


Fig. 4. The visualisation of sequenceID21601 and its process model

Another example of differences in the tactic detection is SequenceID21601 which contained 256 events. The sequence began by interacting with a quiz in a lecture video, followed by transitions between exam_start, exam progress, a correct/incorrect exam answer, and exam complete; the command to execute the code was observed towards the end of the session, as presented in Fig. 4. The sequence and network analytic approaches associated this session with the *Long Practice-oriented* tactic. This is reasonable, since this sequence was relatively long, and the events showed dynamic transitions between the exam related actions. Meanwhile, the process analytic considered this sequence as an instance of the *Diverse* tactic. This is presumably because the sequence began by interacting with the video lecture. The Diverse tactic exhibited events about a variety of learning activities in a session.

4.3 Learning Strategy Groups

Learning strategies were identified as patterns of how students regulated the tactics according to the study topic. Detail characteristics of each strategy group are provided in the supplementary document (see footnote 1).

Sequence Approach. Three strategy groups were extracted based on how the students employed the tactics identified with the Sequence approach. Figure 5 presents the mean number of tactics employed according to the studied topic. ***Strategy Group 1*** (N = 151 students, 41.03%) exhibited a low level of engagement. The dominant tactic was *Lecture-oriented* with short sessions. The mean number of sessions declined as the course topic progressed for all tactics except for the *Short Practice-oriented* tactic. The students who employed this strategy pattern had a high rate of failing the course (77.48%); their median course grade was 36.14 over 100, and the median number of passed graded items was 9 (out of 22). ***Strategy Group 2*** (N = 151 students, 41.03%) exhibited a high level of engagement when interacting with the first two topics by utilising the *Lecture-oriented* tactic. The *Short* and *Long practice-oriented* tactics increased when the course reached the second topic. However, the level of engagement dropped remarkably after completing the third topic. This strategy group had the highest failure rate (88.74%). The median of the completed graded items was four, and the median course grade was 18.04. ***Strategy Group 3*** (N = 66 students, 17.94%) had the highest course grade (Mdn = 82.86/100), highest number of passed graded items (Mdn = 20 items), and the smallest failure rate (54.55%). Similar to the other strategy groups, the students frequently used the *Lecture-oriented* and *Short practice-oriented* tactics. Unlike the first two strategy groups, the mean number of sessions increased as the students moved to more difficult topics.

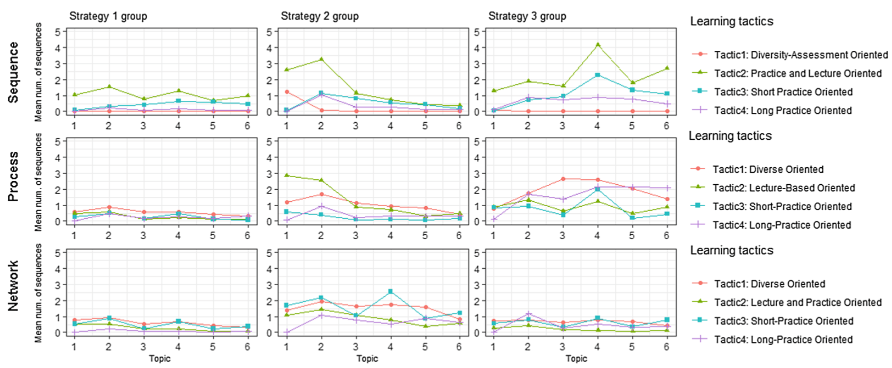


Fig. 5. Frequency of tactics used for each topic and for each strategy group as detected by the three analytic approaches

Process Approach. The mean number of employed tactics detected based on the process analytics approach according to the studied topic is presented in Fig. 5. ***Strategy Group 1*** (N = 215 students, 58.42%) exhibited a low engagement level. The mean number of sessions was consistently below one per study topic. The students who adopted this strategy had a high failing rate (82.79%); their median course grade was 29.33 over 100, and the number of passed graded item was 7 out of 22 items. ***Strategy Group 2*** (N = 89 students, 24.18%) included the students who were quite selective. The *Lecture-oriented* and *Diverse* tactics were dominant at the beginning of the course. The level of engagement dropped constantly from Topic 3 onwards. Despite putting a higher level of effort than *strategy group 1*, the students in this group passed less graded items (Mdn = 5), and had lower course grade (Mdn = 20.41). *Strategy Group 3* (N = 64 students; 17.39%) showed the highest passing rate (43.75%) and grades (Mdn = 82.71/100). Unlike the other strategy groups, the students in this group were consistently increasing their engagement with the course topics. As the MOOC progressed and the topics became more challenging, this group put more effort and used diverse learning tactics, as shown by the high use of the *Practice-oriented* and *Diverse* tactics.

Network Approach. Figure 5 shows three strategy groups with similar tactic enactment patterns. ***Strategy Group 1*** (N = 188 students, 51.09%) used multiple tactics but with a low frequency, and the frequency decreased as the course progressed. The group had a high failure rate (83.51%) with the median score of 27.29 (over 100), and passed, on average, 7 (out of 22) graded items. ***Strategy Group 2*** (N = 94 students, 25.54%) included the students who were the most active. They employed a variety of tactics to study each topic. The use of the *Diverse* and *Lecture-oriented* tactics slightly declined as the course progressed. There was some fluctuation in the use of the *Short Practice-oriented* tactic, especially during the fourth topic. The students with this strategy had the highest course score (Mdn = 56.95), and passed more graded items (Mdn = 15 items) than those following the other two strategies. ***Strategy Group 3*** (N = 86 students, 23.37%) had a similar pattern as the first one. Yet, the rate of students who failed was lower (74.42%), and the median grade was higher (Mdn = 37.04) than for strategy 1.

Association with Performance. The strategy groups detected by using the sequence approach showed no significant association with the course grade, nor with the number of item passed (Table 3). However, we detected a significant association of the strategy and the potential of failing/passing the course. The pairwise comparison (Table 3) showed statistically significant associations among all the strategy groups and the potential of failing/passing the course. The effect sizes ranged from small to medium.

Table 3. Kruskal-Wallis (above) and pairwise comparison (below) of strategy groups with respect to performance

	Sequence	Process	Network
Course Grade	p = 0.125	p = 0.067	p = 0.14
Passed Graded Items	p = 0.082	p = 0.0004*	p = 0.01*
Passed Course	p = 0.046*	p = 0.0004*	p = 0.025*

Approach	Item	Strategy	Strategy	Z	p	r
Sequence	Passed Course	S1	S2	2.607	0.014*	0.150
		S1	S3	-3.401	<0.001*	0.231
		S2	S3	-5.613	<0.001*	0.381
Process	Passed Course	S1	S2	-0.160	0.88	0.009
		S1	S3	-4.401	<0.001*	0.263
		S2	S3	-3.463	<0.001*	0.28
	Passed Graded Items	S1	S2	0.102	0.87	0.006
		S1	S3	-7.206	<0.001*	0.431
		S2	S3	-6.359	<0.001*	0.514
Network	Passed Course	S1	S2	-2.440	0.05	0.146
		S1	S3	-1.765	0.18	0.107
		S2	S3	0.516	0.57	0.039
	Passed Graded Items	S1	S2	-4.323	<0.001*	0.258
		S1	S3	-2.762	0.05	0.167
		S2	S3	1.613	0.059	0.121

Note: * marks statistically significant differences

The strategy groups detected with the process analytic approach had no significant differences in course grades. A significant association was present between the strategy groups and the number of passed graded items and the potential of failing/passing the course. Pairwise comparisons of strategy groups with respect to the completed performance items showed significant differences between strategy group 1 and 3 and groups 2 and 3. The effect sizes were medium except for the passed graded items between strategy groups 2 and 3 where the effect size was large (r = 0.514).

The strategy groups identified with the network analytic approach had no significant difference on course grades. The strategy groups proved to differ significantly with respect to the number of items passed and the potential of failing/passing the course. Pairwise comparisons showed significant differences between strategy groups 1 and 2 on the number of passed graded items with the small effect sizes.

4.4 Comparison of Detected Strategy Groups

Table 4 summarises the detected strategy groups along several dimensions related to the students' pattern of course engagement and academic achievement.

Table 4. Comparison of the strategy groups as detected by the three analytic approaches

	Sequence	Process	Network
Highly active and multiple tactics used	Strategy3	Strategy3	Strategy 2
Highly active at the beginning	Strategy2	Strategy2	–
Surface engagement	Strategy1	Strategy1	Strategy1, Strategy3

Highly Active and Multiple Tactics Used. These strategy groups reflect the deep learning approach as defined by Biggs (1987). The deep approach is characterised by high efforts, a variety of learning tactics used [7, 10], and associated with the high academic performance [4]. The students employed a variety of tactics when interacting with each topic. Based on the sequence approach, the most dominant tactic used was *Lecture-oriented*. Based on the process and network approaches, the dominant tactics were *Diverse* and *Practice-oriented*. Regardless of the tactic detection method, a similar pattern of interaction with the fourth course topic was observed – high enactment of the *Short Practice-oriented* tactic. This suggested that students might have been facing some challenges with the fourth topic that the instructor should consider when designing the next course iteration.

Highly Active at the Beginning. The sequence and process analytic approaches detected this similar pattern of tactic use, but not the network approach. The students were actively engaged during the first two topics, and then the effort significantly declined. The tactics employed during the first three topics showed that students were strategic in choosing tactics. The dominant tactics were *Lecture-oriented* and *Diverse*. This reflects the *Strategic* approach to learning [24], characterized by the aim of achieving high performance with the strategic choice of tactics [8, 24]. As the students faced more difficulty, their learning strategy shifted from strategic to the surface approach to learning. This suggested that some interventions are needed to maintain the level of students' engagement with the third topic. This group showed high engagement as compared to the Surface group, but the group missed to complete a few graded items.

Surface Engagement. This group represented the surface approach to learning. As defined by Biggs (1987), students who follow this approach to learning employ surface effort and have low academic performance [8, 24]. In our study, the students who followed this strategy group exhibited a low level of engagement and high failure rate.

None of the analytic approaches identified strategy groups that were predictive of performance. A significant association was found between the strategy group and the passed graded items for all cases. The process analytic approach proved the best in detecting strategy groups predictive of the passed graded items.

5 Conclusions

Summary. The findings in this study showed that sequence, process, and network analytic approaches can be used to detect meaningful learning tactics from MOOC trace data. The three approaches resulted in tactics that were similar to some extent (Table 2). The highest similarity (67% of detected tactics) was found between the process and network approaches. As for strategy detection, the results of the network analytic approach differed from the other two approaches. The sequence and process analytic approaches resulted in similar strategy groups.

In general, we observed that sequences with similar learning actions were grouped in the same cluster. The length of the sequences affected the clustering in the sequence analytic approach. For example, short learning sessions were grouped into a single cluster (i.e. short diverse oriented) and this was the key distinguishing characteristic of this tactic group. In contrast, the process and network analytics were less based on the length of the sequences. Therefore, in the tactics detected using these two approaches the number of actions per learning sessions varied, ranging from two to hundred or more.

The proportion of learning sessions that belonged to each of the detected tactics impacted the learning strategy detection. The sequence approach detected one large tactic, i.e. *Short Practice-Lecture oriented*, showed that all strategy groups were dominated by this tactic. Furthermore, we found that all of the strategy groups exhibited a high frequency of using the *Short Practice and Lecture-oriented* tactics. This is unsurprising considering the course design that emphasized the use of video lectures and practice exercises.

Implications. The key finding of the study is that the choice of the data analytic approach for detection of learning tactics and strategies affects the results. Specifically, the three approaches emphasize different dimensions of learning tactics – sequential, process, and network. The differences in the underlying modelling of the three analytic approaches produced different data representations that are then fed to an unsupervised (i.e., clustering) machine learning algorithm. The properties of these underlying representations – sequence, process, and network – had direct implications on the computation of the similarities between individual sessions, and thus, the way how clusters were formed to detect learning tactics. Moreover, the choice of the underlying modelling approaches for tactics had a direct impact on the choice of clustering algorithm. For example, the process approach produced the data structure (i.e., adjacency matrix) that was not suitable for analysis with AHC; EM was used instead as also used in the literature [10]. AHC was more suited for the other two approaches, as commonly applied in the literature on similar tasks [9].

Based on the results of our findings, we cannot indicate which of the approaches is 'best'. Instead, the (dis)similarities in the results the three approaches produced and interpretations of the (dis)similarities in this study can inform decisions of researchers and practitioners who work on the detection of learning tactics and strategies. Given that each of the three approaches used unsupervised machine learning at its core, it is also important that the interpretation of results should be done by considering a well-grounded educational learning theory and the learning context the data originate from. In our case, we offered examples that grounded in the theory of approaches to learning and the design of the MOOC used in the study. The use of these two sources demonstrated that all three approaches produced practically and theoretically meaningful learning tactics and strategies.

The differences in the learning strategies detected by the three approaches can directly be attributed to the differences in the modelling approaches used for the detection of learning tactics. This is due to the use of the identical methodology applied in the second step of the three detection approaches (see Fig. 1). Future research should investigate the extent to which changes in the modeling approaches in the second step will influence the results in the detection of learning strategies.

Limitations. Some limitations of this research must be highlighted. First, the detection of learning tactics and strategies relied primarily on trace data. Although limitations of self-reports are well document [12, 25], self-reports could add to the understanding of students' conditions, intention and motivation. Moreover, using multimodal techniques to capture the data could offer a fine-grained dataset. Second, some degree of subjectivity was evident in the selection of the number of clusters identified, even though the selection was informed by the information generated with the clustering technique (e.g., dendrogram in agglomerative hierarchical clustering) and further informed by the interpretability of the cluster solutions. Future research should explore approaches that can be used to produce a 'stable' number of clusters across different contexts.

Acknowledgements. This work was funded by the LALA project (grant no. 586120-EPP-1-2017-1-ES-EPPKA2-CBHE-JP). This project has been funded with support from the European Commission. This publication reflects the views only of the author, and the Commission and the Agency cannot be held responsible for any use which may be made of the information contained therein.

References

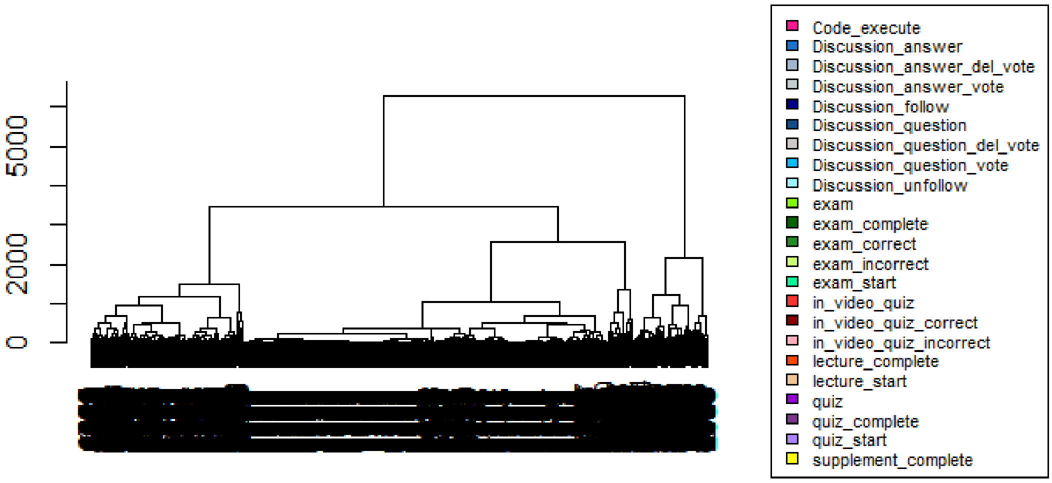
1. Zurita, G., Hasbun, B., Baloian, N., Jerez, O.: A blended learning environment for enhancing meaningful learning using 21st century skills. In: Chen, G., Kumar, V., Kinshuk, Huang, R., Kong, S.C. (eds.) Emerging Issues in Smart Learning. LNET, pp. 1–8. Springer, Heidelberg (2015). https://doi.org/10.1007/978-3-662-44188-6_1
2. Drachsler, H., Kalz, M.: The MOOC and learning analytics innovation cycle (MOLAC): a reflective summary of ongoing research and its challenges. J. Comput. Assist. Learn. **32**, 281–290 (2016)
3. Kizilcec, R.F., Pérez-Sanagustín, M., Maldonado, J.J.: Self-regulated learning strategies predict learner behavior and goal attainment in massive open online courses. Comput. Educ. **104**, 18–33 (2017)
4. Broadbent, J., Poon, W.L.: Self-regulated learning strategies & academic achievement in online higher education learning environments: a systematic review. Internet High. Educ. **27**, 1–13 (2015)
5. Winne, P.H.: How software technologies can improve research on learning and bolster school reform. Educ. Psychol. **41**, 5–17 (2006). https://doi.org/10.1207/s15326985ep4101
6. Yip, M.C.W.: Differences in learning and study strategies between high and low achieving university students: a Hong Kong study. Educ. Psychol. **27**, 597–606 (2007)
7. Maldonado-Mahauad, J., Pérez-Sanagustín, M., Moreno-Marcos, P.M., Alario-Hoyos, C., Muñoz-Merino, P.J., Delgado-Kloos, C.: Predicting learners' success in a self-paced MOOC through sequence patterns of self-regulated learning. In: Pammer-Schindler, V., Pérez-Sanagustín, M., Drachsler, H., Elferink, R., Scheffel, M. (eds.) EC-TEL 2018. LNCS, vol. 11082, pp. 355–369. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-98572-5_27
8. Chonkar, S.P., et al.: The predominant learning approaches of medical students. BMC Med. Educ. **18**, 1–8 (2018)
9. Jovanovic, J., Gasevic, D., Dawson, S., Pardo, A., Mirriahi, N.: Learning analytics to unveil learning strategies in a flipped classroom. Internet High. Educ. **33**, 74–85 (2017)
10. Matcha, W., Gašević, D., Uzir, N.A., Jovanović, J., Pardo, A.: Analytics of learning strategies: associations with academic performance and feedback. In: Proceedings of the 9th International Conference on Learning Analytics and Knowledge, pp. 461–470 (2019)
11. Weinstein, C.E., Husman, J., Dierking, D.R.: Self-regulation interventions with a focus on learning strategies. Handb. Self-Regulation. **22**, 727–747 (2000)
12. Hadwin, A.F., Nesbit, J.C., Jamieson-Noel, D., Code, J., Winne, P.H.: Examining trace data to explore self-regulated learning. Metacogn. Learn. **2**, 107–124 (2007)
13. Derry, S.J.: Putting learning strategies to work. Educ. Leadersh. **47**, 4–10 (1989)
14. Winne, P.H., Jamieson-Noel, D., Muis, K.: Methodological issues and advances in researching tactics, strategies, and self-regulated learning (2002)
15. Winne, P.H., Gupta, L., Nesbit, J.C.: Exploring individual differences in studying strategies using graph theoretic statistics. Alberta J. Educ. Res. **40**, 177–193 (1994)
16. Siadaty, M., Gašević, D., Hatala, M.: Associations between technological scaffolding and micro-level processes of self-regulated learning: a workplace study. Comput. Human Behav. **55**, 1007–1019 (2016). Part B
17. Boroujeni, M.S., Dillenbourg, P.: Discovery and temporal analysis of latent study patterns in MOOC interaction sequences. In: Proceedings of the 8th International Conference on Learning Analytics and Knowledge, pp. 206–215. ACM, New York (2018)
18. Sobocinski, M., Malmberg, J., Järvelä, S.: Exploring temporal sequences of regulatory phases and associated interactions in low- and high-challenge collaborative learning sessions. Metacogn. Learn. **12**, 275–294 (2017)

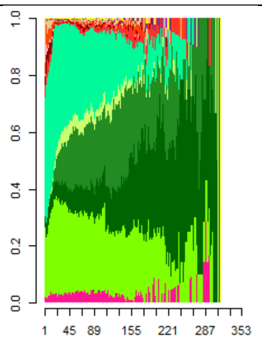
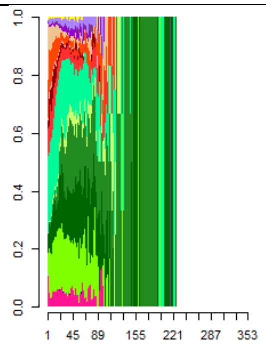
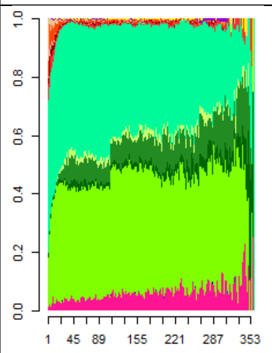
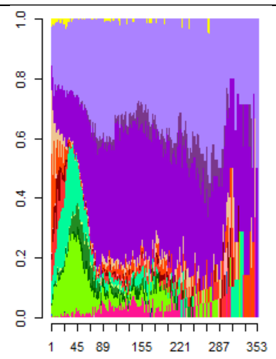
19. Maldonado-Mahauad, J., Pérez-Sanagustín, M., Kizilcec, R.F., Morales, N., Munoz-Gama, J.: Mining theory-based patterns from big data: identifying self-regulated learning strategies in massive open online courses. Comput. Human Behav. **80**, 179–196 (2018)
20. Kovanović, V., Gašević, D., Dawson, S., Joksimović, S., Baker, R.S., Hatala, M.: Penetrating the black box of time-on-task estimation. In: the Fifth International Conference on Learning Analytics And Knowledge, pp. 184–193 (2015)
21. Gabadinho, A., Ritschard, G., Studer, M., Muller, N.S.: Mining sequence data in R with the TraMineR package: a user's guide, vol. 1, pp. 1–124. Department of. Economic Labor and Demographic, University of Geneva, Switzerland (2008)
22. Gatta, R., Lenkowicz, J., Vallati, M., Stefanini, A.: pMineR: processes mining in medicine (2017). https://cran.r-project.org/package=pMineR
23. Shaffer, D.W., Collier, W., Ruis, A.R.: A tutorial on epistemic network analysis: analyzing the structure of connections in cognitive, social, and interaction data. J. Learn. Anal. **3**, 9–45 (2016)
24. Biggs: Student Approaches to Learning and Studying (1987)
25. Zhou, M., Winne, P.H.: Modeling academic achievement by self-reported versus traced goal orientation. Learn. Instr. **22**, 413–419 (2012)

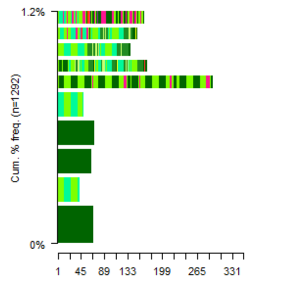
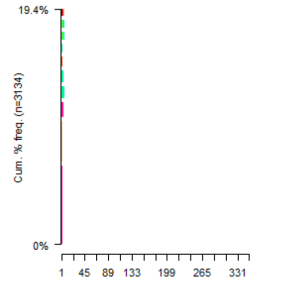
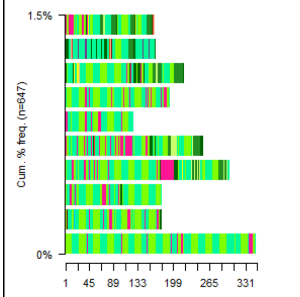
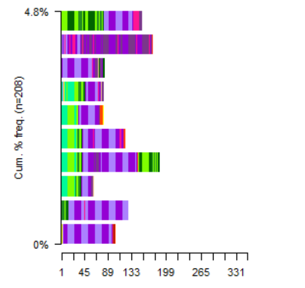
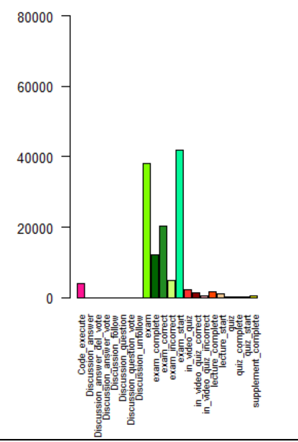
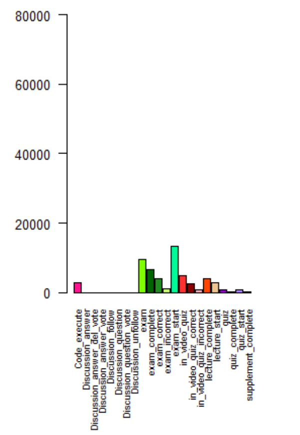
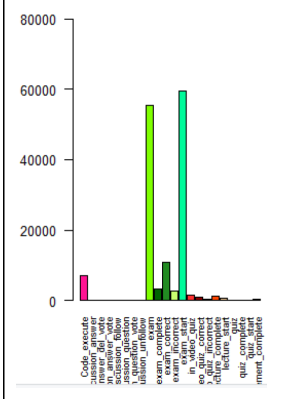
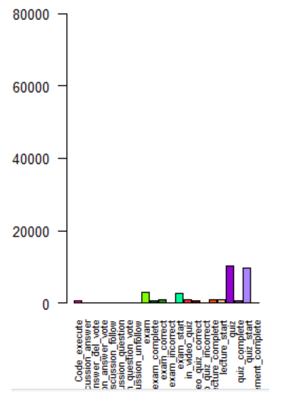
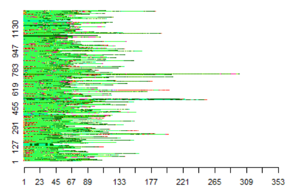
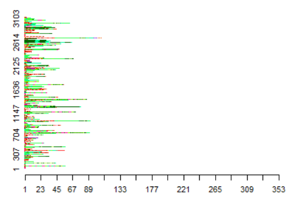
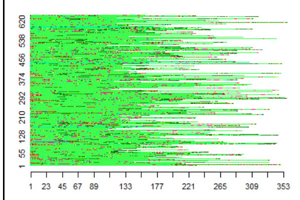
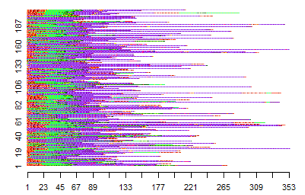
Detection of Learning Tactics: A comparison of Process, Sequence and Network Analytic Approaches (Supplementary Documents)

1. Tactics Detection

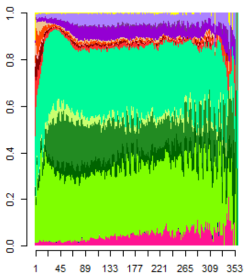
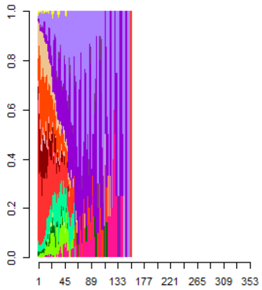
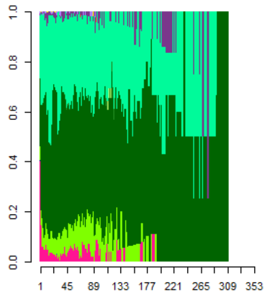
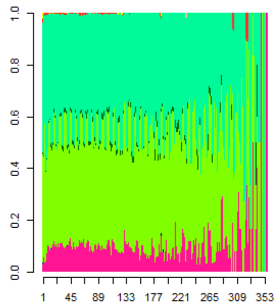
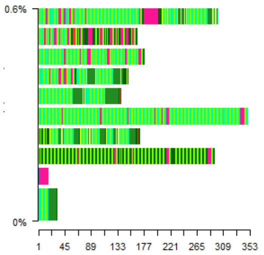
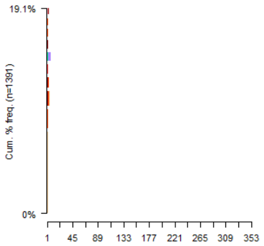
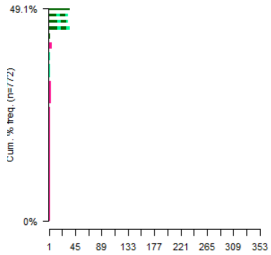
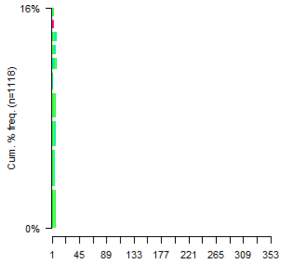
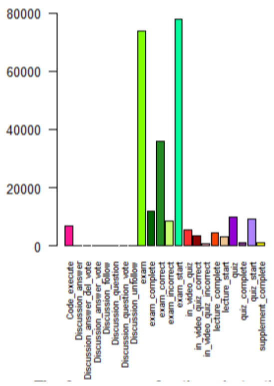
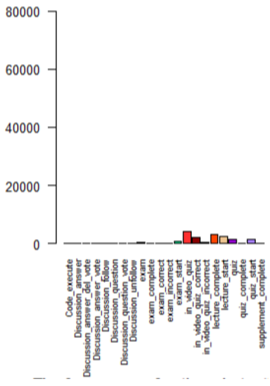
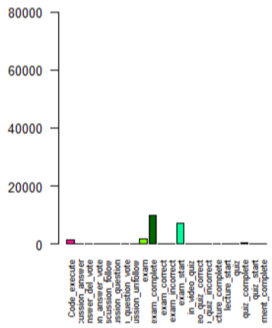
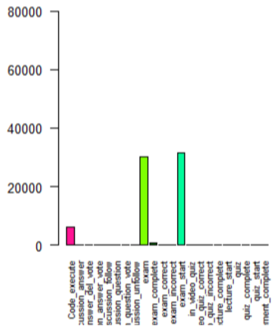
Method1: Sequence Analytics Approach

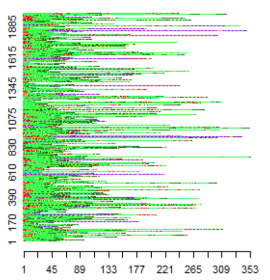
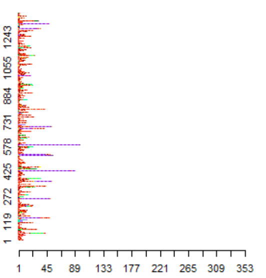
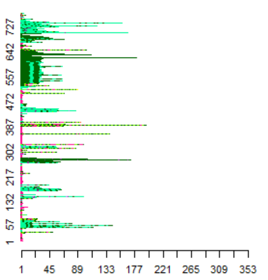
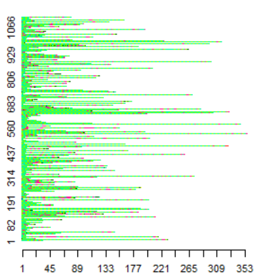


Name	Tactic1: Short Practical Oriented (short session of Exam and code execute Oriented)	Tactic2: Practice- Lecture Oriented (short session of Exam, code execute and lecture Oriented)	Tactic3: Long Practical Oriented (long sequences of exam and code execute)	Tactic4: Diverse- Assessment Oriented (Quiz, Lecture and Exam Oriented)
Sessions	N= 1292 Sessions (24.47 % of all learning sessions)	N = 3134 Sessions (59.34 % of all learning sessions)	N= 647 Sessions (12.25 % of all learning sessions)	N= 208 Sessions (3.94 % of all learning sessions)
Stat Distribution				

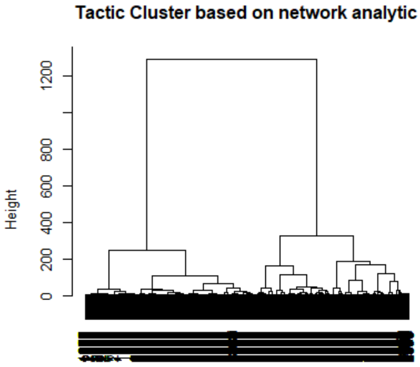
Most Frequent Sequences				
Frequency of Actions				
All Actions in each Learning Session				
Sequences Length	Min. : 33.0 1st Qu.: 68.0 Median : 93.0 Mean :100.1 3rd Qu.:119.0 Max. :315.0	Min. : 2.00 1st Qu.: 4.00 Median : 10.00 Mean : 17.47 3rd Qu.: 25.00 Max. :228.00	Min. :103.0 1st Qu.:164.0 Median :214.0 Mean :221.9 3rd Qu.:278.5 Max. :359.0	Min. : 54.0 1st Qu.:100.8 Median :139.0 Mean :155.7 3rd Qu.:197.8 Max. :355.0

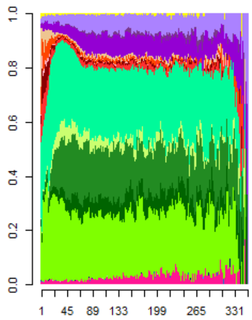
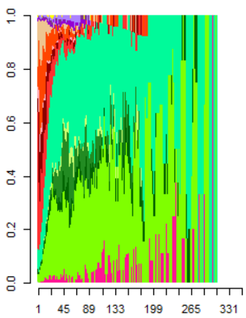
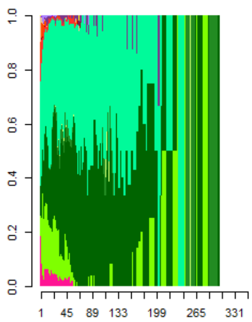
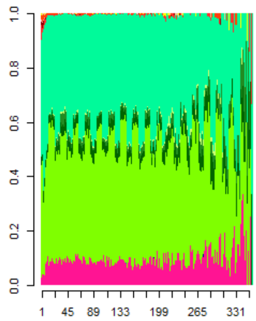
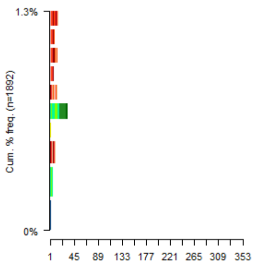
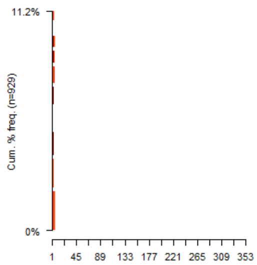
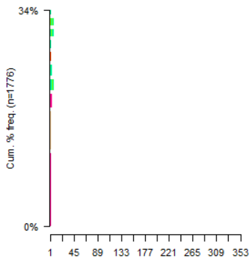
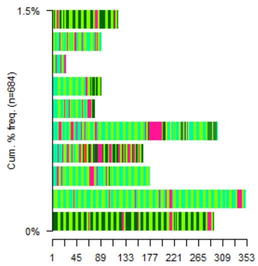
Method2: Process Analytic Approach

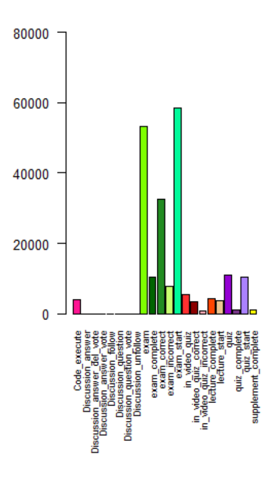
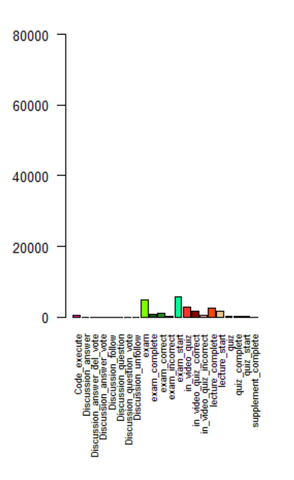
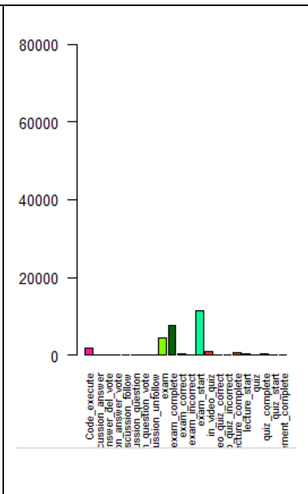
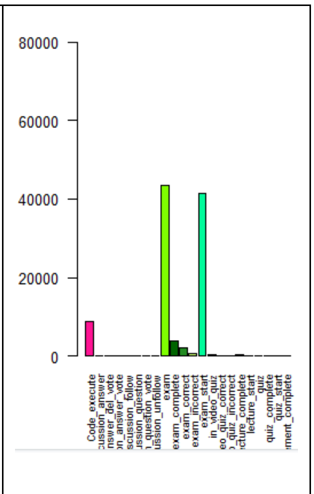
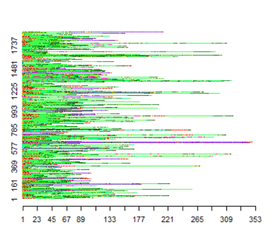
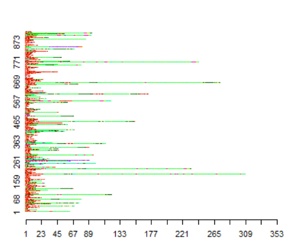
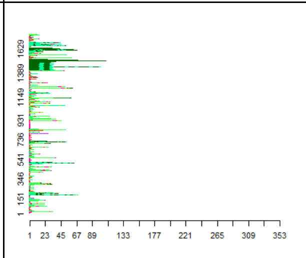
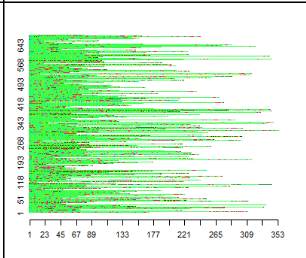
Name	Tactic1: Diverse Oriented	Tactic2: Lecture Oriented (Lecture and Quizzes Oriented)	Tactic3: Short Practice Oriented (Exam and Coding Oriented)	Tactic4: Long - Practice Oriented (Code Executing and Exam Oriented)
Sessions	N= 2000 Sessions (37.87 % of all learning sessions)	N = 1391 Sessions (26.34 % of all learning sessions)	N = 772Sessions (14.62 % of all learning sessions)	N= 1118 Sessions (21.17 % of all learning sessions)
Stat Distribution				
Most Frequent Sequences				
Frequency of Actions				

All Actions in each Learning Session				
Sequences Length	Min. : 3.0 1st Qu.: 63.0 Median :105.0 Mean :126.8 3rd Qu.:172.0 Max. :359.0	Min. : 2.00 1st Qu.: 4.00 Median : 7.00 Mean : 12.02 3rd Qu.: 14.00 Max. :156.00	Min. : 2.00 1st Qu.: 2.00 Median : 8.00 Mean : 26.48 3rd Qu.: 35.00 Max. :310.00	Min. : 2.0 1st Qu.: 9.0 Median : 31.0 Mean : 61.9 3rd Qu.: 91.0 Max. :357.0

Method3: Network Analytics



Name	Tactic1: Diverse Oriented	Tactic2: Lecture and Practice Oriented (Exam and Lecture Oriented)	Tactic3: Short Practice Oriented (Exam and Coding Oriented)	Tactic4: Long - Practice Oriented (Code Executing and Exam Oriented)
Sessions	N= 1892 Sessions (35.83 % of all learning sessions)	N = 929 Sessions (17.59 % of all learning sessions)	N= 1776 Sessions (33.63 % of all learning sessions)	N= 684 Sessions (12.95 % of all learning sessions)
Stat Distribution				
Most Frequent Sequences				

Frequency of Actions				
All Actions in each Learning Session				
Sequences Length	Min. : 2.0 1st Qu.: 53.0 Median : 93.0 Mean :110.2 3rd Qu.:148.0 Max. :355.0	Min. : 2.00 1st Qu.: 6.00 Median : 11.00 Mean : 24.67 3rd Qu.: 22.00 Max. :309.00	Min. : 2.00 1st Qu.: 2.00 Median : 7.00 Mean : 15.63 3rd Qu.: 21.00 Max. :305.00	Min. : 11.00 1st Qu.: 73.75 Median :126.00 Mean :147.56 3rd Qu.:212.00 Max. :359.00

Similarity of tactics detection

Summary: Process Analytics and Network Analytics

Similar sequences: 3526 Sessions (66.77%)

	Network – Diverse (100%)	Network - Lecture-Based (100%)	Network - Long-Practice (100%)	Network - Short-Practice (100%)
Process - Diverse	85.68	17.33	25.44	2.48
Process - Lecture-Based	10.94	78.26	0	25.73
Process - Long-Practice	2.11	3.34	69.44	32.21
Process - Short-Practice	1.27	1.08	5.12	39.58

Summary: Sequence Analytics and Network Analytics

Similar sequences: 1500 Sessions (28.40 %)

	Network – Diverse (100%)	Network - Lecture-Based (100%)	Network - Long-Practice (100%)	Network - Short-Practice (100%)
Sequence - Diverse	10.84	0.32	0	0
Sequence - Lecture-Based	29.49	89.56	14.62	92.57
Sequence - Long-Practice	14.64	2.26	49.71	0.51
Sequence - Short-Practice	45.03	7.86	35.67	6.93

Summary: Sequence Analytics and Process Analytics

Similar sequences: 1861 Sessions (35.24%)

	Process – Diverse (100%)	Process - Lecture-Based (100%)	Process - Long-Practice (100%)	Process - Short-Practice (100%)
Sequence - Diverse	9.25	1.65	0	0
Sequence - Lecture-Based	22.4	98.35	58.94	85.36
Sequence - Long-Practice	21.6	0	18.34	1.3
Sequence - Short-Practice	46.75	0	22.72	13.34

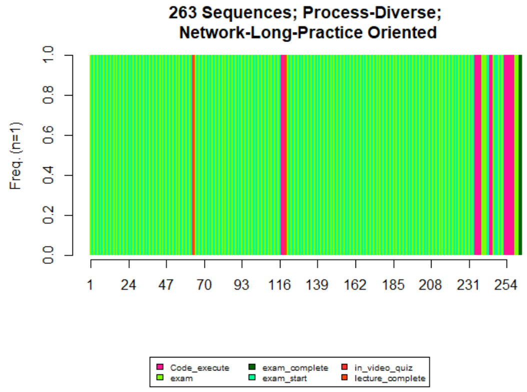
Case Study 1:

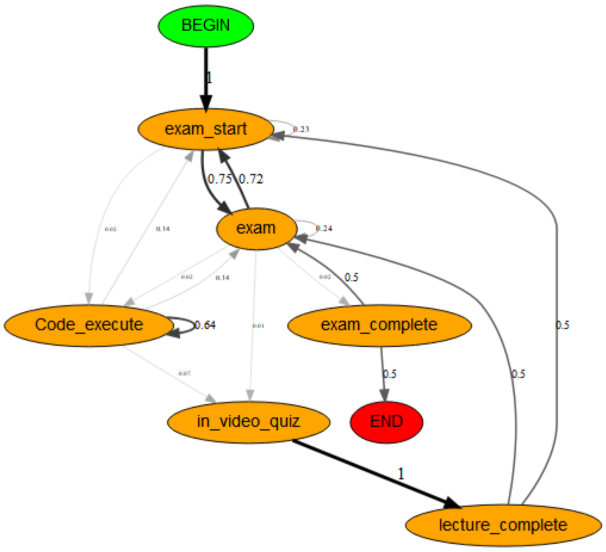
Sequence Length: 263 actions (Seq 10595)

Sequence Approach: Long-Practice Oriented

Process Approach: Diverse Oriented

Network Approach: Long-Practice Oriented





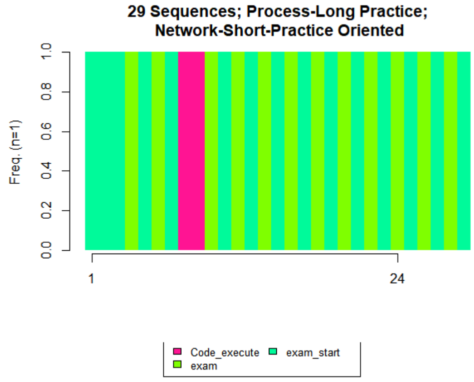
Case Study 2:

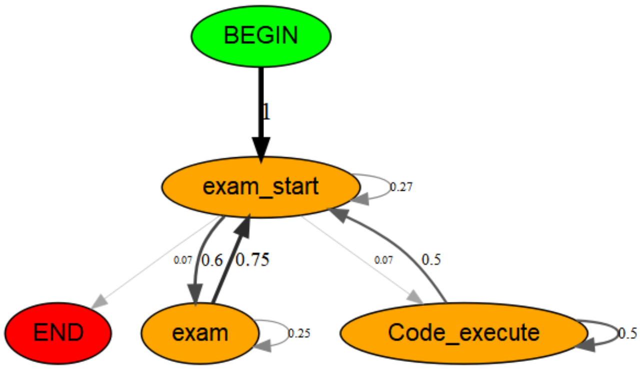
Sequence Length: 29 (Seq13745)

Sequence Approach: Lecture-Based Oriented

Process Approach: Long-Practice Oriented

Network Approach: Short-Practice Oriented





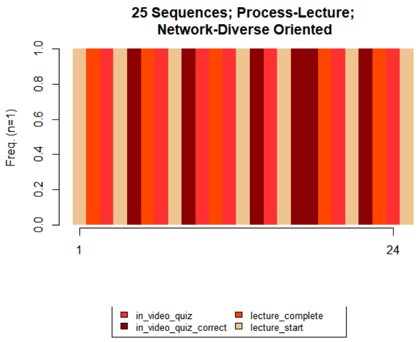
Case Study 3:

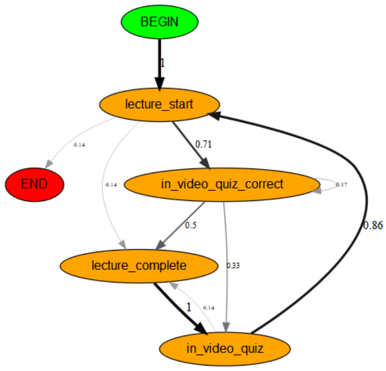
Sequence Length: 25 (Seq11934)

Sequence Approach: Lecture-Based Oriented

Process Approach: Lecture Oriented

Network Approach: Diverse Oriented





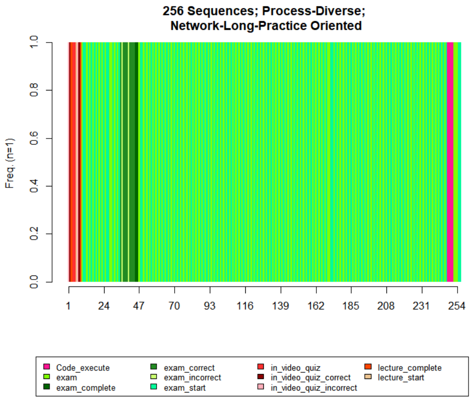
Case Study 4:

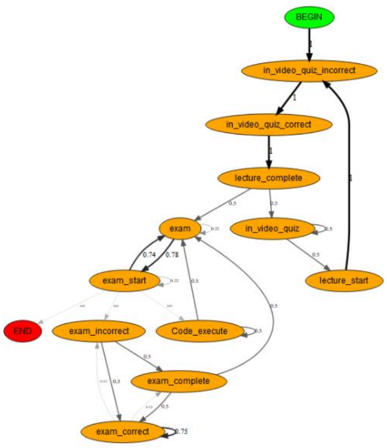
Sequence Length: 256 actions

Sequence Approach: Long-Practice Oriented

Process Approach: Diverse Oriented

Network Approach: Long-Practice Oriented





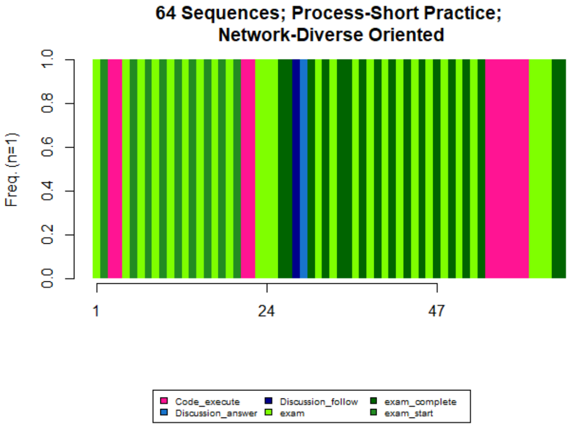
Case Study 5:

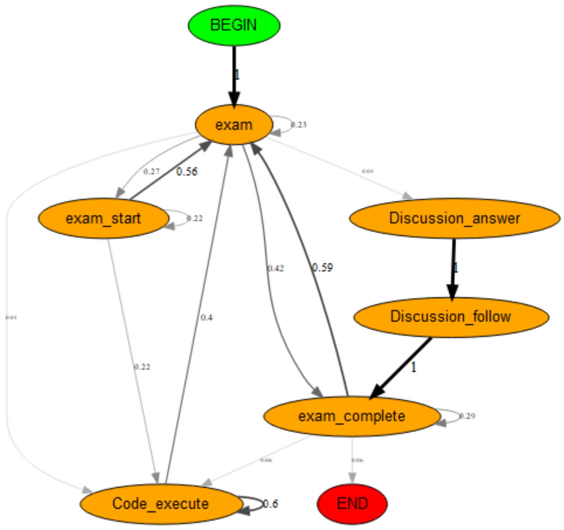
Sequence Length: 64 (Seq22131)

Sequence Approach: Short-Practice Oriented

Process Approach: Short-Practice Oriented

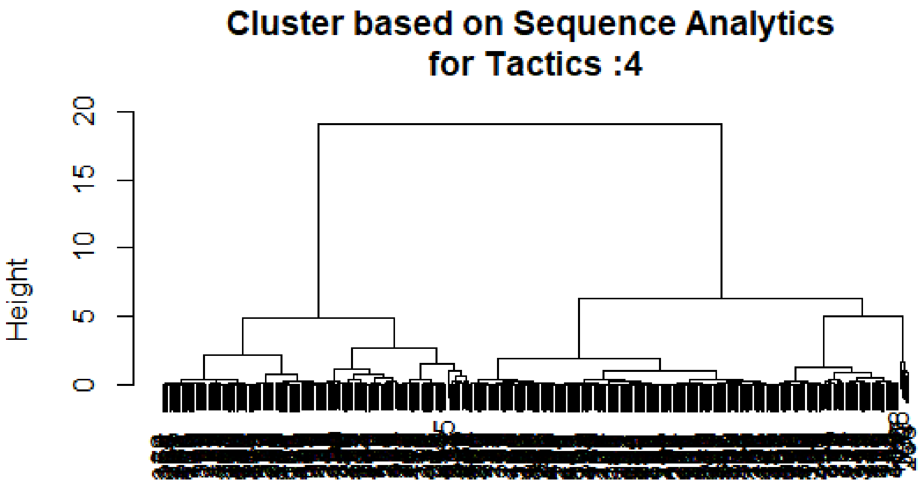
Network Approach: Diverse Oriented

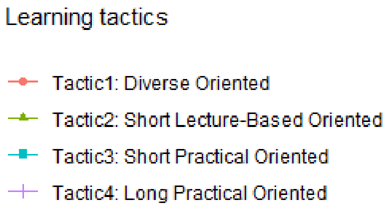
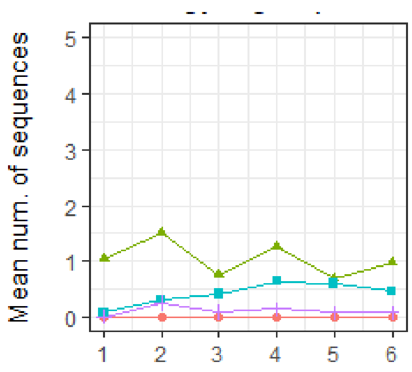
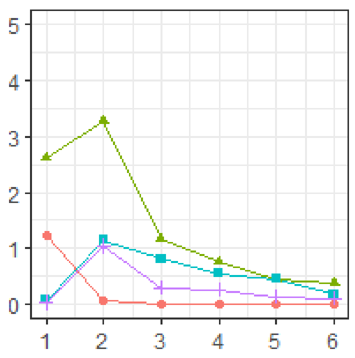
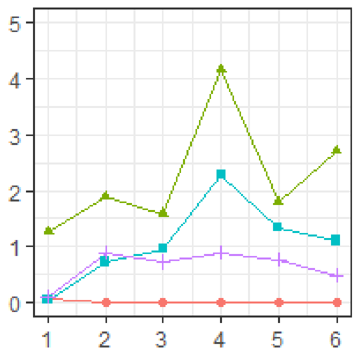




Strategy Detections

Method1: Sequence Analytics



<table>
<tr><td>Name</td><td>Strategy Group1:</td><td>Strategy Group2:</td><td>Strategy Group3:</td></tr>
<tr><td></td><td colspan="3">
</td></tr>
<tr><td>Mean num. of sequences/weekly topic</td><td>
</td><td></td><td></td></tr>
<tr><td>Students</td><td>N= 151 Students
(41.03 % of all students)</td><td>N= 151 Students
(41.03 % of all students)</td><td>N= 66 Students
(17.94 % of all students)</td></tr>
<tr><td>Pass/Fail Rate</td><td>Fail: 117 students
Pass: 34 Students

%Pass = 22.52</td><td>Fail: 134 students
Pass: 17 Students

%Pass = 11.26</td><td>Fail: 36 students
Pass: 30 Students

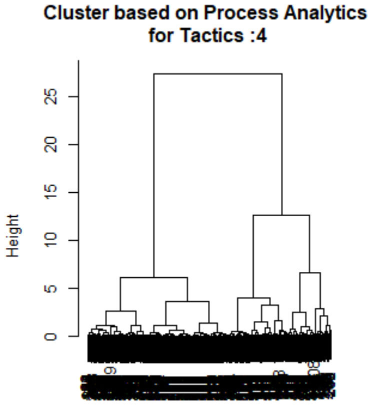
%Pass = 45.45</td></tr>
</table>

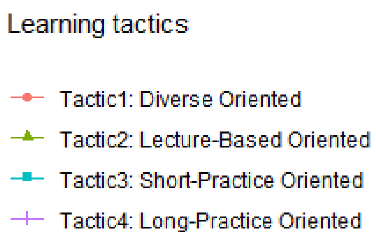
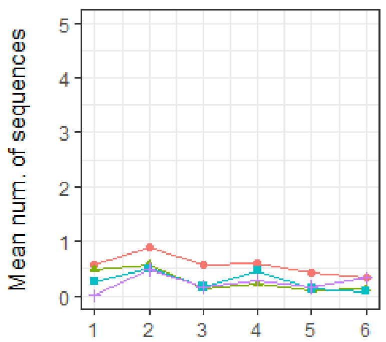
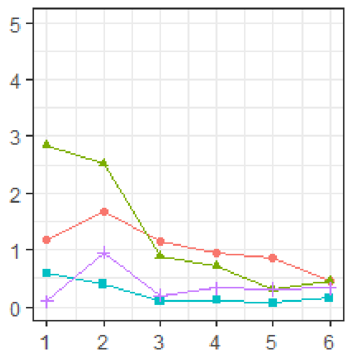
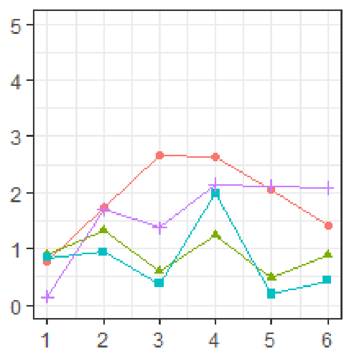
Item Passed (Max = 22)	Mean = 10.50 Median = 9 SD = 7.90	Mean = 7.77 Median = 4 SD = 7.19	Mean = 16.64 Median = 20 SD = 6.59
Course Grade (Max = 100)	Mean = 42.97 Median = 36.14 SD = 33.53	Mean = 31.29 Median = 18.04 SD = 30.37	Mean = 67.79 Median = 82.86 SD = 27.36

PairWise comparison between strategy groups

Item	Strategy	Strategy	Z	p	r
Course Grade	S1	S2	3.254	0.001	0.187
	S1	S3	-4.595	<0.001	0.312
	S2	S3	-7.155	<0.001	0.486
Passed Course	S1	S2	2.607	0.014	0.150
	S1	S3	-3.401	<0.001	0.231
	S2	S3	-5.613	<0.001	0.381
Passed Graded Items	S1	S2	3.145	0.0014	0.181
	S1	S3	-5.101	<0.001	0.346
	S2	S3	-7.389	<0.001	0.502

Method2: Process Analytics



<table>
<tr><td>Name</td><td>Strategy Group1:</td><td>Strategy Group2:</td><td>Strategy Group3:</td></tr>
<tr><td></td><td colspan="3">
</td></tr>
<tr><td>Mean num. of sequences/weekly topic</td><td>
</td><td>
</td><td>
</td></tr>
<tr><td>Students</td><td>N= 215 Students
(58.42 % of all students)</td><td>N= 89 Students
(24.18 % of all students)</td><td>N= 64 Students
(17.39 % of all students)</td></tr>
<tr><td>Pass/Fail Rate</td><td>Fail: 178 students
Pass: 37 Students

%Pass = 17.21</td><td>Fail: 73 students
Pass: 16 Students

%Pass = 17.98</td><td>Fail: 36 students
Pass: 28 Students

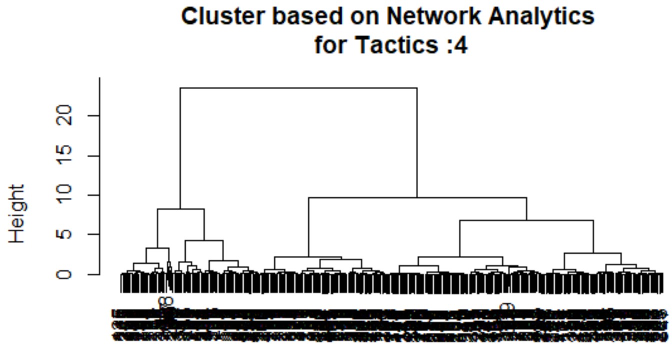
%Pass = 43.75</td></tr>
</table>

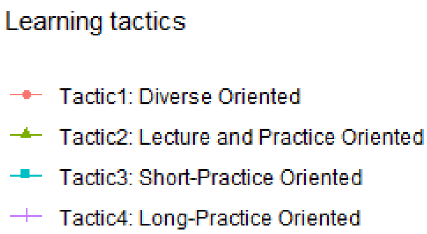
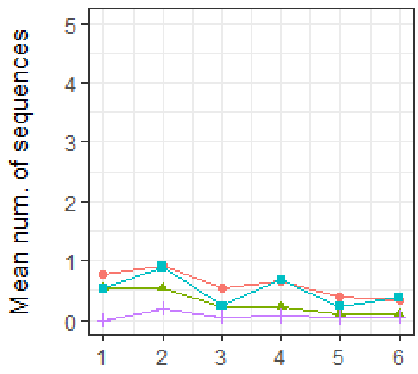
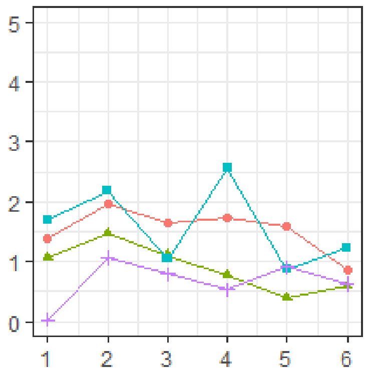
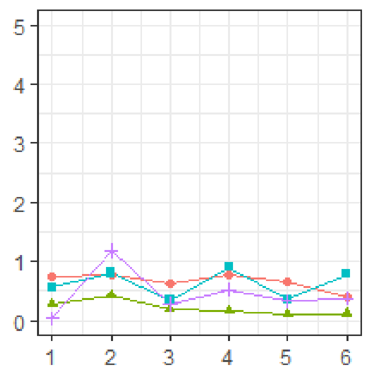
Item Passed (Max = 22)	Mean = 9.06 Median = 7 SD = 7.62	Mean = 8.90 Median = 5 SD = 7.85	Mean = 17.47 Median = 20 SD = 5.47
Course Grade (Max = 100)	Mean = 36.70 Median = 29.33 SD = 32.04	Mean = 36.37 Median = 20.41 SD = 33.66	Mean = 71.26 Median = 82.71 SD = 23.24

PairWise comparison between strategy groups

Item	Strategy	Strategy	Z	p	r
Course Grade	S1	S2	0.108	0.93	0.006
	S1	S3	-6.927	<0.001	0.415
	S2	S3	-5.879	<0.001	0.475
Passed Course	S1	S2	-0.160	0.88	0.009
	S1	S3	-4.401	<0.001	0.263
	S2	S3	-3.463	<0.001	0.28
Passed Graded Items	S1	S2	0.102	0.87	0.006
	S1	S3	-7.206	<0.001	0.431
	S2	S3	-6.359	<0.001	0.514

Method3: Network Analytics



Name	Strategy Group1:	Strategy Group2:	Strategy Group3:
			
Mean num. of sequences/weekly topic			
Students	N= 188 Students (51.09 % of all students)	N= 94 Students (25.54 % of all students)	N= 86 Students (23.37 % of all students)
Pass/Fail Rate	Fail: 157 students Pass: 31 Students %Pass = 16.49	Fail: 66 students Pass: 28 Students %Pass = 29.79	Fail: 64 students Pass: 22 Students %Pass = 25.58

Item Passed (Max = 22)	Mean = 8.78 Median = 7 SD = 7.63	Mean = 13.36 Median = 15 SD = 8.02	Mean = 11.06 Median = 9 SD = 7.91
Course Grade (Max = 100)	Mean = 35.63 Median = 27.29 SD = 32.30	Mean = 54.48 Median = 56.95 SD = 33.66	Mean = 44.99 Median = 37.04 SD = 33.16

PairWise comparison between strategy groups

Item	Strategy	Strategy	Z	p	r
Course Grade	S1	S2	-4.088	<0.001	0.243
	S1	S3	-2.756	0.058	0.166
	S2	S3	1.448	0.064	0.108
Passed Course	S1	S2	-2.440	0.05	0.146
	S1	S3	-1.765	0.18	0.107
	S2	S3	0.516	0.57	0.039
Passed Graded Items	S1	S2	-4.323	<0.001	0.258
	S1	S3	-2.762	0.05	0.167
	S2	S3	1.613	0.059	0.121

3.3 Summary

The study presented in this chapter examined three data analytics approaches of detecting learning tactics and strategies to address research question two (RQ2) as stated in Section 1.1. The sequence analytics-based approach was developed by considering a learning tactic as a sequence of learning actions exhibited within a learning session by Jovanovic et al. (2017). The process analytics-based approach was developed by considering learning tactics as a process of learning that student performed within a learning session. The network analytics-based approach considered learning tactic as the co-occurrence of learning actions performed by a learner within a learning session. The corresponding unsupervised machine learning algorithms were used differently in each of the three approaches to detect similar patterns of learning actions within a learning session. Based on different structures of the three analytics-based approaches to detecting learning tactics, learning strategies were captured by using the agglomerative hierarchical clustering technique to detect patterns of how student employ the detected learning tactics.

Using the dataset about learning activities of MOOC learners enrolled in the Introduction to Python course, the study presented in this chapter revealed that the three approaches detected similar learning patterns to some extend. That is, by examining common characteristics of each learning tactic, the sequence analytics-based approach characterised the learning patterns indicative of tactics clustered 35.24% of learning sessions in the same way as the process analytics-based approach did and 28.40% as the network analytics-based approach did. The process analytics-based approach clustered 66.77% in the same way as the network analytics-based approach did. The agglomerative hierarchical clustering algorithms suggested three groups of learning strategies for each analytics approach. By examining the patterns of each strategies group, we found that sequence analytics and process analytics-based approach detected three similar patterns of learning strategies identified as *Highly active and multiple tactics used, Highly active at the beginning,* and *Surface engagement* strategy groups. The network analytics-based approach detected two strategy groups indicative of *Surface engagement* and one *Highly active and multiple tactics used* strategy groups.

We examine whether the detected learning strategies are supported by relevant educational theory as identified in research question three (RQ3) in Section 1.1. The results show that the learning strategies detected using three different analytics-based approaches coincide with the approaches to learning (Biggs, 1987; Entwistle, 1991). That is, by applying the sequence analytics and process analytics-based approaches, three approaches to learning, including, deep, surface and strategic approaches were detected. Meanwhile, network analytics-based approach detected the deep and surface approaches to learning. The deep approach to learning is characterised by the application of multiple tactics and high level of engagement (Biggs, 1987). Deep approach to learning was reflected through the use of *Highly active and multiple tactics used* strategy. The learners who demonstrated the use of deep approach to learning showed the highest academic performance which was in line with findings reported in the literature of approaches to learning (Chonkar et al., 2018; Mattick et al., 2004). In contrast, the approaches to learning explain the characteristics of surface learners as the learners who intend to obtain the high score with low level of effort (Entwistle, 1991). This characteristic of the surface approach to learning was observed in the *Surface engagement* strategy group. The learners who followed the *Surface engagement* strategy focused on assessment activities with a low level of engagement. Hence, a low level of performance and high dropouts were observed for this group of learners.

The work presented in this chapter demonstrates that three learning analytics-based approaches can be used to detect learning patterns indicative of learning tactics and strategies used by MOOC learners. The work presented in this chapter demonstrates that three learning analytics-based approaches can be used to detect learning patterns indicative of learning tactics and strategies used by MOOC learners. The generalisation can be made from the detected results is that regardless of the applied approaches, the tactics are representative of the course design, as demonstrated in this study and posited by previous research (Gašević et al., 2016). The Python course focused on practical exercise; hence, the long and short practice-oriented tactics were detected by the three analytics approaches. Nonetheless, future research is needed in order to examine of the generalizability of the results across different learning contexts.

Even though, each analytics-based approach demonstrated the ability to detect meaningful learning tactics and strategies, the rest of the thesis will continue to explore the use of process analytics-based approach to capture learning tactics and strategies for reasons described below:

i The ability to detect distinct forms of learning tactics: For instance, in this study, the Introduction of Python course was designed based on problem-solving exercise. Hence, the main learning activities were centred around exam-related activities. Almost every detected learning tactic involved exam-related activities, except the *Lecture oriented* tactic which was only observable through the use of process analytics-based approach. The ability of this approach in detecting distinct forms of learning tactics warrants further exploration.

ii The distribution of the observed items (i.e. the number of sessions in each tactic group): The distribution of the learning sessions that were categorised into each tactic consequently impacts on the detection of learning strategies. For instance, the sequence analytics-based approach categorised 59.34% of the learning sessions as the *Practice-Lecture oriented* tactic and only 3.94% of the learning sessions were recognised as the *Diverse-Assessment oriented* tactics. As a result, when exploring the patterns of learning strategies, we were not able to detect the enactment of the *Diverse-Assessment oriented* tactics due to the small number of detected learning sessions indicative of this tactic. In contrast, the process analytics-based approach was capable of detecting a reasonably even distribution of learning tactics across learning sessions. Hence, the enactment of all learning tactics could be observed.

iii The ability to detect the patterns of all the strategy groups: The process analytics-based approach, similar to the other two analytics-based approaches, enables the detection of the deep and surface approaches to learning based on the identified strategies. Moreover, an additional learning strategy group was identified with the process approach and labeled as *Highly active at the beginning*. A similar result was produced using the sequence analytics-based approach. In contrast, the network analytics-based approach could not detect the *Highly active at the beginning* strategy group. Therefore, the process analytics-based approach demonstrated the ability to detect most of the observable patterns of learning strategies, thus, worthy of a further examination.

Due to the above-mentioned advantages, the rest of the thesis focuses on the study of the process analytics-based approach.

The study presented in this chapter is based on the MOOC context; hence, the applicability of the process analytics-based approach in different learning contexts such as blended learning or flipped classroom still needs a further examination. Moreover, how the SRL constructs contribute to the application of learning tactics and strategies, as detected by the analytics approaches, is yet unexplored. The following chapter examines the use of process analytics-based approach to detecting learning tactics and strategies in the flipped classroom context. Moreover, the chapter investigates the association of the SRL constructs i.e. feedback and products (academic performance) with the choice of learning tactics and strategies.

4 Analytics of Learning Strategy: Associations with Academic Performance and Feedback

> Being supportive and building students' confidence is not accomplished by blindly telling them they are doing a great job every day. It involves assessing weaknesses and strengths and delivering feedback in a timely manner so that they can build their skills to complete the task at hand.
>
> — Oran Tkatchov, *Success for Every Student: A Guide to Teaching and Learning*

4.1 Introduction

Self-regulated learning is an iterative process that is influenced by several factors. Students employ learning tactics and strategies by considering the learning tasks that they are dealing with, the cognitive conditions they have, and feedback they receive (Winne & Hadwin, 1998). Students generate internal feedback when they evaluate products of their learning against an 'expectation' or 'standard' and learning goal (Pardo et al., 2017). Feedback can influence how students perceive and adopt learning strategies (Hattie & Timperley, 2007). Receiving feedback may result in the adaptation of learning tactics and strategies or changes of expectations and learning goals (Greene & Azevedo, 2007). However, this type of feedback is not always accurate. For instance, Bjork et al. (2013) reports that students with high performance tend to underestimate their learning while students with low performance are more likely to overestimate their learning. This over and under-confidence may lead to the use of sub-optimal learning tactics and strategies.

Previous research on learning strategies has revealed that students often lack the ability to choose and adapt their strategies to the course requirements (Lust et al., 2013). This implies that students need guidance in effective strategy usage. Similarly, Dunlosky et al. (2013) assert that guiding students to employ effective learning tactics and strategies is a promising method to improve learning. Hence, research suggests that students need feedback provided by instructors. Nonetheless, research into the impact of feedback on student uptake of learning strategies is hardly present. Moreover, offering personalised guidance based on the detection of strategies adopted by learners remains a challenge (Matcha, Ahmad Uzir, et al., 2020). Simply providing feedback without customisation to the individual needs and performance has proven to be ineffective (Kizilcec et al., 2016). Effective feedback needs to address the regulation process of learning (Hattie & Timperley, 2007) in order to improve the learning process and learning outcome (Pardo et al., 2017). The extensive use of technology in contemporary education offers opportunities to address this gap by making use of learning analytics tools to provide feedback. Two types of learning analytics-based tools have been adopted widely, including LADs and customised-text based feedback or also referred to as "personalised feedback" in this thesis.

Unlike LADs (explored in Chapter two) that present feedback through the graphical visualisation (Bodily & Verbert, 2017; Schwendimann et al., 2016), personalised feedback is provided by sending customised textual messages to individual student (Pardo et al., 2017); customised textual messages are semi-automatically generated. That is, feedback is generated by processing the results of the analysis of students data with a set of rules, which is manually written by instructor, to automatically produce the customised message. Although, the literature suggests that this type of feedback shows promising results (Pardo et al., 2019), the studies evaluating its effectiveness are scarce. This study aims to fill this gap in the literature by exploring the association of the provision of LAD and personalised feedback with the choices of learning tactics and strategies. More specifically, this chapter examines the learning tactics and strategies detection by using the process analytics-based approach by focusing on how the approach enables to capture the learning tactics and strategies in different learning context i.e. flipped classroom.

4.1.1 Chapter overview

In this chapter, we demonstrate the use of the process analytics-based approach to capture learning tactics and strategies; the approach has previously been explained in Chapter three. In this chapter, learning tactics detected by using the process analytics-based approach are explained in detail. That is, the analysed results are interpreted by considering the sequence of actions and the process model of learning actions. The chapter also studies the association of the provision of personalised feedback with the choices of learning tactics and strategies. Moreover, the chapter explores the association of the choices of learning strategies and academic performance.

A significant contribution of this thesis is that it demonstrates the importance of considering the sequential and temporal dimensions of learning tactics and strategies. By exploring the sequence of actions and considering the process as a temporal dimension of detected learning tactics, this thesis shows that the learning patterns detected with the proposed process analytics-based approach reflect the learning design; that is, the learning patterns are indicative of tactics used to complete course tasks. Learning strategies were detected based on patterns of how students employed the tactics. Then, the detected learning strategies were theoretically interpreted with respect to approaches to learning (Biggs, 1987; Entwistle, 1991) as a response to research question three (RQ3) in Section 1.1. Moreover, the associations of the learning strategies with academic performance were analysed.

Another contribution of this work is that the process analytics-based approach proposed in Chapter three was applied to a different learning context, i.e. flipped classroom. Data used in the present chapter was obtained from online interactions of the students who participated in a course that was designed according to a flipped classroom model. The key concept of the flipped classroom is that it highly emphasises on promoting active learning and consists of two main parts: pre-class preparation and in-class activities (Freeman et al., 2014; Pardo et al., 2018). Introduction to relevant concepts and other learning tasks that require subject matter skills such as understanding and remembering of facts, typically take place online during the preparation phase (Pardo, 2018). The in-class sessions are typically designed to develop high-order thinking skills (Freeman et al., 2014; O'Flaherty & Phillips, 2015; Pardo, 2018).

Table 3. Description of the learning actions recorded in the trace data and the corresponding SRL phases

SRL Phases	Learning actions observed in the trace data	Description
1: Goal setting and planning	MC_ORIENT	Access to the course schedule and learning objective pages; this is considered a metacognitive orientation action
2: Tasks identification		
3: Enactment of learning tactics and strategies	EXE_CO	A correctly solved exercise item
	EXE_IN	An incorrectly solved exercise item
	MCQ_CO	A correctly solved MCQ item
	MCQ_IN	An incorrectly solved MCQ item
	MCQ_SR	A solution requested for an MCQ item
	VIDEO_LOAD	Loading the video
	VIDEO_PL	Play the course video
	VIDEO_PA	Pause the video
	VIDEO_END	End of the video
	CONTENT_ACCESS	Access to a page that contains reading materials
	INDEX_ACCESS	Access to the index page
	PROJECT_ACCESS	Access to the project information page
	SEARCH	Search for keyword
	DOWNLOAD	Download documents or files
4: Adaption	MC_EVAL	Access to the dashboard; this is considered a metacognitive evaluation action

Table 3 presents the descriptions of learning actions collected according to the SRL phases. SRL comprises a recursive cycle of the learning process (Winne & Hadwin, 1998). Most of the collected learning actions represented the operations carried out during phase 3 (enactment of learning tactics and strategies) of the Winne and Hadwin model of SRL. In addition, one of the aims of this study is to examine the association of feedback and learning strategies. Hence, actions representing phase 1 (Goal setting and planning) and phase 4 (adaption) of the Winne and Hadwin model of

SRL, both of which were influenced by students' evaluation process through the feedback provision, were also considered in this study. However, phase 2 (tasks identification) of the Winne and Hadwin model of SRL was not observed in this study. Task identification involves the understanding of task requirements and objectives; and choosing the relevant tactics and strategies to complete the task (Winne & Hadwin, 1998). The course was designed to help the students during the task identification phase by providing the instruction on how to complete the tasks. However, the trace data did not record the students' decisions of how they intended to complete the learning tasks.

Another major contribution of the work presented in this chapter explores the association of analytics-based feedback, which represents the external evaluation construct according to the Winne and Hadwin's (1998) model of SRL, with the choices of learning strategies. Two types of analytics-based feedback were provided through the use of LAD and personalised feedback. This investigation contributes to answering research question four (RQ4 in Section 1.1) which aims to explore how SRL constructs, particularly, the external evaluation (i.e. feedback) contribute to the adoption of learning strategies.

4.2 Publication: Analytics of Learning Strategies: Associations with Academic Performance and Feedback

The following section includes the verbatim copy of the following publication:

> Matcha, W., Gašević, D., Ahmad Uzir, N., Jovanović, J., & Pardo, A. (2019). Analytics of Learning Strategies: Associations with Academic Performance and Feedback, *Proceedings of the 9th international conference on learning analytics & knowledge*. https://doi.org/10.1145/3303772.3303787

Analytics of Learning Strategies: Associations with Academic Performance and Feedback

Wannisa Matcha
School of Informatics,
University of Edinburgh,
United Kingdom
w.matcha@ed.ac.uk

Dragan Gašević
School of Informatics,
University of Edinburgh,
United Kingdom,
dragan.gasevic@ed.ac.uk

Nora'Ayu Ahmad Uzir
School of Informatics,
University of Edinburgh,
United Kingdom
n.uzir@ed.ac.uk

Jelena Jovanović
Faculty of Organizational Sciences,
University of Belgrade
Serbia
jeljov@fon.rs

Abelardo Pardo
Information Technology Engineering and the Environment Divisional Office, University of South Australia,
Australia
abelardo.pardo@unisa.edu.au

ABSTRACT

Learning analytics has the potential to detect and explain characteristics of learning strategies through analysis of trace data and communicate the findings via feedback. However, the role of learning analytics-based feedback in selection and regulation of learning strategies is still insufficiently explored and understood. This research aims to examine the sequential and temporal characteristics of learning strategies and investigate their association with feedback. Three years of trace data were collected from online pre-class activities of a flipped classroom, where different types of feedback were employed in each year. Clustering, sequence mining, and process mining were used to detect and interpret learning tactics and strategies. Inferential statistics were used to examine the association of feedback with the learning performance and the detected learning strategies. The results suggest a positive association between the personalised feedback and the effective strategies.

CCS CONCEPTS

• **Applied computing**~Computer-assisted instruction

KEYWORDS

Learning Analytics, Learning Strategies, Learning Tactics, Data Mining, Self-regulated Learning, Feedback

LAK19, March, 2019, Tempe, AZ, USA
© 2019 Association for Computing Machinery.
ACM 978-1-4503-6256-6/19/03…$15.00
https://doi.org/10.1145/3303772.3303787

ACM Reference Format:

W. Matcha, D. Gašević, N. Ahmad Uzir, J. Jovanović and A. Pardo. 2019 Analytics of Learning Strategies: Associations with Academic Performance and Feedback. In *The 9th International Learning Analytics and Knowledge Conference (LAK19), March, 2019, Tempe, AZ, USA.* ACM, New York, NY, USA. 10 pages. https://doi.org/10.1145/3303772.3303787

1 INTRODUCTION

Flipped classroom is an approach that promotes active learning and consists of two main parts: pre-class preparation and in-class activities. Introduction to the relevant concepts and other passive learning activities are typically carried out online, during the preparation. The face to face time focuses on the crucial aspects of learning such as developing deep understanding, rectifying misunderstandings and confusion [48], and promoting problem-solving and critical thinking skills [19, 42, 43].

Pre-class preparation sessions may serve as an indicator of the students' overall learning performance [42, 48]. Lack of engagement with preparation activities may result in inattentiveness during in-class activities and poor course performance, even failing the course [48]. Therefore, knowing how students interact with pre-class learning activities could help teachers to provide the necessary assistance, and address any difficulties the students might have faced during the preparation sessions. However, research into how students prepare for face-to-face sessions in a flipped classroom is under-explored [48], particularly, student's ability to self-regulated their learning during the preparation.

The design of preparation activities commonly aims for individual self-managed learning, requiring from students the ability to regulate their learning productively [33]. Self-regulation involves selection of effective learning strategies and knowing when and how to apply them [53]. However, research has revealed that students often lack the ability to choose and adapt their strategies to the course requirements [34]. This implies that students may need guidance in effective strategy use. Feedback may influence the selection of learning strategies [27]. Nonetheless, research into the impact of feedback towards the selection of learning strategies in flipped classrooms is hardly present. The study presented in this paper aimed at examining strategies that students adopt to complete pre-class preparation activities. It also examined how academic performance and feedback were associated with the students' pre-class learning strategies.

2 BACKGROUND

Learning strategies and tactics are often used interchangeably in the literature, however, these two terms are different [9, 37]. A learning tactic is a learning technique or a cognitive operation that is used by a student to perform a particular task [9, 36]. Learning strategies are defined as *"any thoughts, behaviours, beliefs or emotions that facilitate the acquisition, understanding, or later transfer of new knowledge and skills"* [15, p 727]. Students adopt a certain strategy to accomplish their learning target. Learning strategies change and develop according to a learning situation, and involve selecting, combining, coordinating and utilising cognitive operations and techniques (i.e., learning tactics), directed by the learning goal [53].

Based on the model of self-regulated learning developed by Winne & Hadwin [53], the selection of a learning tactic or strategy is driven by internal (e.g., cognitive) and external (e.g. task-related) conditions. In particular, acting as agents, learners select tactics and strategies based on several factors including their knowledge of task, tactics, and strategies, the available learning materials, instructional cues, as well as their motivation, beliefs, and goals [52]. By examining the products of their learning and how they stand with respect to the expectations or learning goals, students reflect on their performance and the effectiveness of their tactics and strategies. As a consequence, the conditions are updated, and the selection of strategies and tactics could be changed too [53].

Even though research has highlighted the effectiveness of one strategy over another [2], this does not imply that less effective learning strategies are not required in the learning process [15]. For instance, re-reading and highlighting are useful to gain a basic understanding of a learning concept before students can take practice tests, which are, in turn, useful for mastering specific skills [51]. Learning is designed to involve different learning activities. Students need to adopt a variety of tactics and strategies to accomplish their learning goal. Therefore, knowing when and how to choose or adjust a learning strategy to suit the given learning situation is crucial. For example, Lust et al. [34] stress that students have to adjust the use of learning tools in accordance with the learning phases of the course. However, research has reported that students often have underdeveloped skills to regulate and modify their learning strategies to meet the course requirements [18, 34]. Though clearly important, such temporal changes in learning strategy have not been sufficiently explored [34, 40].

2.1 Detection of Tactics and Strategies

Traditionally, learning tactics and strategies have been detected by using self-reporting instruments such as surveys and think-aloud protocols. However, learners are not always accurate in reporting how they learn [52]. For instance, by comparing actual learning activities to the self-reports, Hadwin et al. [26] found that self-reports did not reflect the students' actual behavior. Think-aloud protocols can introduce cognitive overload as students are required to elaborate and justify their actions out loud while learning [52]. Moreover, self-reports usually fail to capture how strategies develop over time. The use of learning trace data allows for understanding the students' actual learning behaviour without intervening in their learning or inadvertently increasing their cognitive overload. Zhou and Winne [59] asserted that trace data were better correlated to the students' learning achievement than self-reports. Still, self-reports are successful in capturing students' perceptions and intentions and could assist in understanding the choices of actions the students make.

The use of trace data allows for examining the dynamics, that is, temporal dimension of learning tactics and strategies [55]. In particular, learning tactics and strategies can be considered as sequences of events, with each event being centered on one learning action [26, 40]. Such events are dynamic, shift from one action to another, and, develop over time [54]. A sequence of learning actions performed to complete a learning task is recognised as a tactic [9]. Scrutinising learning tactics at the session level could provide a closer observation on how students engage in the learning process [18]. Learning strategies can be identified by recognizing regularity in students' learning paths, that is, sequences of learning tactics. To sum up, understanding the temporal and sequential dimensions of learning events could shed some light on how tactics and strategies have developed and changed, and allow for detecting situation where transient state changes happen [40]. Such patterns of events evolve over time to become a characteristic of one's learning and may be considered as aptitudes that could predict one's future behaviour [54].

Research into detection of learning strategies by using trace data has gradually increased. For example, Maldonado-Mahauad et al. [35] extracted learner regulation strategies by using self-reports and applied process mining to examine the regulation process of each strategy group in three massive open online courses. They found that *comprehensive* learners, who followed the course structure step by step, and *targeting* learners, who focused on specific learning activities to help them pass the course, showed higher learning performance as compared to the *sampling* learners, who showed low engagement and inconsistency. Jovanovic et al. [28] used sequence mining and agglomerative hierarchical clustering to detect learning strategies based on the students' sequences of actions. Fincham et al. [18] extracted learning tactics from trace data by using Hidden Markov Models (HMM) and identified learning strategies by examining sequences of the detected tactics based on the first and second half of the semester. The detected tactics and strategies were explained based on the distribution of learning actions. They found that students who employed multiple learning tactics and were highly engaged in learning activities tended to perform better. This study is different from the work of [18] as it examined tactics and strategies based on their temporal and sequential characteristics of the learning sequences. Moreover, this study aimed to examine the association of the feedback, performance and the selection of the strategy at the course level, rather than dividing the course into the first and second half of the semester.

The studies presented thus far clearly indicate that learning strategies can be derived from trace data and can be interpreted according to the research on self-regulated learning and learning strategies [2, 53]. However, the distinction between tactics and strategies have rarely been made [18], while studies of how tactics articulate in a strategy are lacking. As such the following research question has been defined:

RQ1: *Given a sequence of actions across several time frames, can we detect theoretically meaningful learning tactics and strategies applied by students when preparing for face-to-face sessions in a flipped classroom?*

2.2 Learning Strategy and Learning Outcomes

Learning strategies have impact on learning performance [51, 57]. Research findings denote that not all students use effective learning strategies [12, 37]. Particularly, low and high performing students tend to apply different learning approaches [10, 47]. For instance, Nandagopal and Ericsson [41] found that upper-level college students were regulating their learning strategies differently. Students with higher performance engaged more in applying a variety of learning strategies and had a higher tendency to review lessons. DiFrancesca et al. [10] reported that low performing students perceived less effective learning strategies as more important than effective learning strategies, and relied more on less effective strategies during their learning process. Being based on an instructional approach that is substantially different than those of the traditional classroom, flipped classroom could impose further challenges for regulation of learning and selection of effective strategies [18]. Therefore, how students regulate their learning strategies and how that regulation is associated with their learning outcomes in a flipped classroom might differ from the research findings obtained in traditional classrooms. Therefore, our second research question is formulated as follows:

RQ2: *Is there an association between the identified learning strategies and the students' academic performance in a flipped classroom?*

2.3 Feedback and Learning Strategies

Feedback helps students to reflect on their studies, clarify misunderstanding, reflect on the use of tactics and strategies, and enhance their self-regulation [30]. Even though much research reports positive effect of feedback, not every feedback is equally effective [30, 31]. For example, internal feedback, i.e. feedback generated when students evaluate their learning products against the standards and/or learning goals, is often inaccurate. For instance, Bjork et al. [2] found that students were deceptive in assessing their learning, which could lead to the selection of ineffective learning approaches. Teachers can intervene by providing external feedback. However, the provision of external feedback has become more challenging due to the large cohorts of students per small number of teachers [44] and the time-consuming task of generating feedback for each student [46]. Furthermore, considering the importance of applying potent learning strategies and the necessity of guiding students in the selection of learning strategy [52], feedback targeted at guiding students to regulate their learning is considered the most useful [27]. However, students often do not receive feedback on how to select, modify and articulate learning strategies to suit the given learning situation and course design [6]. Learning analytics has a potential to support the administration of timely, accurate, strategy-focused feedback [22]. For example, rich information produced by learning analytics may serve as the bases for the provision of personalised feedback.

Besides the challenge in feedback provision, research also highlights some of the barriers preventing students from turning feedback into action, such as the lack of ability to interpret the meaning and decode the required actions from the feedback received [7, 56]. Students may also fail to recognize the benefits of feedback, or have difficulty in relating the feedback to the current learning situation [16, 46, 56]. Even though learning analytics-based feedback shows promising results, its impact (i.e. how students change their behaviour after receiving the feedback) is under-explored [3, 49]. Considering all the above, we have formulated our third research question as follows:

RQ3: *Are feedback interventions associated with the students' choice of learning strategies and their performance in a flipped classroom?*

3 METHODS

3.1 Study Context

The data for the study were collected from class preparation activities of a flipped classroom deployed in a first year Computer Engineering course at an Australian university. The study examined data from three successive course editions, in years 2014, 2015, and 2016. The number of enrolled students steadily increased over the three years ($N_{2014} = 290$, $N_{2015} = 368$, and $N_{2016} = 477$). In all years the course was scheduled for 13 weeks with ten studied topics.

Students were required to complete online pre-class learning activities and attend face-to-face classes. This study was focused on the online preparation activities, which were crucial for the success of the overall flipped classroom design [48]. The following preparation activities were available to students:

1. *Videos with multiple-choice questions (MCQs)*: students were provided with short videos explaining concepts relevant for the weekly topic. Each video was followed by a set of MCQs that served as formative assessment and were aimed at recalling the knowledge explained in the video. Prompt evaluation, multiple attempts, and solutions were provided.
2. *Reading materials with MCQs*: these MCQs also served as formative assessment and were aimed at provoking a simple recall of the concepts introduced in the reading materials.
3. *Problem-solving activities (exercises)*: these activities were part of summative assessment as the obtained scores contributed to the final course mark, provided that the exercises were completed prior to the week's lecture. This requirement was introduced to ensure that students prepare for the in-class activities.

Aside from learning activities, students were required to do a project. They were also provided with feedback. Two types of feedback were introduced over the three years: personal dashboards and analytics-based feedback in the form of personalised emails. The latter will be referred to as *personalised feedback* in the reminder of the paper.

The learning analytics dashboard used gauges to represent the level of individual learning performance and engagement with the pre-class preparation activities [29]. Students could compare their scores against the class average scores to reflect on their engagement and performance. The dashboard was provided across the three years.

Personalised feedback was in the form of textual messages that students received by email. These were semi-automatically generated based on a specific set of rules set by the instructor to create feedback for individual students at the learning process and regulation levels. Personalised feedback messages were customised based on an algorithm that compared a student's level of engagement and performance with several predefined ranges of performance and engagement (i.e. ranges set by the instructor). This way, each student received a weekly elaborated message that addressed every pre-class learning activity. Students received an advice on how to guide their learning process, focusing on the actions and some extra learning materials. The following examples illustrate the feedback messages the students received: *"make sure you review again the whole content explained in the video of the activity 2.3.1. You could use a piece of paper and try to replicate the developments that are explained in the video."* or *"Good initial work with the video in the activity. You may try creating your circuits only with interconnected flip-flops and try to derive their behaviour"*. Further details about this type of feedback are given in [44]. Personalised feedback was introduced in 2015 (2nd year of data collection) and was sent out from week 2 until the mid-term examination (week 6). In 2016, students received personalised feedback weekly throughout the entire semester.

The examinations were arranged twice and included mid-term and final exams. The scores obtained from the problem-solving activities during the weekly preparation were accumulated into the final course grade.

3.2 Data

The collected trace data consisted of a series of events where each event included session ID, student ID (anonymised), type of learning action (Table 1), topic, and timestamp.

Table 1: Types of learning actions extracted from the data

Action Code	Description
EXE_CO	A correctly solved exercise item
EXE_IN	An incorrectly solved exercise item
MCQ_CO	A correctly solved MCQ item
MCQ_IN	An incorrectly solved MCQ item
MCQ_SR	A solution requested for an MCQ item
VIDEO_LOAD	Loading the video
VIDEO_PL	Play the course video
VIDEO_PA	Pause the video
VIDEO_END	End of the video
CONTENT_ACCESS	Access to a page that contains reading materials
INDEX_ACCESS	Access to the index page
PROJECT_ACCESS	Access to project information page
MC_EVAL	Access to the dashboard; this is considered a metacognitive evaluation action
MC_ORIENT	Access to the schedule and the learning objective pages; this is considered a metacognitive orientation action
SEARCH	Search for keyword
DOWNLOAD	Download documents or files

Learning sessions were extracted from trace data and considered learning sequences consisting of consecutive learning actions. Any two events were within 21min of one another within each sequence. Currently there is no widely accepted approach to estimation of time on task [35]. This time of 21min was selected by analysis the duration of videos. This analysis indicated that the 96th percentile of 17.5min was too short, whereas the 98th percentile of 41.2min was too long. The sequences varied, both in terms of length and composition of learning actions. To normalise the data, the outliers, i.e. overly short sequences (consisting of one action) and overly long sequences (above the 95th percentile of the number of learning actions) were trimmed off following [28]. As a result, 65,710 learning sequences were generated from the dataset. The length of learning sequences ranged from 2 to 175 actions.

3.3 Data Analysis Techniques

3.3.1 Detection of Students' Learning Tactics and Strategies. A learning tactic can be considered as a sequence of actions that a student performs to complete the specified task [18, 26]. To automatically detect learning tactics from learning sequences, we began by inspecting the learning process through process mining lenses. Process mining generates a process model based on a set of timestamped actions. By observing the overall process model, the potential number of learning tactics can be inferred based on the density of connections among actions. Process mining was performed by using first order Markov models (FOMMs) and the pMineR R package [25]. The output of a FOMM is an adjacency matrix of transition probabilities between events shown in Table 1. This output is suitable for cluster analysis using the Expectation – Maximization (EM) algorithm [17]. Thus, EM was used to cluster learning sequence to detect meaningful learning tactics.

According to Derry [9], a learning strategy employs one or more tactics. Therefore, learning strategies can be inferred from the way individuals employed tactics. Agglomerative Hierarchical Clustering based on the Ward's algorithm [20] was used to extract patterns of how individual students used the identified learning tactics. As the input for the clustering process, for each student we used the number of each tactic and the total number of all tactics. The Euclidian metric was used to calculate the (dis)similarity between vectors of the input. The dendrogram was used to determine the optimal number of clusters. This process of the use of hierarchical clustering has already been established for detection of learning strategies [28, 32].

3.3.2 Exploring the Detected Learning Tactics and Strategies. Behavioural patterns detected in the trace data require content-based interpretation to provide actionable insights [35, 43]. To understand the characteristics of the detected clusters (i.e. tactics and strategies), sequence mining and process mining were used. The TraMineR R package allows for constructing and examining sequences of actions [20]. It can be used to explore the frequency, the ordering, and the distribution of actions within sequences, and to explore clusters of sequences.

The pMineR R package [25]was also used to explore temporal characteristics of the learning strategy groups. Specifically, we analyzed changes of learning tactics for each of the strategy groups within each weeks of the course. We decided to use week as the unit of analysis given that activity cycles in the course were weekly as common for higher education [43].

Descriptive and inferential statistics were used to characterise the identified clusters (i.e., tactics and strategies). Kruskal Wallis tests were used to examine the association between learning strategies and the performance.

3.3.3 Association with the Feedback. Different types of feedback were provided each year (Sect. 3.1). Therefore, the effects of feedback can be observed from the changes in learning tactics, strategies or performance of students across the years. To observe these changes, the proportions of the students' use of different tactics and strategies were examined. Proportions were used instead of the counts data due to the increasing number of students enrolled in the course in each year. Observing changes in the proportions could help us to reflect on the adoption of the learning tactics and strategies.

4 RESULTS

4.1 RQ1: Detection of Learning Tactics and Strategies

4.1.1 Detection of Learning Tactics from Trace Data. Table 2 presents the characteristics of the five detected tactics in term of the state distribution, frequency, and arrangement of the sequences. The supplementary document can be found online[1].

Tactic 1 – Diverse (N = 8,288, 12.61% of all sequences) contained sequences with the highest number of learning actions (median = 53 actions per sequence). Students exhibited a variety of actions, with a relatively even distribution of exercises, MCQs, and video views.

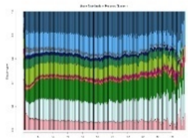
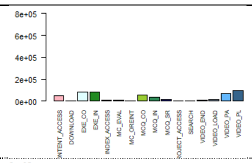
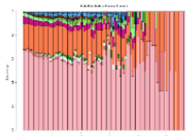
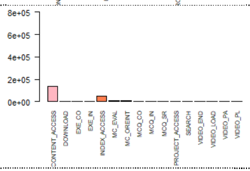
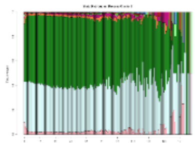
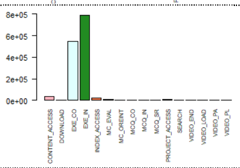
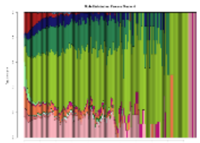
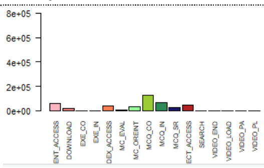
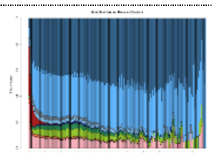
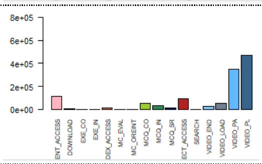
Tactic 2 – Reading oriented (N = 17,024, 25.91% of all sequences) contained the shortest sequences of learning actions (median = 4 actions in a sequences). The dominant kind of action was access to the reading materials.

Tactic 3 – Exercise oriented (N = 16,287, 24.79% of all sequences) contained a moderate number of learning actions (median = 24 actions per sequence). The most dominant learning actions were problem-solving actions (EXE_CO and EXE_IN). Unlike other tactics, most of the learning sequence in this cluster began by direct access to the problem-solving activities rather than access to the reading materials.

Tactic 4 – MCQ oriented (N = 11,915, 18.13% of all sequences) contained relatively short learning sessions (median = 5 actions per session) that often began by accessing reading materials (CONTENT_ACCESS), followed by MCQ answering. MCQ related actions (MCQ_CO, MCQ_IN, and MCQ_SR) were the most dominant type of action.

Tactic 5 – Video oriented (N = 12196, 18.56% of all sequences) consisted of relatively short sessions (median = 9 actions per session). Based on the sequence length and dominant type of action, two types of learning sequences could be distinguished (Table 2). Long sessions often comprised of content access followed by video playing/pausing actions, which were in turn followed by MCQ-related actions. Shorter sessions consisted mainly of access to the project information pages.

Table 2: Characteristics of the detected tactics

Tactics	State Distribution	Frequency of Action
1		
2		
3		
4		
5		

CONTENT_ACCESS, DOWNLOAD, EXE_CO, EXE_IN, INDEX_ACCESS, MC_EVAL, MC_OREINT, MCQ_CO, MCQ_IN, MCQ_SR, PROJECT_ACCESS, SEARCH, VIDEO_END, VIDEO_LOAD, VIDEO_PA, VIDEO_PL

Overall, metacognitive actions, which consisted of access to the dashboard (MC_EVAL) and course syllabus (MC_ORIENT), were noticeable in every tactic but showed relatively low presence compared to other types of actions (Table 2).

4.1.2 Detection of Learning Strategies from the Students' Choice of Tactics. The dendrogram produced by the employed clustering algorithm suggested three strategies as the best solution. Table 3 presents the summary statistics for the variables that served as the input for the clustering: the number of times each of the learning tactics was used as well as the overall number of learning tactics per student. To better understand the detected learning strategies, we have also examined, for each strategy, how the use of learning tactics changed throughout the course. Fig. 1 illustrates, for each detected strategy group, median number of tactics applied in each week of the course.

[1]https://drive.google.com/open?id=1ltIp-HlCu9hJiq-GJknAFIaPNJ1q9poJg3a6nIif0FI

Strategy Group 1: Strategic – Moderate engagement: This was the largest cluster (N = 519, 45.73%). Students in this strategy group tended to use different learning tactics in different weeks of the course. Only exercise-oriented tactic was consistently used throughout the semester. In the first half of the semester (week 2-week 6), in addition to exercises, students also focused on reading materials and the associated MCQs. In the second half of the semester (weeks 7-13), exercise-oriented tactic was combined with video- and reading-oriented tactics.

Strategy Group 2: Highly selective – low engagement: The proportion of students in this strategy group was relatively high (N=418, 36.83%). Students in this group had low engagement with the preparation activities (Table 3). They chose to focus on specific types of learning tactics, namely, exercises and reading oriented. The exercise-oriented tactic was used throughout the semester, whereas the reading-oriented tactic was present only up until the midterm exam (week 6).

Strategy Group 3: Intensive – high engagement: This was the smallest group (N = 198, 17.44%), comprised of students with the highest engagement level. They applied a variety of learning tactics. Reading, video, MCQ and exercise-oriented tactics were used each week throughout the semester. Diverse tactic was observed on certain weeks (week 6 and 10).

Table 3: Summary statistics (median, 1st and 3rd quartile) for variables used for detecting learning strategies

Tactics	Strategy 1	Strategy 2	Strategy 3
1	8.0 (5.0-11.0)	3.0 (1.0-5.0)	13.0 (8.0-17.0)
2	15.0 (11.0-22.5)	7.0 (4.0-11.0)	21.5 (14.0 30.0)
3	15.0 (12.0-18.0)	12.0 (9.0-15.0)	17.0 (13.0-20.0)
4	10.0 (7.0-15.0)	5.0 (2.0-7.0)	16.5 (10.0-27.0)
5	11.0 (8.0-14.0)	4.0 (2.0-6.0)	23.0 (19.0-27.0)
Total	62.0 (53.0-72.0)	33.0 (25.0-41.0)	94.5 (77.2-111.0)

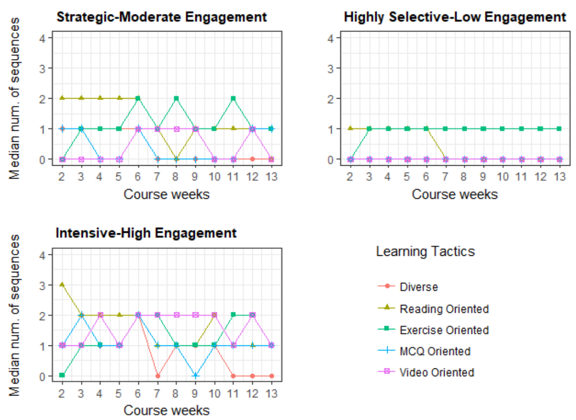


Figure 1: Weekly changes in the applied learning tactics for each of the three detected learning strategies group.

To further our understanding of the detected learning strategies, first order Markov models were fitted to explore transitions from one learning tactic to another on within each learning strategy. Fig. 2 presents the process models of each learning strategy based on transition of tactics.

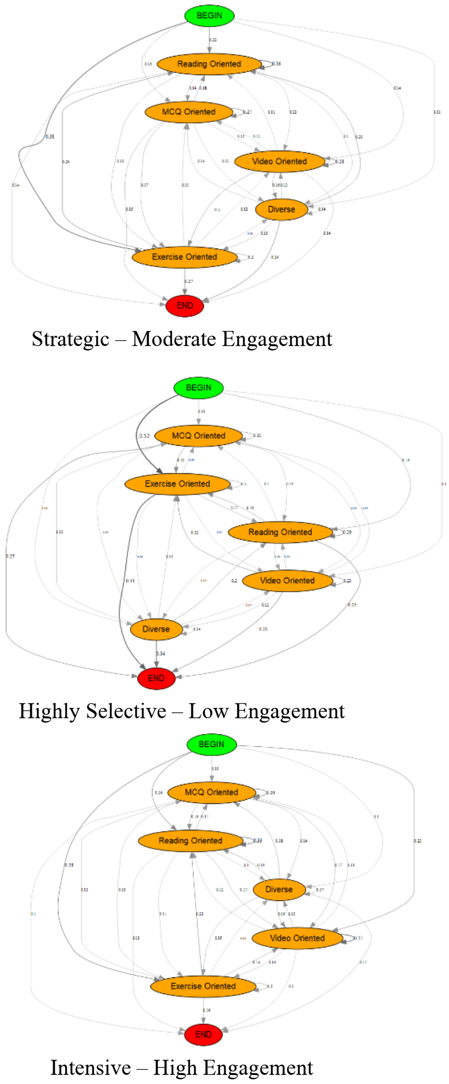


Figure 2: The process models of the detected strategies.

The main focus of the *Strategic – moderate engagement* strategy was on the exercises. There was a high probability (p=0.38) for students in this group to have weeks based on exercise-oriented tactic only. When it comes to changing learning tactics from one session to another one within a week, the most notable were the shift from exercise-oriented to reading-oriented tactics (p=0.24) and the shift from reading-oriented to diverse tactics (p=0.21). The weeks often ended by using the exercise oriented and diverse oriented tactics.

The main characteristics of the *Highly selective – low engagement* strategy was its concentration on the exercise-oriented tactic. There is a strong probability that students in this strategy group began and ended their weeks by doing exercises (p= 0.52, p= 0.34, respectively).

Students who adopted the *Intensive – high engagement* strategy tended to begin their study weeks with a variety of learning tactics, rather than relying on one specific tactic. More precisely, the beginning of the sessions was almost equally distributed across reading-oriented (p=0.24), exercise-oriented (p=0.28), and video-oriented tactics (p=0.23). Transition between learning tactics were clearly observable and roughly equally distributed (probabilities ranged from 0.15 to 0.23). This indicates that students in the *Intensive – high engagement* group used a variety of learning tactics, which can be interpreted as an indication of their ability to self-regulate their learning [52].

4.2 RQ2: Associations between Learning Strategies and Academic Performance

The scores for all learning strategy groups are presented in Table 4. Kruskal Wallis tests showed a significant association between student learning strategy groups and the students' course performance (p-value < 0.0001) for both midterm and final exam scores. To further examine the associations between the detected learning strategies and the students' academic performance, we did pairwise comparisons of learning strategy groups with respect to the midterm and final exam scores.

Table 4: Summary statistics (median, 1st and 3rd quartile) for midterm and final exams for each strategy group

Strategy Group	Strategic – moderate engagement	Highly selective – low engagement	Intensive – high engagement
Midterm	15 (12 – 17)	12 (10 – 15)	15 (12 – 17)
Final Exam	21 (14 – 30)	15 (10 – 20)	22 (16 – 32)

As shown in Table 5 and 6, significant differences with respect to both midterm and final exam scores were present between *Strategy 1: strategic – moderate engagement* and *Strategy 2: highly selective – low engagement* and between *Strategy 3: Intensive – high engagement* and *Strategy 2: highly selective – low engagement.* However, no statistically significant difference between *Strategy 1: strategic – moderate engagement* and *Strategy 3: Intensive – high engagement with respect with midterm and final exam score*.

Table 5: Pairwise comparison of strategy groups with respect to the midterm scores

Strategy	Strategy	Z	p	r
1	2	9.329	< 0.0001	0.31
1	3	-0.1577	0.8749	0.01
2	3	-7.0258	<0.0001	0.28

Table 6: Pairwise comparison of strategy groups with respect to the final exam scores

Strategy	Strategy	Z	p	r
1	2	9.338	< 0.0001	0.311
1	3	-1.8512	0.06412	0.077
2	3	-8.8471	<0.0001	0.366

4.3 RQ3: Association with the Feedback

To explore the association between the students' choice of learning tactics and the feedback interventions over the three years of the course (2014-2016), the trace data of each year were first explored separately and then compared as feedback was incrementally introduced (see Section 3.1). Fig. 3 shows the proportions of learning tactics applied by students in each year.

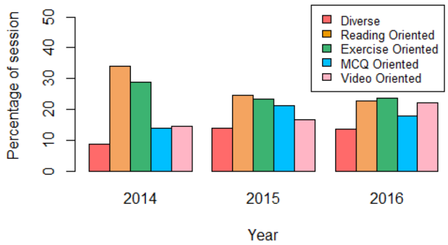


Figure 3: The proportion of learning tactics in each year.

In 2014, the students mostly focused on the use of the reading and exercise-oriented tactics. About 15 percent of learning sessions were focused on the use of the MCQ and video-oriented tactics. In 2015 and 2016, when the personalised feedback was introduced, the use of video-oriented tactic increased, whereas the proportion of reading and exercise-oriented tactics decreased. Overall, Fig. 3 indicates a much more balanced use of learning tactics in 2015 and 2016, when the personalised feedback was provided. In spite of the observed change, a Chi-squared test applied to investigate the association between the course year and the choice of learning tactics found no statistically significant association; that is, no significant association was found between the feedback interventions and the students' use of learning tactics (p-value = 0.4286).

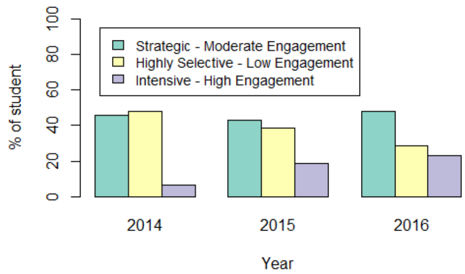


Figure 4: The proportion of learning strategies across years.

Examining the proportions of students in each strategy group across the three consecutive course offerings (Fig. 4), we found that in the years when personalised feedback was administered (in 2015 and 2016), the proportion of students in the *Intensive-high engagement* strategy group (violet bar, Fig. 4) steadily increased, whereas, the proportion of students adopting the *Highly selective – low engagement strategy* (yellow bar) steadily decreased over time. This observation was further confirmed by the Chi-squared test that revealed a significant association between the course year and student strategy group ($p < 0.005$).

5 DISCUSSION

5.1 RQ1: Tactics and Strategies Detection

The identified clusters are well distinguished. Each cluster reflects different dominant actions and the composition of actions in learning sequences varied (Tables 2 and 3, Fig. 1). As such, these clusters are representative of the patterns in students' learning behaviour and indicative of the learning tactics and strategies that students used.

The detected learning strategies reflect the well-defined approaches to learn as described by Entwistle [14], Marton & Säljö, [38], and Biggs [1]. Approach to learning are situation, content, and intuition dependent. In other words, the goal and motivation of individual students (intuition), the way in which learning was carried out (the situation), and, the content that students needed to learn (the content) play highly important roles in the selection of the approach to learn [13]. This notion is well-aligned with self-regulated learning theories, which emphasize influences of task and cognitive conditions on the selection of learning strategies [53]. Entwistle [14] and Biggs [1] defined three learning approaches, extended from the initial concept of approaches to learning proposed by Marton & Säljö, [38], namely, surface, deep, and strategic learning.

Deep learning is characterised by active engagement in various learning activities. This learning approach is considered a desirable learning approach [39] since a number of research studies has reported a positive impact of deep learning [5]. In our study, students who used learning strategy 3 (*intensive – high engagement*) reflect the use of this learning approach. The highest level of engagement of this strategy group reflects an active concentration and high amount of effort the students exerted. Moreover, they used varied learning tactics, concentrating on various learning activities. The performance of students who applied the *Intensive – high engagement* strategy tended to be the highest among the three groups, which is consistent with previous research findings that the deep learning approach is associated with a high level of performance [39, 58].

Surface learning is considered a superficial method of learning where the concentration is mainly on the assessment, with the lack of understanding of the content [14]. The students who used *Highly selective – low engagement* strategy are good representatives of this learning approach. The low level of engagement demonstrates the lack of intention and effort towards accomplishing higher order of thinking. The assessment tactic is the only consistent tactic of these. Surface learning is also characterised by focusing on a portion of learning facts and often jumping to conclusions as demonstrated by the use of a single learning tactic (reading oriented) and directly jumping to the assessment activities [5, 28]. Students in this group did not fully utilise the learning activities that were offered to them.

Strategic learning is also referred to as achieving learning approach [1]. It is the study approach of those who intend to achieve high performance [11]. Similar to surface learning, focus is on the assessment, but a considerable amount of effort is put into understand the topic of study. Therefore, strategic learners combine surface and deep learning approaches [5], and they do well in time and study management. As demonstrated in the detected learning strategy group 1 (*strategic – moderate engagement*), students in this group put in moderate effort as compared to the other two groups, yet, they achieved high performance level. They were consistently concentrated on the assessment activities, as evident in the extensive use of the exercise-oriented tactic.

5.2 RQ2: Association between Learning Strategies and Academic Performance

This study found that students who used a variety of learning tactics tended to have higher performance than those who used a single tactic. This finding implies that higher performers use many different learning tactics, some are general, passive, and less effective such as re-reading [12] or re-watching lecture videos, but some are specific and more effective such as practical testing as demonstrated by the frequent use of exercise-oriented tactic and MCQ oriented tactic. Similarly, [18, 23, 41] found that the students who engaged in various learning tactics had a tendency to perform better. Good learners know when, where, and how to apply learning tactics and strategies [45].

Referring again to the students' approaches to learning, the deep and strategic learning approaches are the desired learning approaches, associated with higher performance [5, 39]. Students who apply these two learning approaches tend to achieve high performance whereas the surface learning approach is less desirable and associated with a low performance. Consistent with the literature, this research found that students who used the *Highly selective – low engagement* strategy, thus reflecting the surface learning approach, performed poorer than those who used one of the other two learning strategies [4]. Meanwhile, this study shows no significant association in term of pairwise comparison between the *Strategic – moderate engagement* strategy and the *Intensive – high engagement* strategy. This corresponds to the use of strategic learning and deep learning approaches, respectively. In the first half of the semester, high level of engagement was evident among the students of these two groups. Considering that these two learning approaches are known to be associated with the high academic performance[4, 58], it was not surprising that we found no difference in academic performance of these two groups. The amount of effort exerted by the *strategic – moderate engagement* group declined after the midterm test. Prior research has denoted the role of engagement as one of the success factors in learning [8]. Hence, our finding that the *Intensive – high engagement* strategy group performed better on the final exam than the *Strategic – moderate engagement* group corroborates the findings of several empirical studies that demonstrated that higher level of engagement shows significant impact on the course achievement. Effective learners are those who did not only choose the effective learning strategies but have also realised that learning requires effort [45, 52].

5.3 RQ3: Association with the Feedback

Feedback could have positive or negative effects [27]. In general, we found a positive association between the personalised feedback with elaborated learning advice and the application of learning strategies. In particular, the number of students who applied the *Intensive – high engagement* strategy had progressively increased from 2014 (no personalised feedback intervention) to 2016 (full implementation of the personalised feedback intervention). Moreover, from 2014 to 2016, there was a gradual decrease in the number of students who adopted the *Highly selective – low engagement strategy*. This is consistent with the literature, which states that effective feedback promotes higher level of engagement and quality of study [56]. Similarly, [16] explored the association with the feedback by dividing the course structure into two halves, separated by the midterm exam, and found a small significant association of the feedback in the choice of learning strategies.

A student is an active agent [52]. Students decide how to respond to external feedback by considering the usefulness of the feedback [46], the difficulty in decoding and interpreting the feedback, and turning it into action [56]. Students' motivation to engage with feedback is also found to be relevant [16, 46]. Further research needs to explore the features of feedback that can influence the adoption of a learning strategy.

Research on the impact of learning analytics-based feedback on behavioural changes is scarce. Most of the existing research has examined the usefulness of the feedback by using a survey and examining the association with the learning performance [3]. This research has contributed to closing this gap by demonstrating the association between the feedback and the increased use of effective learning strategies as well as reduced adoption of less effective strategies.

The study is also significant as it showed the value of analytics-based feedback in the form of personalized textual messages over the use of personal learning analytics dashboards. Although learning analytics dashboards has been challenged to produce either negative or inconclusive effects on different learning outcomes [24], further research needs to investigate if analytics-based personalised feedback messages would be associated to similar effects in cases where no learning dashboards are not used at all (n.b., in this study, personalised feedback messages were provided in addition to a learning analytics dashboard).

6 CONCLUSION AND IMPLICATIONS

This research provides some novel insight into the field. It showed promising results for tactic and strategy detection from trace data and their associations with academic peformance and feedback. However, as any study, it has some limitations. First, the study was based on the data about online preparation activities in a flipped classroom. Students' engagement in face-to-face sessions was not analyzed. Even though preparation activities play an important role in the development of conceptual understanding and aids the in-class activities, in-class activities contribute to deepening of students' understanding and consequently affect their course performance. Furthermore, the students' demographics and previous knowledge were not considered due to the terms and conditions imposed by the institutional ethics approval.

The methods used to detect learning tactics and strategies (EM and hierarchical clustering) belong to the group of the unsupervised machine learning methods, which, by their nature, introduce a certain level of subjectivity. The explanation of the identified tactics and strategies is subjected to the theory as to how tactics and strategies are composed [35, 43]. The use of self-reports [59] or multimodal techniques to capture the students' motivation and goals could help us understand the study results better.

The role of feedback requires further exploration. For example, identifying the extent to which students modify their learning according to the feedback provided. This research could not answer if students modified their learning strategies and tactics according to the feedback provided in a causal manner – only associations were established. Exploring this research gap could contribute to the understanding of effective approaches to provide feedback to students; the resulting guidelines could be put into action.

In spite of the stated limitations, our findings are consistent with the literature. The detected learning tactics and strategies are meaningful in the learning context and in accordance with learning theories [1, 14], which demonstrates that the findings are feasible for future research. In particular, the methodology used has some important implications to research practice.

As latent constructs, learning strategies cannot be directly observed in the collected learning traces, but have to be extracted using appropriate statistical or machine learning algorithms. We have detected tactics at the level of learning sessions, extracted learning strategies from patterns of tactics used, and interpreted the meaning of the detected learning tactics and strategies by considering their temporal and sequential attributes. This method proved to provide detailed observations of how student interact with preparatory learning activities in the context of a flipped classroom. Still, the method should be equally well applicable in any blended learning context, provided that trace data from the online part of the course design could be collected. On the other hand, the detected tactics and strategies are heavily context-dependent. That is, the specific learning tactics and strategies have to be interpreted in the particular learning context the data originate from [21, 22, 53]. The interpretation should consider both the chronological ordering as well as their temporal dimension as learning tactics and strategies are "dynamic constructs".

Our findings related to the association of the feedback interventions, learning strategies, and academic performance, suggest that elaborated personalised feedback is useful and pragmatic. Literature has long appealed for an approach to guide students in selecting effective learning strategies. This study demonstrates that elaborated personalised analytics-based feedback at the level of process and regulation offers a potential approach to effectively communicate with the students on the selection of the effective learning strategy. Still, further research needs to be conducted to inform a holistic framework of how to adopt personalised analytics-based feedback in flipped classroom setting.

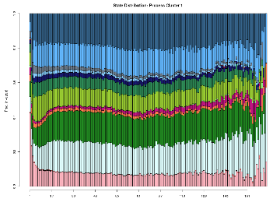
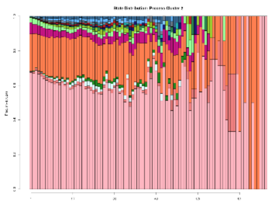
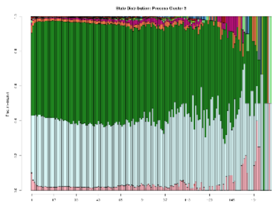
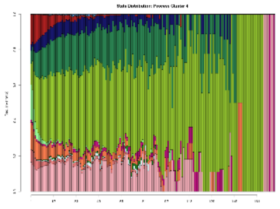
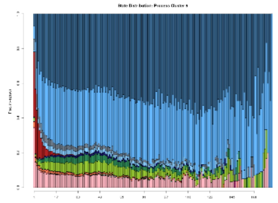
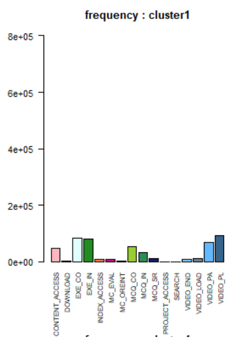
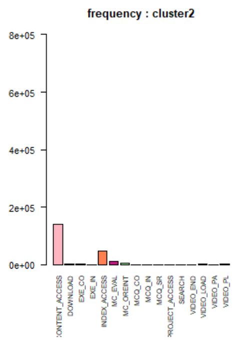
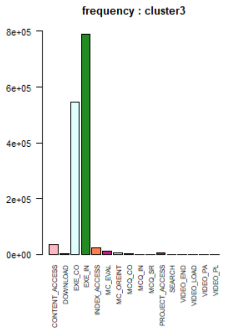
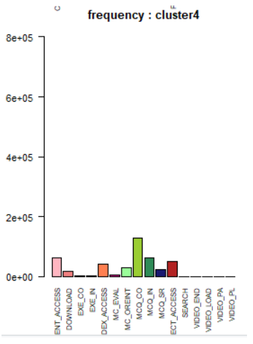
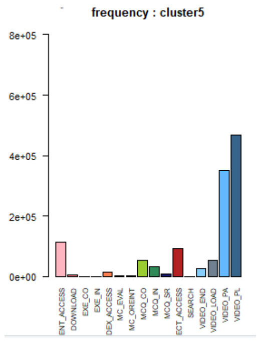
REFERENCES

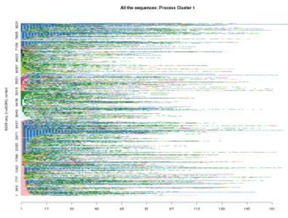
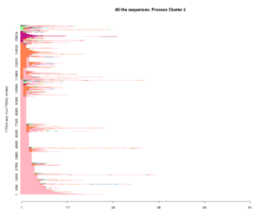
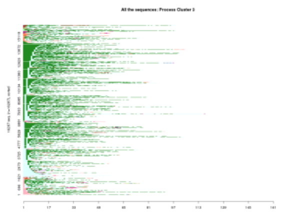
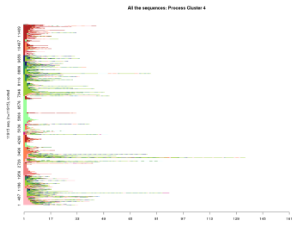
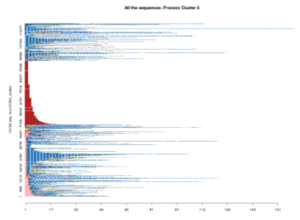
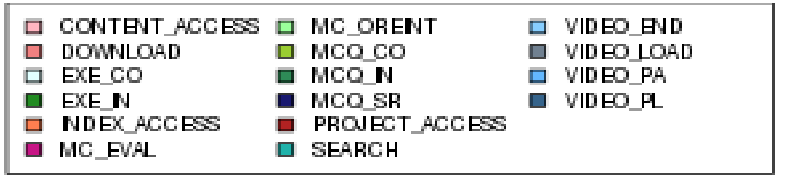
[1] Biggs 1987. Student Approaches to Learning and Studying.
[2] Bjork, R. a., Dunlosky, J. and Kornell, N. 2013. Self-Regulated Learning: Beliefs, Techniques, and Illusions. Annual Review of Psychology. 64, 1 (2013), 64, 417-444.
[3] Bodily, R. and Verbert, K. 2017. Trends and issues in student-facing learning analytics reporting systems research. Proceedings of the Seventh International Learning Analytics & Knowledge Conference on - LAK '17. 1, 212 (2017), 309–318.
[4] Byrne, M., Flood, B. and Willis, P. 2010. The relationship between learning approaches and learning outcomes : a study of Irish accounting students The relationship between learning approaches and learning outcomes : a study of Irish. 9284, (2010).
[5] Chonkar, S.P., Ha, T.C., Chu, S.S.H., Ng, A.X., Lim, M.L.S., Ee, T.X., Ng, M.J. and Tan, K.H. 2018. The predominant learning approaches of medical students. BMC Medical Education. 18, 1 (2018), 1–8.
[6] Cicchinelli, A., Veas, E., Pardo, A., Pammer-Schindler, V., Fessl, A., Barreiros, C. and Lindstädt, S. 2018. Finding traces of self-regulated learning in activity streams. Proceedings of the 8th International Conference on Learning Analytics and Knowledge - LAK '18. (2018), 191–200.

[7] Corrin, L. and de Barba, P. 2015. How Do Students Interpret Feedback Delivered via Dashboards? International Conference on Learning Analytics and Knowledge. (2015), 430–431.
[8] Dabbagh, N. 2007. The online learner: Characteristics and pedagogical implications. Contemporary Issues in Technology and Teacher Education. 7, 3 (2007), 217–226.
[9] Derry, S.J. 1989. Putting learning strategies to work. Educational Leadership. 47, 5 (1989), 4–10.
[10] DiFrancesca, D., Nietfeld, J.L. and Cao, L. 2016. A comparison of high and low achieving students on self-regulated learning variables. Learning and Individual Differences. 45, (2016), 228–236.
[11] Diseth, Å. and Martinsen, Ø. 2003. Approaches to Learning, Cognitive Style, and Motives as Predictors of Academic Achievement. Educational Psychology. 23, 2 (2003), 195–207.
[12] Dunlosky, J. 2013. Strengthening the Student Toolbox. American Educator. 37, 3 (2013), 12–21.
[13] Entwistle, N. 2007. Research into student learning and university teaching. The British Psychological Society. October (2007), 1–18.
[14] Entwistle, N.J. 1991. Approaches to Learning and Perceptions of the Learning Environment : Introduction to the Special Issue. Higher Eduation. 22, 3 (1991), 201–204.
[15] Entwistle, N.J. 2009. Teaching for understanding at university: Deep approaches and distinctive ways of thinking. Palgrave Macmillan.
[16] Evans, C. 2013. Making Sense of Assessment Feedback in Higher Education. Review of Educational Research. 83, 1 (2013), 70–120.
[17] Ferreira, D.R. and Gillblad, D. 2009. Discovering Process Models from Unlabelled Event Logs. Business Process Management. 5701, (2009), 143–158.
[18] Fincham, O.E., Gasevic, D. V., Jovanovic, J.M. and Pardo, A. 2018. From Study Tactics to Learning Strategies: An Analytical Method for Extracting Interpretable Representations. IEEE Transactions on Learning Technologies. 1382, c (2018), 1-14.
[19] Freeman, S., Eddy, S.L., McDonough, M., Smith, M.K., Okoroafor, N., Jordt, H. and Wenderoth, M.P. 2014. Active learning increases student performance in science, engineering, and mathematics. Proceedings of the National Academy of Sciences. 111, 23 (2014), 8410–8415.
[20] Gabadinho, A., Ritschard, G., Studer, M. and Muller, N.S. 2008. Mining sequence data in R with the TraMineR package: A user's guide. Department of Econometrics and Laboratory of Demography, University of Geneva, Switzerland. 1, (2008), 1–124.
[21] Gašević, D., Dawson, S., Rogers, T. and Gasevic, D. 2016. Learning analytics should not promote one size fits all: The effects of instructional conditions in predicting academic success. The Internet and Higher Education. 28, (2016), 68–84.
[22] Gašević, D., Dawson, S. and Siemens, G. 2015. Let's not forget: Learning analytics are about learning. TechTrends. 59, 1 (2015).
[23] Gašević, D., Jovanović, J., Pardo, A. and Dawson, S. 2017. Detecting Learning Strategies with Analytics: Links with Self-Reported Measures and Academic Performance. Journal of Learning Analytics. 4, 2 (2017), 113–128.
[24] Gašević, D., Kovanović, V. and Joksimović, S. 2017. Piecing the Learning Analytics Puzzle: A Consolidated Model of a Field of Research and Practice. Learning: Research and Practice. 3, 1 (2017), 63–78.
[25] Gatta, R., Lenkowicz, J., Vallati, M. and Stefanini, A. 2017. pMineR: Processes Mining in Medicine.
[26] Hadwin, A.F., Nesbit, J.C., Jamieson-Noel, D., Code, J. and Winne, P.H. 2007. Examining trace data to explore self-regulated learning. Metacognition and Learning. 2, 2–3 (2007), 107–124.
[27] Hattie, J. and Timperley, H. 2007. The Power of Feedback. Review of Educational Research. 77, 1 (2007), 81–112.
[28] Jovanovic, J., Gasevic, D., Dawson, S., Pardo, A. and Mirriahi, N. 2017. Learning analytics to unveil learning strategies in a flipped classroom. The Internet and Higher Education. 33, (2017), 74–85.
[29] Khan, I. and Pardo, A. 2016. Data2U : Scalable Real time Student Feedback in Active Learning Environments. LAK '16 6th International Conference on Learning Analytics and Knowledge, April 25 - 29, 2016. (2016), 249–253.
[30] Van der Kleij, F.M., Feskens, R.C.W. and Eggen, T.J.H.M. 2015. Effects of Feedback in a Computer-Based Learning Environment on Students' Learning Outcomes: A Meta-Analysis. Review of Educational Research. 85, 4 (2015), 475–511.
[31] Kluger, A.N. and DeNisi, A. 1996. Effects of feedback intervention on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. Psychological Bulletin. 119, 2 (1996), 254–284.
[32] Kovanović, V., Gašević, D., Joksimović, S., Hatala, M. and Olusola, A. 2015. Analytics of communities of inquiry: Effects of learning technology use on cognitive presence in asynchronous online discussions. Internet and Higher Education. 27, (2015), 74–89.
[33] Lai, C.L. and Hwang, G.J. 2016. A self-regulated flipped classroom approach to improving students' learning performance in a mathematics course. Computers and Education. 100, (2016), 126–140.
[34] Lust, G., Elen, J. and Clarebout, G. 2013. Regulation of tool-use within a blended course: Student differences and performance effects. Computers and Education. 60, 1 (2013), 385–395. DOI:https://doi.org/10.1016/j.compedu.2012.09.001.
[35] Maldonado-Mahauad, J., Pérez-Sanagustín, M., Kizilcec, R.F., Morales, N. and Munoz-Gama, J. 2018. Mining theory-based patterns from Big data: Identifying self-regulated learning strategies in Massive Open Online Courses. Computers in Human Behavior. 80, (2018), 179–196.
[36] Malmberg, J., Järvenoja, H. and Järvelä, S. 2010. Tracing elementary school students' study tactic use in gStudy by examining a strategic and self-regulated learning. Computers in Human Behavior. 26, 5 (2010), 1034–1042.
[37] Malmberg, J., Sanna, J. and Kirschner, P.A. 2014. Elementary school students' strategic learning: Does task-type matter? Metacognition and Learning. 9, 2 (2014), 113–136.
[38] Marton, F. and Säljö, R. 1976. on Qualitative Differences in Learning: I-Outcome and Process*. British Journal of Educational Psychology. 46, 1 (1976), 4–11.
[39] Mattick, K., Dennis, I. and Bligh, J. 2004. Approaches to learning and studying in medical students: Validation of a revised inventory and its relation to student characteristics and performance. Medical Education. 38, 5 (2004), 535–543.
[40] Molenaar, I. 2014. Advances in temporal analysis in learning and instruction. Frontline Learning Research. 6, (2014), 15–24.
[41] Nandagopal, K. and Ericsson, K.A. 2012. An expert performance approach to the study of individual differences in self-regulated learning activities in upper-level college students. Learning and Individual Differences. 22, 5 (2012), 597–609.
[42] O'Flaherty, J. and Phillips, C. 2015. The use of flipped classrooms in higher education: A scoping review. Internet and Higher Education. 25, (2015), 85–95.
[43] Pardo, A., Gasevic, D., Jovanovic, J.M., Dawson, S. and Mirriahi, N. 2018. Exploring Student Interactions with Preparation Activities in a Flipped Classroom Experience. IEEE Transactions on Learning Technologies. (2018).
[44] Pardo, A., Jovanovic, J., Dawson, S., Gasevic, D. and Mirriahi, N. 2017. Using Learning Analytics to Scale the Provision of Personalised Feedback. British Journal of Educational Technology. (2017).
[45] Pressley, M., Borkowski, J.G. and Schneider, W. 1987. Cognitive Strategies : Good Strategy Users Coordinate Metacognition and Knowledge. Annals of Child Development. 4, 2 (1987), 89–129.
[46] Price, M., Handley, K., Millar, J. and O'Donovan, B. 2010. Feedback : all that effort , but what is the effect ? Assessment & Evaluation in Higher Education. 35, 3 (2010), 277–289. DOI:https://doi.org/10.1080/02602930903541007.
[47] Proctor, B.E., Prevatt, F.F., Adams, K.S., Reaser, A. and Petscher, Y. 2006. Study Skills Profiles of Normal-Achieving and Academically-Struggling College Students. Journal of College Student Development. 47, 1 (2006), 37–51.
[48] Rahman, A.A., Aris, B., Rosli, M.S., Mohamed, H., Abdullah, Z. and Zaid, N.M. 2015. Significance of preparedness in flipped classroom. Advanced Science Letters. 21, 10 (2015), 3388–3390. DOI:https://doi.org/10.1166/asl.2015.6514.
[49] Verbert, K., Duval, E., Klerkx, J., Govaerts, S. and Santos, J.L. 2013. Learning Analytics Dashboard Applications. American Behavioral Scientist. February (2013), 1500-1509.
[50] Weinstein, C.E., Husman, J. and Dierking, D.R. 2000. Self-regulation interventions with a focus on learning strategies. Handbook of Self-Regulation. 22, (2000), 727–747.
[51] Winne, P.H. 2006. How Software Technologies Can Improve Research on Learning and Bolster School Reform. Educational Psychologist. 41, 1 (2006), 5–17.
[52] Winne, P.H. 2013. Learning Strategies, Study Skills, and Self-Regulated Learning in Postsecondary Education. Higher Education: Handbook of Theory and Research. 28 (2013), 337–403.
[53] Winne, P.H. and Hadwin, A.F. 1998. Studying as Self-Regulated Learning. Metacognition in educational theory and practice. 93, (1998), 277–304.
[54] Winne, P.H., Jamieson-Noel, D. and Muis, K. 2002. Methodological issues and advances in researching tactics, strategies, and self-regulated learning.
[55] Winne, P.H., Vh, D.J.H., Ode, R.U., Suhglfwv, J. and Riihuv, R. Chapter 21 : Learning Analytics for Self-Regulated Learning. 241–249.
[56] Winstone, N.E., Nash, R.A., Rowntree, J. and Parker, M. 2017. 'It'd be useful, but I wouldn't use it': barriers to university students' feedback seeking and recipience. Studies in Higher Education. 42, 11 (2017), 2026–2041.
[57] Yip, M.C.W. 2007. Differences in Learning and Study Strategies between High and Low Achieving University Students: A Hong Kong study. Educational Psychology. 27, 4 (2007), 597–606.
[58] Zeegers, P. 2001. Approaches to learning in science : A longitudinal study. (2001), 115–132.
[59] Zhou, M. and Winne, P.H. 2012. Modeling academic achievement by self-reported versus traced goal orientation. Learning and Instruction. 22, 6 (2012), 413–419.

Supplementary Material: Analytics of Learning Strategies: Associations with Academic Performance and Feedback

Table 2: Characteristics of the detected tactics

	Diverse	Reading oriented	Exercise oriented	MCQ oriented	Video oriented
State distribution					
Frequency of actions					

<table>
<tr><th></th><th>Diverse</th><th>Reading oriented</th><th>Exercise oriented</th><th>MCQ oriented</th><th>Video oriented</th></tr>
<tr><td>Order of sequence</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>Legend</td><td colspan="5">
</td></tr>
</table>

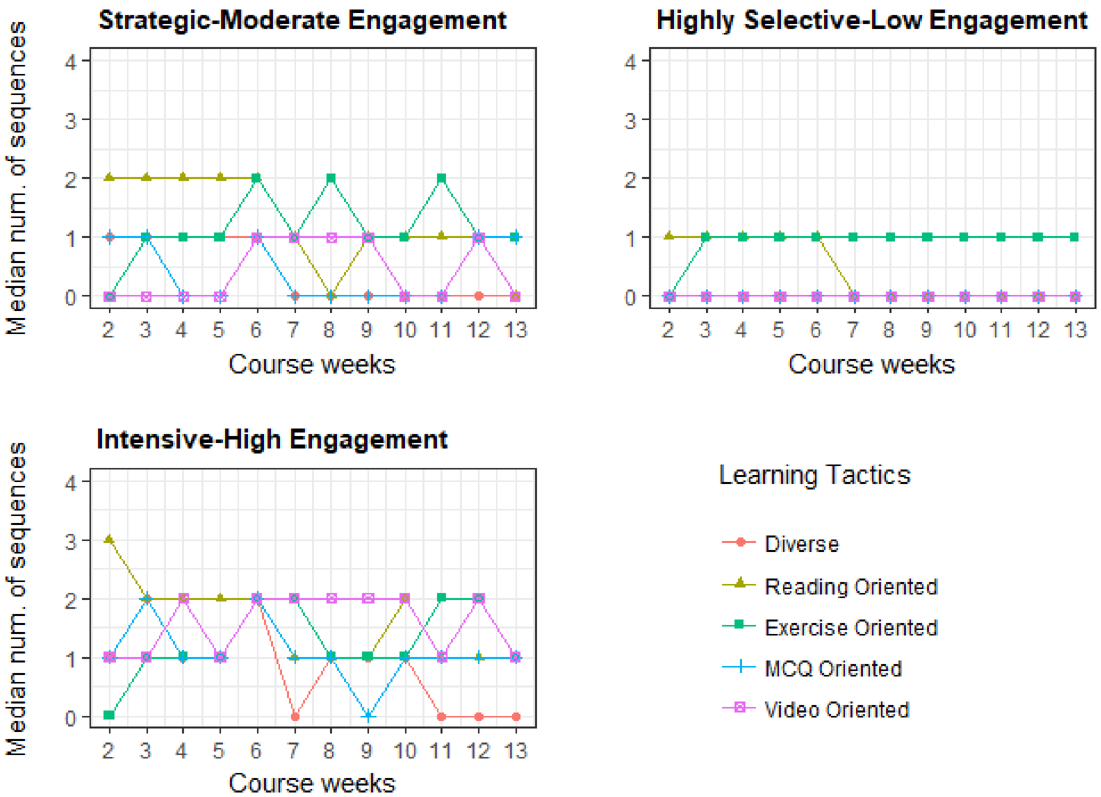


Figure 1: Weekly changes in the applied learning tactics for each of the detected learning strategies

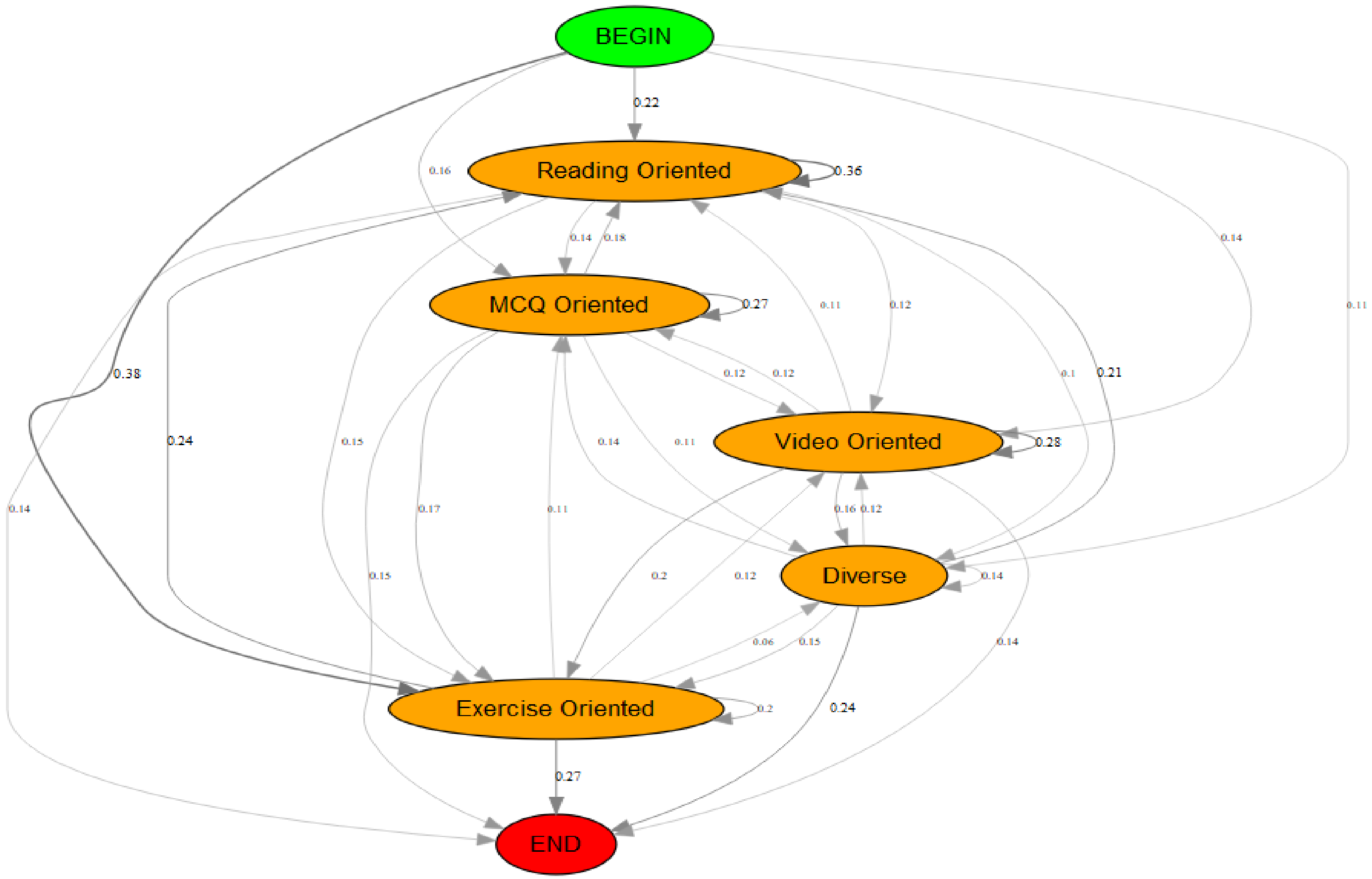


Figure 2: The Process models of the detected learning strategies: Strategic-Moderate Engagement Group

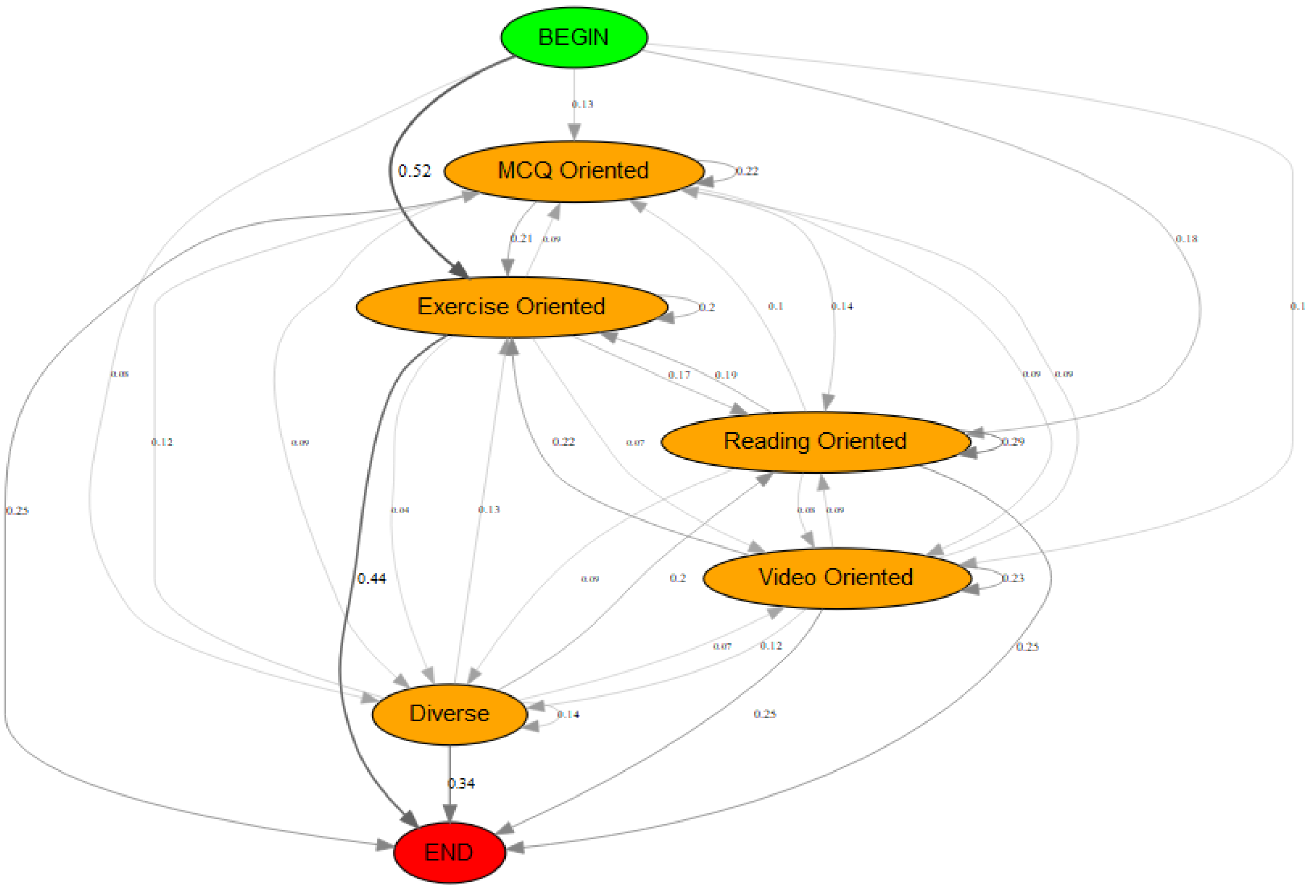


Figure 2: The Process models of the detected learning strategies: Highly Selective – Low Engagement Group

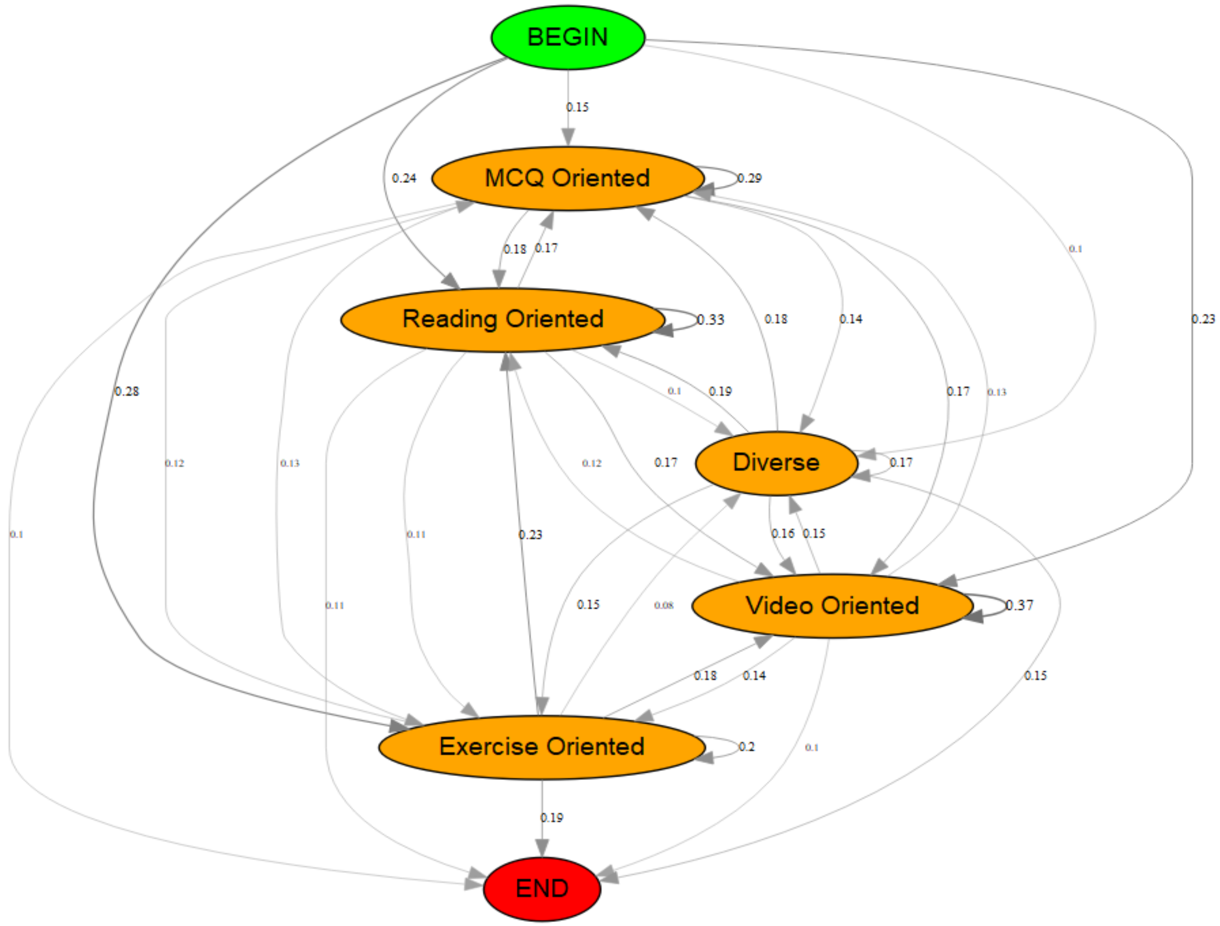


Figure 2: The Process models of the detected learning strategies: Intensive – High Engagement Group

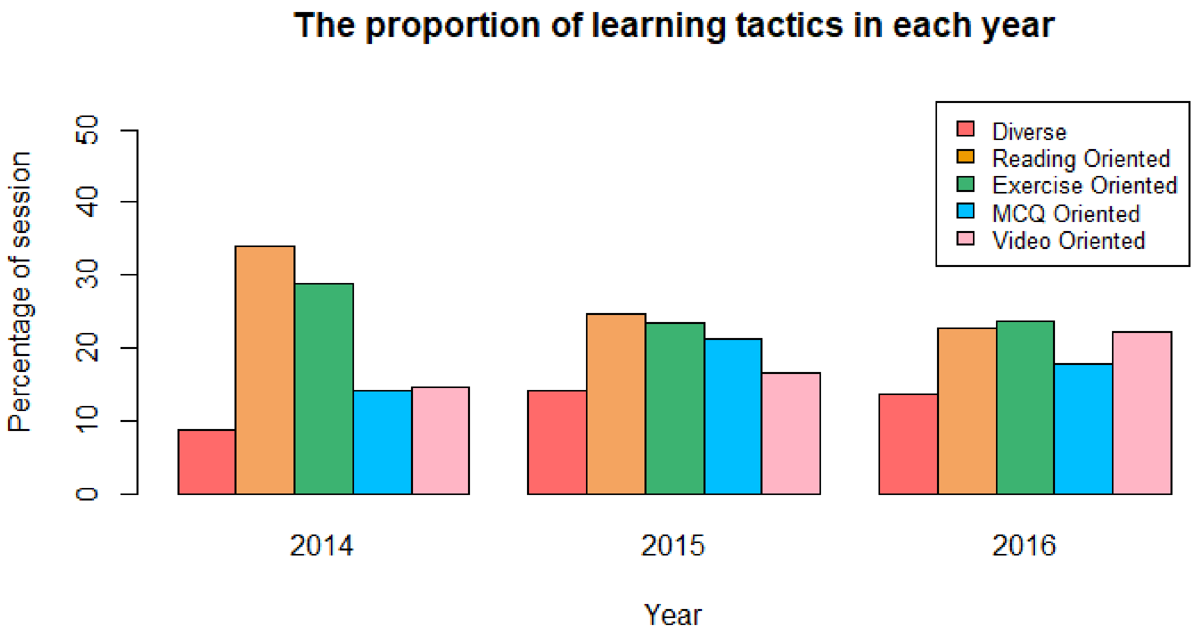


Figure 3: The proportion of learning tactics in each year

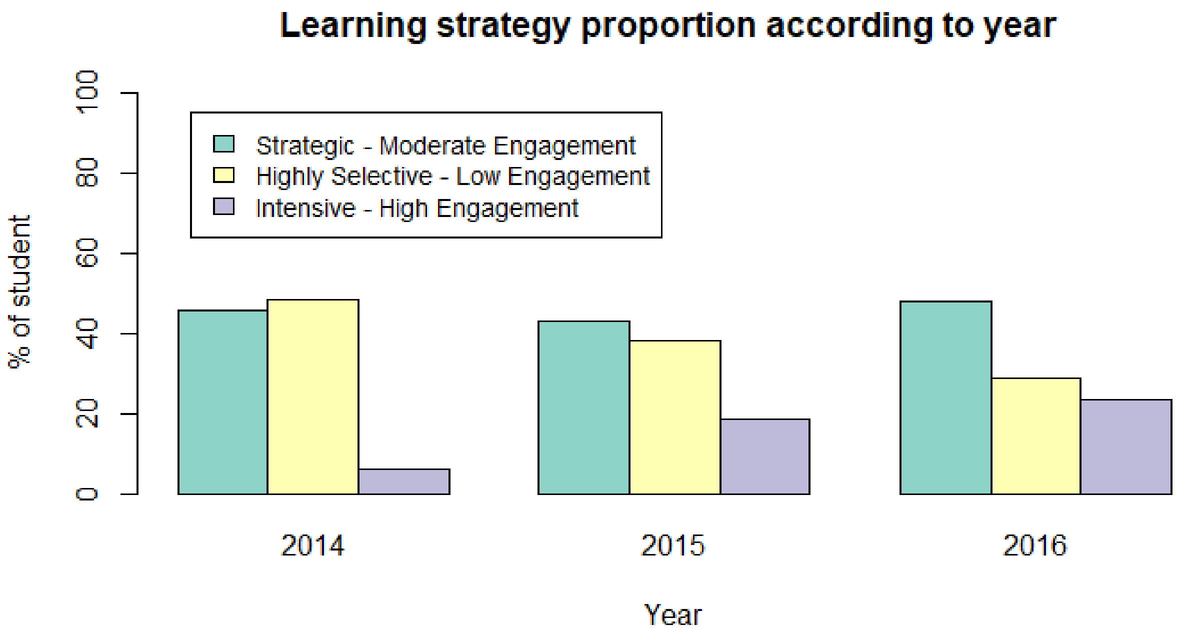


Figure 4: The proportion of learning strategy in each year

4.3 Summary

In this chapter, the process analytics-based approach, initially discussed in Chapter three, is examined in detail. The application of the approach is demonstrated in a different learning context i.e. flipped classroom. The data was collected from an online preparation session of the flipped classroom. Online preparation is considered as an important step of the flipped classroom, where students are expected to equip themselves with relevant knowledge required to be able to profit from the benefits of active face-to-face sessions of the flipped classroom (O'Flaherty & Phillips, 2015; Rahman et al., 2015). The study presented in the chapter offers in-depth insights into learning tactics and strategies. That is, sequences of actions showed the distribution of actions and dominant actions in each of the detected learning tactic. Meanwhile, the process mining showed the transitions between learning tactics as employed by the students during the course.

To align our analysis with the consolidated model of learning analytics (Figure 3) and to address the research question three (RQ3) in Section 1.1, we interpret the results based on the theory of approaches to learning (Biggs, 1987; Entwistle, 1991, 2007). The use of the process analytics-based approach detected learning strategies that were reflective of the approaches to learning. That is, three learning strategies were detected, i.e. *Highly Selective-Low Engagement, Intensive-High Engagement*, and *Strategic-Moderate Engagement*. The *Highly Selective-Low Engagement* strategy group is representative of a surface approach to learning, demonstrated by low level of engagement and high focus on assessment (Mattick et al., 2004). The *Intensive-High Engagement* strategy resembles a deep approach to learning. Students who applied this strategy showed a high level of effort to understand the learning materials by applying multiple learning tactics (Fincham et al., 2018; Jovanovic et al., 2017). The *Strategic-Moderate Engagement* strategy is representative of a strategic approach to learning. That is, students strategically focused on assessment-related activities with a considerable amount of effort spent to understand the learning materials. This result indicates that learning strategies as theorised in the model of approaches to learning can be detected with learning analytics approaches and trace data.

Both the *Intensive-High Engagement* and *Strategic-Moderate Engagement* strategy showed positive significant associations with high academic performance. Two similar features, including i) the tactics used and ii) high level of engagement were observed in these two strategy groups. In contrast to the *Strategic-Moderate Engagement* strategy, the students in the *Intensive-High Engagement* strategy group realised the benefit of formative assessment. Hence, the application of the *MCQ oriented* tactic, which was considered as a form of formative assessment, was applied in every week of the studying by the *Intensive-High Engagement* strategy group. Moreover, the students in the *Intensive-High Engagement* strategy group used the *Video and Reading oriented* tactics to gain deeper understanding on the content, rather than focusing only on the reading materials as done by students in the *Strategic-Moderate Engagement* strategy group. Thus, the grades obtained by students who followed the *Intensive-High Engagement* strategy was higher than the grades of those who followed the *Strategic-Moderate Engagement* strategy group. This showed that tactics applied in each strategy group contributed to academic performance. However, the extent to which the level of engagement could contribute to the academic performance was not explored in this study, although some trend is observed that higher engagement is associated with the strategies that led to higher academic performance. Future research is needed to explore how engagement plays in the role in explaining learning strategies and the association between learning strategies and academic performance.

Finally, one of the key contributions of this chapter is to explore the relation of the detected learning strategies with one of the constructs of SRL as posited in Winne and Hadwin's (1998) COPES model, namely, external evaluation or feedback. The investigation of the relationship addressed research question four (RQ4), which was formulated to explore how the SRL constructs are associated with the selection of learning tactics and strategies. Feedback has been identified as one of the factors that contributes to the application of learning strategies (Greene & Azevedo, 2007). The results of the study reveal that the introduction of feedback offered in terms of textual messages generated based students' data (i.e., 'personalised feedback'), has a positive association with the increase in the use of deep learning strategies. As demonstrated in this chapter, in the year when the students received personalised feedback, there was an increase in the use of the learning strategy (i.e., *Intensive-High Engagement*) which is indicative of the deep approach to learning. The increase in the application of this learning strategy is associated with high academic performance. This finding suggests that personalised feedback is a promising approach to the optimisation of student learning.

This study affirms the finding of the previous study (Chapter three) in which we report that the automatically detected learning strategies are reflective of the approaches to learning. The association of the SRL constructs with the choice of learning strategies is partially investigated in the current study (RQ4 in Section 1.1). That is, this chapter focuses on the external evaluation (i.e. feedback). The next chapter focuses on another important dimension of SRL constructs, i.e. the task (internal) conditions, and investigates how it shapes the selection of learning tactics and strategies.

5 Analytics of Learning Strategy: Role of Course Design and Delivery Modalities

Many external influences are driving change.

— Steven Redhead, *Life Is A Cocktail*

5.1 Introduction

The SRL model proposed by Winne and Hadwin (1998) emphasises the influence of task conditions on the selection of learning tactics and strategies. Task conditions refer to the constraints that are external to the cognitive process. Task conditions involve time constraints, availability of learning resources, social context, and instructional cues, to name a few (Winne & Hadwin, 1998). These constructs are in one way or another manipulated by the learning context.

Learning context is driven by many factors. Two main factors that have been found to be the most common features of a learning context are course instructional design and the delivery modalities (Elen & Clarebout, 2005). Course design outlines learning tasks created to guide students. Delivery modality refers to ways in which the course is delivered to the students (e.g., fully online, blended, flipped classroom, or face-to-face lectures).

The literature posits that learners modify the use of learning tactics and strategies according to a learning context (Broadbent, 2017; Lust et al., 2013; Morehead et al., 2019). For instance, the use of note-taking has been defined as one of the most commonly practiced methods used by learners when attending face-to-face classes (Dunlosky et al., 2013). However, the literature also finds that students often decide not to take any notes when studying online (Morehead et al., 2019). This chapter intends to investigate how learning contexts contribute to the application of learning tactics and strategies.

The learning analytics-based detection of learning tactics and strategies is introduced in Chapter three and Chapter four. These two chapters offer compelling evidence of a potentially reliable mechanism to extract learning tactics and strategies from trace data generated as a by-product of the use of real-world learning environments. However, most of the research in this area is context-dependent. That is, the proposed analytics approach is used in a specific context as outlined in

Chapter four. However, the generalisability of the method is an open research question. The second goal of this chapter, therefore, aims to examine the generalisability of the analytics approach to capture learning tactics and strategies.

5.1.1 Chapter overview

In this chapter, we present the results of a study which examined the used of the process analytics-based approach to detect learning tactics and strategies from trace data as proposed in Chapter three and Chapter four. The study used trace datasets from three different contexts. The first learning context was a course based on a flipped classroom delivery modality. The course was designed by emphasising a problem-solving based design. The dataset used was previously discussed in Section 4.1.1 of Chapter four (refer to Table 3).

The second dataset was collected in a course that was based on a blended learning delivery modality. The course was structured based on a practice-based design that focused on laboratory activities and workshops. Table 4 presents the description of learning activities observed in the trace data. In particular, most of the learning actions corresponded to phase 3 (enactment of learning tactics and strategies) of the Winne and Hadwin model of SRL. The goal-setting and planning (phase 1 of the Winne and Hadwin model of SRL) were observed by the access to course outline and structure.

The third dataset was collected in a course offered through a MOOC delivery modality. The course was centred around a problem-solving based design. The dataset was previously described in Table 2 presented in Chapter three, Section 3.1.2. By applying the process analytics-based approach, learning tactics were detected and explained in terms of sequences of actions using a process model. The learning strategies were detected based on the pattern of how students employ the detected learning tactics. The association of the detected learning strategies with academic performance was also examined. The exploration of three different learning contexts aims to validate the generalisability of the proposed analytics approach to detect the learning tactics and strategies. The validation of the method aims to address the research question two (RQ2) in Section 1.1.

To answer the research question three (RQ3), the results of this study are interpreted according to the approaches to learning (Biggs, 1987; Entwistle, 1991). Finally, the investigation across the different course instructional designs and delivery modalities contributes to the understanding of how task conditions (i.e. course designs and delivery modalities) may influence the adoption of learning tactics and strategies, as asked in research question four (RQ4).

Table 4. Description of the learning actions recorded in the trace data and the corresponding SRL phases

SRL Phases	Learning actions observed in the trace data	Description
1: Goal setting and planning	MetaCog: View Course Outline	Access to the course outline pages; this is considered a metacognitive orientation action
	MetaCog: View Course Structure	Access to the course structure information; this is considered a metacognitive orientation action
2: Tasks Identification		
3: Enactment of learning tactics and strategies	Assignment Submission Status	View the status of the submitted assignment
	Connect: Access	Access to external learning tool called 'Connect'
	Connect: Register	Register to access the external learning tool
	Discussion: Create	Create a discussion
	Discussion: Delete	Delete a posted discussion
	Discussion: Post	Post or reply to the discussed forum
	Discussion: Update	Update or edit the posted discussion
	Discussion: View	View the discussion posted in the forum
	Discussion: View News and Updates	View the announcement posted by instructors
	Home Page: Access	Access to the home page
	Pre-Lab: Launch	Start working on learning activities designed to help the students to prepare for the practical laboratory session
	Pre-Lab: Submit	Submit the pre-lab activities
	Pre-Lab: View	View the information and instruction to prepare for the physical laboratory
	Pre-Lab: View Status	View the status of submitted pre-lab activities
	Reading: Lecture Materials	Access to the reading materials
	Reading: Useful Information	Access to the recommended external reading material or other useful information
	YourTutor: Access	Access to the external tool called 'YourTutor'
	Other: Search	Search for keyword
	Other: Print	Print documents or files
4: Adaption		

5.2 Publication: Analytics of Learning Strategies: Role of Course Design and Delivery Modality

The following section includes the verbatim copy of the following publication:

> Matcha, W., Gašević, D., Ahmad Uzir, N., Jovanovic, J., Pardo, A., Lim, L., Maldonado-Mahauad, J., Gentili, S., Perez-Sanagustín, M., & Tsai, Y.-S. (2020). Analytics of Learning Strategies: Role of Course Design and Delivery Modality. *Journal of Learning Analytics*

JOURNAL OF LEARNING ANALYTICS

Volume 1(1), 1—10. http://dx.doi.org/xxx-xxx-xxx

Analytics of Learning Strategies: Role of Course Design and Delivery Modality

Wannisa Matcha[1]*, Dragan Gašević[2], Nora'ayu Ahmad Uzir[3], Jelena Jovanović[4], Abelardo Pardo[5], Lisa Lim[6], Jorge Maldonado-Mahauad[7], Sheridan Gentili [8] Mar Pérez-Sanagustín[9] Yi-Shan Tsai[10]

Abstract
Generalisability of the value of learning analytics based methods remains one of the big challenges in the field of learning analytics. One approach to test generalisability of a method is to apply it consistently in different learning contexts. This study extends a previously published work by examining the generalisability of a learning analytics method proposed for detecting learning tactics and strategies from trace data. The method was applied to the datasets collected in three different course designs and delivery modalities, including flipped classroom, blended learning, and massive open online course. The proposed method combines process mining and sequence analysis. The detected learning strategies are explored in terms of their association with the academic performance. The results indicate the applicability of the proposed method across different learning contexts. Moreover, the findings contribute to the understanding of the learning tactics and strategies identified in the trace data: the learning tactics proved to be responsive to the course design, whereas the learning strategies were found to be more sensitive to the delivery modalities than the course design. These findings, well-aligned with self-regulated learning theory, highlight the association of the learning contexts and the choice of learning tactics and strategies.

Notes for Practice

- Exploration on the detected learning tactics and strategies need to consider both sequential and temporal characteristics
- Learning tactics and strategies are context dependent, therefore, specific learning tactics and strategies have to be interpreted in the particular learning context from which the data originate
- Detected learning tactics should reflect the instructional course design.

Keywords
Learning Strategies, Learning Tactics, Data Mining, Self-regulated Learning, Modality, Course Design

Submitted: 07/01/20 — **Accepted:** 07/07/20 — **Published:** 10/08/20

[1] *School of Informatics, University of Edinburgh, Edinburgh, UK. email icon: w.matcha@ed.ac.uk*
[2] *Faculty of Information Technology, Monash University, Melbourne, Australia, and School of Informatics, University of Edinburgh, Edinburgh, UK. email icon: dragan.gasevic@monash.edu*
[3] *School of Informatics, University of Edinburgh, Edinburgh, UK, and Universiti Teknologi MARA, Shah Alam, Malaysia email icon: nora'ayu590@uitm.edu.my*
[4] *Faculty of Organizational Sciences, University of Belgrade, Belgrade, Serbia. email icon: jelena.jovanovic@fon.bg.ac.rs*
[5] *University of South Australia, Australia. email icon: abelardo.pardo@unisa.edu.au*
[6] *University of South Australia, Australia. email icon: lisa.lim@unisa.edu.au*
[7] *Pontificia Universidad Católica de Chile, Santiago Chile, and Universidad de Cuenca, Cuenca Ecuador. email icon: jorge.maldonado@ucuenca.edu.ec*
[8] *Teaching Innovation Unit, University of South Australia, Australia. email icon: sheridan.gentili@unisa.edu.au*
[9] *Pontificia Universidad católica de Chile, Santiago Chile, Université Paul Sabatier Toulouse III, Institute de Recherche en Informatique de Toulouse (IRIT), F-31062 Toulouse CEDEX 9. email icon: mar.perez-sanagustin@irit.fr*
[10] *School of Informatics, University of Edinburgh, Edinburgh, UK. email icon: yi-shan.tsai@ed.ac.uk*

1. Introduction

Self-regulated learning (SRL) skills are essential for successful learning in technology-enhanced learning environments. Among the key SRL skills are those related to the ability to identify effective learning strategies and knowing when and how to apply them (Winne & Hadwin, 1998; Lust, Elen, & Clarebout, 2013a). Winne and Hadwin (1998) asserts that the choice of learning strategies is influenced by the cognitive and task conditions, that is, students take into account the learning context when selecting the learning strategy.

The structure of the learning context is shaped by multiple factors that may facilitate or inhibit the construction of knowledge (Tessmer & Richey, 1997). In a technology-enhanced learning environment, two main factors that shape learning context are instructional design and learning modality. Instructional design or course design refers to the structure of the learning topics and the corresponding learning tasks designed to guide learning. Instructional design is typically driven by the nature of the discipline and the pedagogical approaches chosen to scaffold learning. For example, computer programming courses tend to rely on problem-solving and practical exercises, whereas humanities courses may require more theoretical development and discussion. The design of a course is also influenced by the delivery modalities, which refer to how and when learning activities are facilitated. For example, in a Massive Open Online Course (MOOC) setting, learning activities take place online, and offer flexibility in terms of free enrollment as well as when and where learners engage with the learning activities. Blended learning relies on online learning activities as a complement to support face-to-face learning activities, such that the online component is used during the pre-course preparations, in-class activities, and/or as part of revision. Closely related to the blended learning modality is flipped classroom where the emphasis is on promoting active learning (Pardo, Gasevic, Jovanovic, Dawson, & Mirriahi, 2018). In this mode, two components of learning settings are typically involved: i) online preparatory activities that are offered to students to prepare for ii) face-to-face learning sessions.

The variation in course structures and learning tasks, and delivery modalities all contribute to how learners adapt and adopt learning tactics and strategies as part of their studies (see Section 2.1 for definitions). Still, the role of learning context in the selection and adaptation of learning tactics and strategies remains largely unexplored – only few studies explored learning tactics and strategies adopted by students across different learning contexts (Broadbent, 2017). One possible reason is the difficulty in accurately detecting learning tactics and strategies used by students (Matcha, Ahmad Uzir, Gasevic, & Pardo, 2019). To overcome this limitation in the literature, research in learning analytics has focused on the development of data analytic methods to mine trace data to detect learning patterns indicative of learning tactics and strategies. Nonetheless, the methods currently reported in the learning analytics literature are usually context specific and raise the question of their applicability across different learning contexts.

In this study, we replicate the previous study done by Matcha, Gašević, Ahmad Uzir, Jovanović, and Pardo (2019) and validate an analytics method to detect learning tactics and strategies used by students when they interact with online learning activities in different learning contexts. The previous study has demonstrated the use of this analytic method in detecting learning strategies in trace data about students' online activities performed during their preparation for the face-to-face component of a flipped classroom course (Matcha, Gašević, Ahmad Uzir, Jovanović, & Pardo, 2019). That work opened up a question regarding the applicability of the proposed method in different learning contexts. This paper therefore, extends our original study to explore the applicability of the proposed analytic method in different learning contexts. The examined learning contexts are based on three different delivery modalities (blended learning, flipped classroom, and MOOC) and two different course designs (problem-solving and practice-based designs). We also examined how academic performance was associated with the students' choice of learning strategies.

2. Literature Review

2.1 Learning strategies and Learning Modalities

The terms 'learning strategy' and 'learning tactic' are often used interchangeably in the literature, although these terms refer to quite different concepts (Derry, 1989; Malmberg, Järvelä, & Kirschner, 2014). A learning tactic is a learning technique or a cognitive operation that is used by a student to perform a particular task (Derry, 1989; Malmberg, Järvenoja, & Järvelä, 2010). Students often combine two or more tactics to accomplish their learning objectives (Derry, 1989; Rachal, Daigle, & Rachal, 2007). Learning strategies, on the other hand, can be defined as "a coordinated set of study tactics that are directed by learning goal, and that are aimed at acquiring a new skill or gaining understanding" (Malmberg et al., 2014)[p 116]. Learning strategies, therefore, change and develop according to a learning situation, and involve selecting, combining, coordinating and utilising cognitive operations and techniques (i.e., learning tactics), directed by the learning goal (Winne & Hadwin, 1998).

Based on the model of self-regulated learning proposed by Winne and Hadwin (1998), the selection of learning tactics or strategies is driven by internal (e.g., cognitive) and external (e.g. task-related) conditions. In particular, acting as agents, learners select tactics and strategies based on several factors including their knowledge of the task at hand, tactics, and strategies, the available learning materials, instructional cues, as well as their motivation, beliefs, and goals (Winne, 2013). By examining the products of their learning and where they stand with respect to the expectations or learning goals, students can reflect on their performance and the effectiveness of their tactics and strategies. As a consequence, the internal conditions are updated, and the selection of strategies and tactics could be changed (Winne & Hadwin, 1998).

It has long been clearly noted by many scholars e.g. Winne and Hadwin (1998); Zimmerman (1986) that learning tactics and strategies are influenced by external conditions. These conditions are shaped by how the courses were designed and delivered. For instance, the technology has been diversely applied in current education to support learning. This has resulted in different learning modalities, blended and fully online learning being amongst the most prominent forms.

Blended Learning (BL) refers to the learning approach that combines face-to-face classroom with technology supported online learning activities (Bernard, Borokhovski, Schmid, Tamim, & Abrami, 2014). Technology is used to support various aspects of learning within this setting, such as preparation, in-class participation, and revision. The online section of blended learning is generally provided through digital learning tools that require students to self-direct their own learning (Bernard et al., 2014; Zhu, Au, & Yates, 2016). It is important to recognise that, online learning materials in this setting are not replacement for the face-to-face learning activities, they serve as complementary activities to support face-to-face learning.

Closely related to blended learning is flipped classroom (FC). Its key distinctive feature is that learning activities for teaching students the basic concepts and facts are carried out online, before the face-to-face sessions, whereas the face-to-face time is devoted to further knowledge construction through active learning (Pardo et al., 2018). Students are responsible to self-regulate their own learning to construct the knowledge required before attending face-to-face sessions (O'Flaherty & Phillips, 2015). The face-to-face sessions seek to promote the development of deep understanding, the rectification of misunderstanding and confusion (Rahman et al., 2015), and the development of problem-solving and critical thinking skills (Freeman et al., 2014; O'Flaherty & Phillips, 2015; Pardo, 2018).

Massive Open Online Courses (MOOCs) rose to prominence in 2012 by offering fully online learning at practically no costs for students. MOOC-based learning relies heavily on students' ability to self-regulate their learning and self-manage time (Kizilcec, Pérez-Sanagustín, & Maldonado, 2017; Eriksson, Adawi, & Stöhr, 2017). Several researchers and practitioners have adopted MOOCs as a way of transforming traditional classroom-based courses into blended learning mode (Rodríguez, Correa, Pérez-Sanagustín, Pertuze, & Alario-Hoyos, 2017; Pérez-Sanagustín, Hilliger, Alario-Hoyos, Kloos, & Rayyan, 2017).

Self-regulated learning skills and application of the effective learning strategies form an essential set of skills required across different learning modalities. For instance, Zhu et al. (2016) found that in a blended learning context, self-regulation strategies were predictive of academic performance. Furthermore, research found that the ineffective use of learning resources and poor study tactics had a negative impact on the learning performance (Lust, Elen, & Clarebout, 2013b). Similarly, Lai and Hwang (2016) found that in a flipped classroom-based course, students with high SRL scores performed significantly better than those who reported having low SRL scores. In the MOOC context, research found that high performing students employed different learning strategies as compared to the lower performance students (Nugent, Guru, & Namuth-Covert, 2018). Eriksson et al. (2017) relied on students' self-reports to examine the factors that contributed to the completion of a MOOC, and identified motivation, time management, and learning strategies as the prominent factors reported by learners.

Research published thus far clearly highlights the importance of employing effective learning strategies across different learning modalities. However, research has also reported that students often have underdeveloped skills to regulate and modify their learning strategies to meet the course requirements (Fincham, Gasevic, Jovanovic, & Pardo, 2018; Lust et al., 2013a). Guiding students to select the effective learning strategies is, therefore, important (Rachal et al., 2007). However, the difficulty in obtaining timely and informative insights into the students' learning tactics and strategies has prevented the provision of the required support and guidance on learning strategy selection and regulation (Matcha, Ahmad Uzir, et al., 2019).

2.2 Detection of Learning Tactics and Strategies

Traditionally, learning tactics and strategies have been detected by using self-reporting instruments such as surveys and think-aloud protocols. However, learners are not always accurate in reporting how they learn (Winne, 2013). For instance, think-aloud protocols can introduce cognitive overload as students are required to elaborate and justify their actions out loud while learning (Winne, 2013). By comparing actual learning activities to the self-reports, Hadwin, Nesbit, Jamieson-Noel, Code, and Winne (2007) found that self-reports did not reflect the students' actual behavior. Moreover, self-reports usually fail to capture how strategies develop over time. The use of learning trace data allows for understanding the students' actual learning behaviour without intervening in their learning or inadvertently increasing their cognitive overload. Zhou and Winne (2012) asserted that trace data were better correlated to the students' learning achievement than self-reports. Still, self-reports are successful in capturing students' perceptions and intentions and could assist in understanding the choices of actions that students make. Thus, Zhou and Winne (2012) referred to the insights obtained through self-reports as "perceived intentions". On the other hand, the data about students' actual learning behaviors, as recorded in the database of the learning platform, reflect the "realized intentions" of the students.

The use of trace data allows for examining the temporal dimension of learning tactics and strategies (Winne, 2017). In particular, learning tactics and strategies can be considered sequences of events, with each event being centered on one learning action (Hadwin et al., 2007; Molenaar, 2014). Such representation recognizes tactics and strategies as being dynamic, based on shifts from one action to another, and developing over time (Winne, Jamieson-Noel, & Muis, 2002). As already highlighted, a learning tactic is considered an operation performed by a student to complete a learning task. This operation is composed of a sequence of actions, and hence, is recognised as a tactic (Derry, 1989). Scrutinising Scrutinising learning tactics at the session level (i.e. within a particular period of time during which students were continuously interacting with learning materials) can provide a close observation on how students engaged in the learning process (Fincham et al., 2018). A learning strategy is defined as the application of one or more learning tactics (Malmberg et al., 2014), and therefore, it can be identified by detecting patterns in students' learning paths; that is, the regulation of sequences of learning tactics. To sum up, understanding the temporal and sequential dimensions of learning events could shed some light on how tactics and strategies have developed and changed, and allow for detecting situations where transitions between states happen (Molenaar, 2014). Such patterns of events evolve over time and become a characteristic of one's learning. This characteristic may be considered as an aptitude that could predict one's future behaviour (Winne et al., 2002).

The detection of learning strategies using trace data has been the focus of several research studies. At the core of these studies is the use of process and sequence mining techniques to analyze trace data. For example, Maldonado-Mahauad, Pérez-Sanagustín, Kizilcec, Morales, and Munoz-Gama (2018) classified the learning patterns, detected at the session level by using a process mining technique, into seven types 1) only video-lecture – the sessions in which students mainly interacted with the video contents; 2) only assessment – the learning sessions in which students only accessed the assessment activities; 3) explore – the sessions in which students were superficially accessing the assessment activities and video contents; 4) assessment try to video-lecture – the sessions in which students began to learn by accessing the assessment items and then switched to video content; 5) video-lecture complete to assessment try – the sessions in which students completed a video lecture but only accessed assessment activities; 6) video-lecture to assessment pass – students watched a video lecture, followed by passing an assessment; and 7) others. These types of learning patterns were triangulated with students' self-reported data about self-regulation strategies c. This resulted in the identification of three strategy groups of learners based on the data collected from three MOOCs. Specifically, Maldonado-Mahauad and colleagues found that comprehensive learners followed the course structure step by step, while targeting learners were focused on specific learning activities to help them pass the course. Both of these groups showed higher academic performance when compared to the sampling learners, who showed low and inconsistent engagement. The combined use of sequence mining and unsupervised machine learning is another common approach to detecting learning strategies from trace data. For example Jovanovic, Gasevic, Dawson, Pardo, and Mirriahi (2017) used a sequence mining technique together with the agglomerative hierarchical clustering to detect learning strategies based on the students' sequences of actions recorded when preparing online for a flipped classroom course. They identified four groups of learning tactics based on the patterns of learning actions within learning sessions, including sessions highly focused on 1) reading materials; 2) summative assessment; 3) formative assessment; and 4) video watching followed by assessment activities. Based on the regularity of the identified session-level learning patterns, they detected five groups of learners including intensive–highly active group who applied a variety of learning tactics; strategic–highly active group with emphasis on the interaction with summative and formative assessments; highly strategic who focused on summative assessment and reading materials; selective who focused on summative assessments with a low engagement with reading materials; and highly selective who only applied the summative assessment related tactic. Using the same course as the context of their study, Fincham et al. (2018) extracted learning tactics from trace data across three consecutive course offerings by using Hidden Markov Models (HMM) and identified learning strategies by examining sequences of the detected tactics separately in the first and second half of the semester. The detected tactics and strategies were explained based on the distribution of learning actions. Given the same data, Matcha, Gašević, Ahmad Uzir, Jovanović, and Pardo (2019) combined first order Markov models and the Expectation-Maximization algorithm to detect learning tactics across the three course offerings of the same course as Fincham et al. (2018). Despite the methodological differences, the results of the two studies were consistent. That is, the students who employed multiple learning tactics and were highly engaged in learning activities (i.e., those who adopted the deep approach to learning) tended to perform better. Those who applied a surface learning strategy with a low level of engagement, and were highly focused on the assessment were more likely to have lower performance (Fincham et al., 2018; Matcha, Gašević, Ahmad Uzir, Jovanović, & Pardo, 2019).

The studies presented thus far indicate that learning strategies can be derived from trace data and can be interpreted according to the theory of self-regulated learning and learning strategies (Maldonado-Mahauad et al., 2018; Fincham et al., 2018; Jovanovic et al., 2017). However, whether the same analytics-based tactic and strategy detection methods are applicable across different learning designs and modalities remains questionable since it has not been examined how general the applied methods are. As highlighted by Baker (2019), generalisability remains a big challenge in the field of learning analytics. Generalisability is difficult to achieve as the application of the same methodology to different learning contexts may produce different results. This inconsistency was observed, for example, in a study by Gašević, Dawson, Rogers, and Gasevic (2016). The study aimed to explore the factors that contributed to the success of nine undergraduate courses offered in blended learning modes. Gašević et al. (2016) found that the variables that were significantly predictive of the academic performance in one course were not significantly predictive when applied to another course. This effect was due to the diversity of course

JOURNAL OF LEARNING ANALYTICS

instructional designs. On the other hand, replicating an analytics method across different learning contexts is essential for the method to reach the maturity and acceptance. However, research focused on examining the general applicability of learning analytics methods is still limited (Gašević, Kovanović, & Joksimović, 2017).

To our knowledge, no studies have investigated how applicable an data analytic approach is in the detection of learning tactics and strategies across different learning contexts. Therefore, the goal of this study is to explore if the method proposed in our recent work Matcha, Gašević, Ahmad Uzir, Jovanović, and Pardo (2019) is applicable across different learning contexts and to identify general factors that may have contributed to the choices of learning tactics and strategies based on well-defined educational theories. As such the first research question has been defined as:

RQ1: Given a sequence of learning actions across several time frames, can we detect theoretically meaningful learning tactics and strategies applied by students when interacting with online learning activities across different course designs that are based on different delivery modalities?

2.3 Learning Strategy and academic performance

Learning strategies have impact on learning performance (Winne, 2006; Yip, 2007). Research findings denote that not all students use effective learning strategies (Dunlosky, 2013; Malmberg et al., 2014). Research done in traditional classrooms found that low and high performing students tend to apply different learning approaches (DiFrancesca, Nietfeld, & Cao, 2016; Proctor, Prevatt, Adams, Reaser, & Petscher, 2006). For instance, Nandagopal and Ericsson (2012) found that upper-level college students were regulating their learning strategies differently. Students with higher performance engaged more in applying a variety of learning strategies and had a higher tendency to review lessons. DiFrancesca et al. (2016) used self-reports to capture the students' perception of important learning strategies. They classified learning strategies into 'less effective' (e.g., repeating words, attending class, and reading the textbook), 'effective' (e.g., elaboration and connections, finding themes or main ideas, and application), and 'help-seeking'. Based on the self-reports they found that low performing students perceived less effective learning strategies as more important than effective learning strategies and relied more on less effective strategies during their learning process. Being based on an instructional approach that is substantially different from those of the traditional classroom, current technology-enhanced learning contexts could impose further challenges for regulation of learning and selection of effective strategies. Therefore, how students regulate their learning strategies and how that regulation is associated with their academic performances might differ across different learning contexts.

In the flipped classroom setting, Gasevic, Mirriahi, Dawson, and Joksimovic (2017) explored the association between learning strategies detected from traca data with data analytic methods and approaches to learning reported by the students and how the detected learning strategies were associated with the students' academic performances. They found that students who reported using the deep approach to learning tended to have significantly higher learning performance. Broadbent and Poon (2015) conducted a literature review to investigate the relationship of self-regulated learning strategies and academic performance of online learners. They found that four learning strategies namely, meta-cognition, time management, effort regulation and critical thinking, were significantly associated with course grades. Broadbent (2017) further explored the differences in self-reported self-regulated learning strategies across two different learning contexts, namely, online learning and blended learning. They found that time management strategy and effort regulation strategy were significantly correlated with higher performance in online learning settings. In blended learning settings, however, elaboration, organisation, meta-cognition, time management, and effort regulations strategies were found to be significantly correlated with higher academic performance. Broadbent (2017) observed that students who reported high frequency of using the rehearsal strategy were more likely to get a lower grade across two learning contexts. Students also reported a low frequency of using the peer learning and help seeking strategies in the online learning. Finally, Broadbent (2017) recognised the drawbacks of using self-reports to obtain information about students' learning strategies, mainly that students might not be able to recognize that they utilised certain strategies.

The studies presented so far have demonstrated the differences in the association between learning strategies used and academic performance across different learning contexts. However, the association of learning strategies and academic performances across different course designs and delivery modalities has remained insufficiently explored. Moreover, the majority of the extant research that examined this association relied on self-reports for collecting information about the students' learning strategies. To our knowledge, limited research reported in the literature explored the association between students' actual learning strategies (i.e., strategies derived from the learning activity records in a learning platform) and academic performance across different course designs and modalities. Therefore, our second research question is formulated as follows:

RQ2: Is there an association between learning strategies automatically detected with data analytic methods from trace data and students' academic performance in different course designs that are based on different delivery modalities?

3. Methods

JOURNAL OF LEARNING ANALYTICS

3.1 Data

3.1.1 Description of the data sets

Three data sets collected from three different learning contexts were used in this study. Table 1 presents a brief summary of the learning contexts and Table 2 provides a summary of the trace data sets collected in the three contexts. The data sets used in this study originate from different disciplines: two engineering courses and one science course. The structure and design of the courses were different. Table 3 offers a glimpse into the activities structure by presenting percentages of learning activities as observed in the trace data.

Table 1. Summary of the Learning Context

Dataset	Course	Modality	Years	Course Duration	Learning Activities
DatasetA	Computer Engineering	Flipped Classroom	2014-2016	13 Weeks	Lecture Videos, Reading materials, Quiz, Exercises
DatasetB	Biology	Blended Learning	2016-2017	13 Weeks	Reading materials, Pre-Lab External tools, Revision tools, Quiz
DatasetC	Introduction to Python	MOOC	2018	8 Weeks	Lecture Videos, Reading materials, Quiz, Summative Assessment

Table 2. Summary of the Collected Trace Data

Activities	Items	DatasetA	DatasetB	DatasetC
Learning Activities	Lecture Videos	Yes	-	Yes
	Quizzes Embedded in the Lecture Videos	Yes	-	Yes
	Reading Materials	Yes	Yes	Yes
	Quizzes Embedded in the Reading Materials	Yes	-	-
	Coding	-	-	Yes
	Conceptual Quizzes	Yes	-	Yes
	Practical Exercises	Yes	-	Yes
	Pre-Lab Activity	-	Yes	-
	Links to External Tools	-	Yes	-
Learning Supports	Discussion Forum	Yes	Yes	Yes
	Course Structure/ Overview/ Syllabus	Yes	Yes	-
Feedback Supports	Dashboard	Yes	-	-
	Personalised Message	Weekly	Week4 & Week9	-

DatasetA: The data set was collected in a first year Computer Engineering course that was organised based on a flipped classroom model and offered at an Australian university. The study examined data from three successive course editions, in years 2014, 2015, and 2016. The number of enrolled students steadily increased over the three years ($N_{2014} = 290$, $N_{2015} = 368$, and $N_{2016} = 477$). In all years the course was scheduled for 13 weeks with ten topics studied.

Students were required to complete online pre-class learning activities and attend face-to-face classes. The current study was focused on the online preparation activities, which were crucial for the success of the overall flipped classroom design (Rahman et al., 2015). Students were provided with a set of lecture videos, reading material, multiple-choice questions (MCQs) were embedded in both video and reading material, and problem-solving exercises. The details of instructional design is provided by Pardo and Mirriahi (2017); Pardo et al. (2018); Matcha, Gašević, Ahmad Uzir, Jovanović, and Pardo (2019).

Aside from learning activities, students were required to do a project. They were also provided with feedback. Two types of feedback were introduced over the three years: personal dashboards and analytics-based feedback in the form of personalised emails.The final assessment score of each student was determined by the completed problem-solving exercises during weekly preparation (20%); a laboratory report (5%); a laboratory project (15%); the midterm exam in week 6 (20%), and the final exam in week 13 (40%) (Pardo et al., 2018).

DatasetB: The second set of data was collected in the Introduction to Biology course, offered to first-year students at an Australian university (different from the university where DatasetA was collected). The data were collected over two years

(N_{2016} = 255, and N_{2017} = 232). The course was based on the blended learning approach that involved online weekly activities and face-to-face sessions. The course consisted of 13 learning weeks and covered 10 topics. Students were required to attend three face-to-face lectures and one tutorial per week. The course also included three obligatory workshops and seven sessions of obligatory laboratory practice. The students were required to prepare for laboratory practical sessions by completing the online activities offered through the Learning Management system (LMS). The students needed to complete these pre-lab activities in weeks 4-6 and weeks 9-12.

The student were also provided with online learning activities aimed to support self-revision after face-to-face lectures. The main learning materials consisted of lecture materials, other useful information, and external learning tools. The LMS provided students with links to these learning tools. The assessments were conducted four times. Two of them were administered in weeks 7 and 13 in the form of quizzes that contributed 20 percent of the final score (10% each). Practicals in the laboratories were counted for 25 percent towards the final mark and the final exam was 55 percent of the final mark. The students were provided with discussion forums to discuss the learning topics. In 2017, students received personalised messages as feedback during week 4 and week 9. The feedback aimed to create awareness of individual's learning progress, including attendance to face-to-face learning sessions, engagement with self-revision tool, and performance on the course assessment.

DatasetC: The third set of data was collected from the Introduction to Python MOOC offered by a university in Chile. The course was delivered fully online via the Coursera MOOC platform. The trace data were collected for one course offering in 2018 (N_{2018} = 368). The course was self-paced, scheduled for seven weeks and covered six programming topics. Each topic was divided into two to three subtopics. For each subtopic, a set of short video lectures with embedded questions (to provoke recollections of the concepts) and a set of reading materials were provided to support conceptual understanding. To support the practice of programming, students were required to complete two main learning tasks, namely, conceptual exercises (11 quizzes) and practical exercises (13 tasks). Among the quizzes and tasks, 22 items were graded. Students were allowed to make multiple attempts on the quizzes, and the best score was accumulated to calculate the students' final marks. The students were required to correctly solve at least 80 percent of these items in order to pass the course. The discussion forum was provided in the Coursera platform to support the students' learning.

3.1.2 Dataset Comparison

Table 3 presents the proportion of learning activities captured in each data set. The proportion of learning activities reflected how students distributed their time and effort to complete the given learning tasks. The learning activities are categorised according to the Activity Type Framework (Olney, Rienties, & Toetenel, 2018).The computer science and engineering courses (i.e. Computer Engineering and Introduction to Python) relied heavily on the practical exercises. Especially, in the Python course, the students spent 80 percent of their effort to practice by doing the available exercises. The distribution of student activities in the Computer Engineering course was rather diverse. Even though the majority of learning activities were focused on practical exercises (42.25%), interaction with the videos was also well represented (25.14%). Activities related to reading and conceptual quizzes had almost equal distributions of efforts. The biology course structure was different from the other two courses. Students relied heavily on the reading materials (44.88 %). The pre-lab activities (9.13 %) served as the exercises to ensure that students prepare before the physical attendance to the laboratory sessions during weeks 4-6 and weeks 9-12. Students participated in the discussion more often as compared to the other two courses (10.29 %).

3.2 Data analysis

Learning sessions were extracted from trace data, as sequences of consecutive learning actions, by assuming that any two consecutive actions within a sequence were within 30 min of one another. The sequences varied, both in terms of length and composition of learning actions. To normalise the data, the outliers, i.e. overly short sequences (consisting of one action) and overly long sequences (above the 95th percentile of the number of learning actions) were removed following the approach proposed by Jovanovic et al. (2017). As a result, DatasetA contained 65,710 learning sessions. The length of learning sequences ranged from 2 to 175 actions. DatasetB consisted of 25,648 learning sessions, ranging from 2 to 47 actions. DatasetC contained 5,281 learning sessions and comprised of 2 to 359 learning actions.

3.2.1 Detection of Students' Learning Tactics and Strategies

Figure 1 illustrates the pipeline of the analytic techniques used in this study. The first step in the pipeline is the detection of learning tactics using learning sessions as the input. A learning tactic can be considered as a sequence of actions that a student performs to complete the specified task (Fincham et al., 2018; Hadwin et al., 2007). To automatically detect learning tactics from learning sessions and address *research question RQ1*, we began by inspecting the learning process through a process mining lens. Process mining generates a process model based on a set of timestamped actions. By observing the overall process model, the potential number of learning tactics could be inferred based on the density of connections among actions. Process mining was performed by using the first order Markov models (FOMMs) as implemented in the pMineR R package (Gatta, Lenkowicz, Vallati, & Stefanini, 2017). The output of a FOMM is an adjacency matrix of transition probabilities

Table 3. Proportion of Learning Activities Collected from each Dataset According to the Activity Types Framework Adopted from Olney et al. (2018)

Activity Types	Description	Actions	DatasetA: Flipped Classroom Mode (100%)	DatasetB: Blended Learning Mode (100%)	DatasetC: MOOC Mode (100%)
Assimilative	Attending to information	Home and Updates	6.05	28.17	-
		Metacognitive - Orientation	0.96	1.25	-
		Lecture Videos	25.14	-	3.7
		Reading Materials	11.5	44.88	0.34
		Links to External Tools	-	4.73	-
Finding and handling information	Searching for and processing information	Utility Function	0.69	1.55	-
Communication	Discussing module related content	Discussion Forum	-	10.29	0.05
Productive	Actively constructing an artefact	Code Execute	-	-	4.09
Experiential	Applying learning in a real-world setting	Workshop/Lab	F2F	F2F	-
Interactive/ adaptive	Applying learning in a simulated setting	Pre-Lab Activities	-	9.13	-
Assessment	All forms of assessment	Conceptual Quizzes (inc. quizzes Embedded in the lecture videos and reading material)	12.08	-	11.09
		Practical Exercises	42.25	-	80.73
		Metacognitive - Evaluation	1.32	-	-

between events (i.e. learning actions). This output is suitable for cluster analysis using the Expectation — Maximization (EM) algorithm (Ferreira & Gillblad, 2009). Thus, EM was used to cluster learning sequences to detect meaningful learning tactics.

According to Derry (1989), a learning strategy employs one or more tactics. Therefore, learning strategies can be inferred from the way individuals employed tactics. Hence, in the second step, the Agglomerative Hierarchical Clustering based on Ward's algorithm (Gabadinho, Ritschard, Studer, & Muller, 2008) was used to extract patterns of how individual students used the identified learning tactics. As the input for the clustering process, for each student, we used the number of each identified tactic and the total number of all tactics. The Euclidean metric was used to calculate the (dis)similarity between vectors of the input. The dendrogram was used to determine the optimal number of clusters. This process, based on the hierarchical clustering, has already been established for the detection of learning strategies (Jovanovic et al., 2017; Kovanović, Gašević, Joksimović, Hatala, & Olusola, 2015).

3.2.2 Exploring the Detected Learning Tactics and Strategies

Patterns detected in the trace data require content-based interpretation to provide actionable insights (Maldonado-Mahauad et al., 2018; Pardo et al., 2018). To understand the characteristics of the detected clusters (i.e. tactics and strategies) and further address *research question RQ1*, sequence analysis and process mining were used. The TraMineR R package allows for constructing and examining sequences of actions (Gabadinho et al., 2008). It can be used to explore the frequency, the ordering, and the distribution of actions within sequences, and to explore clusters of sequences.

The pMineR R package (Gatta et al., 2017) was also used to explore the temporal characteristics of the learning strategy groups. Specifically, in computer science/engineering and biology courses, we analysed changes of learning tactics for each strategy group within each week of the course. We decided to use the week as the unit of analysis given the weekly activity cycles in those courses, as is common in formal higher education settings (Pardo et al., 2018). But in the self-paced learning course (i.e. Introduction to Python Course), it was more meaningful to explore the changes in tactics at the level of learning topics as the students were studying at their own pace and thus the schedules of individual students differed.

JOURNAL OF LEARNING ANALYTICS

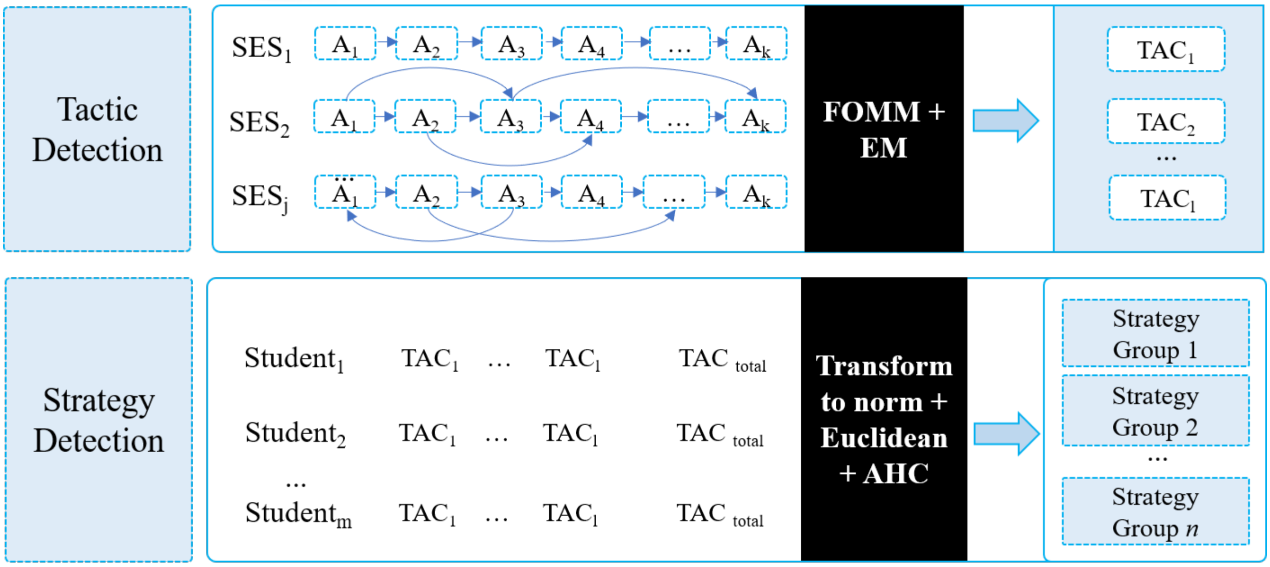


Figure 1. Learning Tactics and Strategies Detection Process (Matcha, Gašević, Ahmad Uzir, Jovanović, Pardo, Maldonado-Mahauad, & Pérez-Sanagustín, 2019)

3.2.3 Association with academic performance

Descriptive and inferential statistics were used to characterise the identified clusters (i.e., tactics and strategies). Kruskal Wallis tests were used to examine the association between learning strategies and course performance to address *research question RQ2*.

4. Result

4.1 RQ1: Detection of Learning Tactics and Strategies

4.1.1 Detection of Learning Tactics from Trace Data

Computer Engineering Course. Figure 2 presents the overall state distribution and state distribution in each year for the five detected tactics in the Computer Engineering course. The supplementary document that can be found online [1] presents the details of the detected learning tactics, Each tactic was labeled according to the characteristics of the detected patterns by considering the state distribution, ordering of sequences, frequency and the length of the sequences.

The five detected tactics can be summarised as follows:

- Diverse (N = 8,288, 12.61% of all sequences): characterised by the longest sequences of learning actions (median = 53 actions per sequence). The adoption of this tactic meant a variety of actions, with a relatively even distribution of exercises, MCQs, and video views.

- Reading oriented (N = 17,024, 25.91%): distinguished by the shortest sequences of learning actions (median = 4 actions in a sequence). The dominant kind of action was access to the reading materials.This similar patterns were observed across the three years of data collected.

- Exercise oriented (N = 16,287, 24.79%): characterised by a moderate number of learning actions (median = 24 actions per sequence). The most dominant learning actions were correctly (EXE_CO) and incorrectly (EXE_IN) solved the exercise questions. Unlike other tactics, most of the learning sequences in this tactic began by direct access to the problem-solving activities rather than access to the reading materials (Refer to Table 1 in the supplementary document).This tactic was presented similarly across the three years.

- MCQ oriented (N = 11,915, 18.13%): characterised by relatively short learning sessions (median = 5 actions per session) that often began by accessing reading materials (CONTENT_ACCESS), followed by MCQ answering. MCQ related

[1] https://bit.ly/37bzbQB

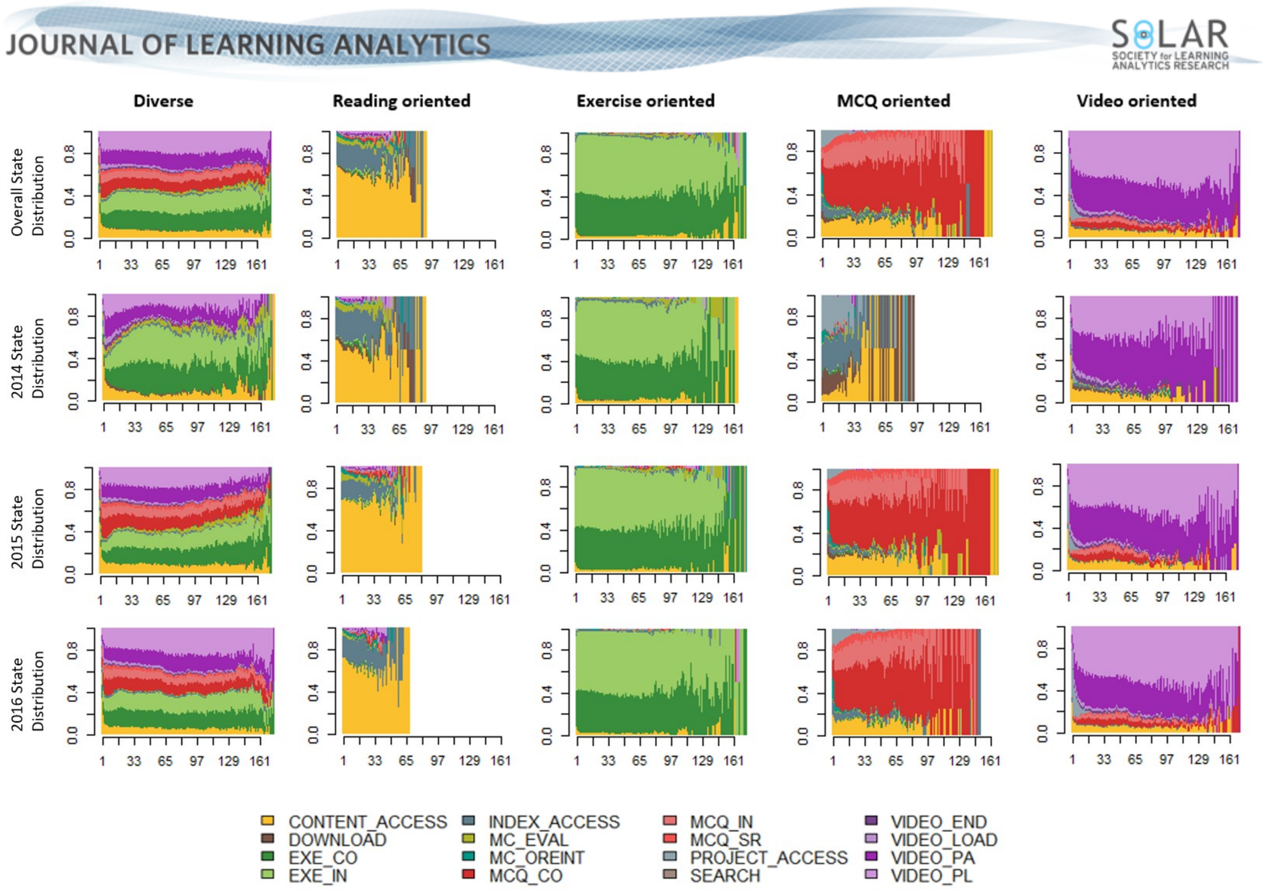


Figure 2. Learning tactics detected from trace data collected in the Computer Engineering course

actions including correctly answered (MCQ_CO) incorrectly answered (MCQ_IN), and requesting for help (MCQ_SR) were the most dominant type of action. These patterns of learning tactic were presented in 2015 and 2016. However in 2014, the most dominant actions were accessing to the index, project and other learning content rather than MCQ access. This is due to the proportion of MCQ related interaction collected in trace data for year 2014 was much lower than 2015 and 2016.

- Video oriented (N = 12,196, 18.56%): this tactic was associated with relatively short sessions (median = 9 actions per session). Based on the sequence length and dominant type of action, two types of learning sequences could be distinguished (Table 2). Long sessions often comprised of content access followed by video playing/pausing actions, which were in turn followed by MCQ-related actions (Refer to Table 1 in the supplementary document). Shorter sessions consisted mainly of access to the project information pages. Similar behaviors were observed across the three years.

Overall, metacognitive actions, which consisted of access to the dashboard (MC_EVAL) and course syllabus (MC_ORIENT), were noticeable in every tactic but showed relatively low presence compared to other types of actions (Figure 2 and Table 1 in the supplementary document)[2].

Biology Course. Three learning tactics were detected from the trace data collected in the Biology course. Figure 3 shows the overall state distribution and yearly state distribution of learning actions in each detected tactic. In general, the length of the learning sessions were relatively short.

- Reading (N = 11,358, 44.28% of all sequences): This tactic is characterised by relatively short learning sessions (median = 5 actions per session). The most dominant learning action was access to the reading materials and home page, which contained the general information about the course. Other types of learning actions were hardly observed. This tactic was observable across the two year of collected data.
- Reading and Pre-Lab (N = 4,287, 16.71%): The sessions grouped in this tactic were the longest learning sessions (median = 9 actions). The most dominant learning actions included access to the external reading materials and the course

[2] `https://bit.ly/37bzbQB`

JOURNAL OF LEARNING ANALYTICS

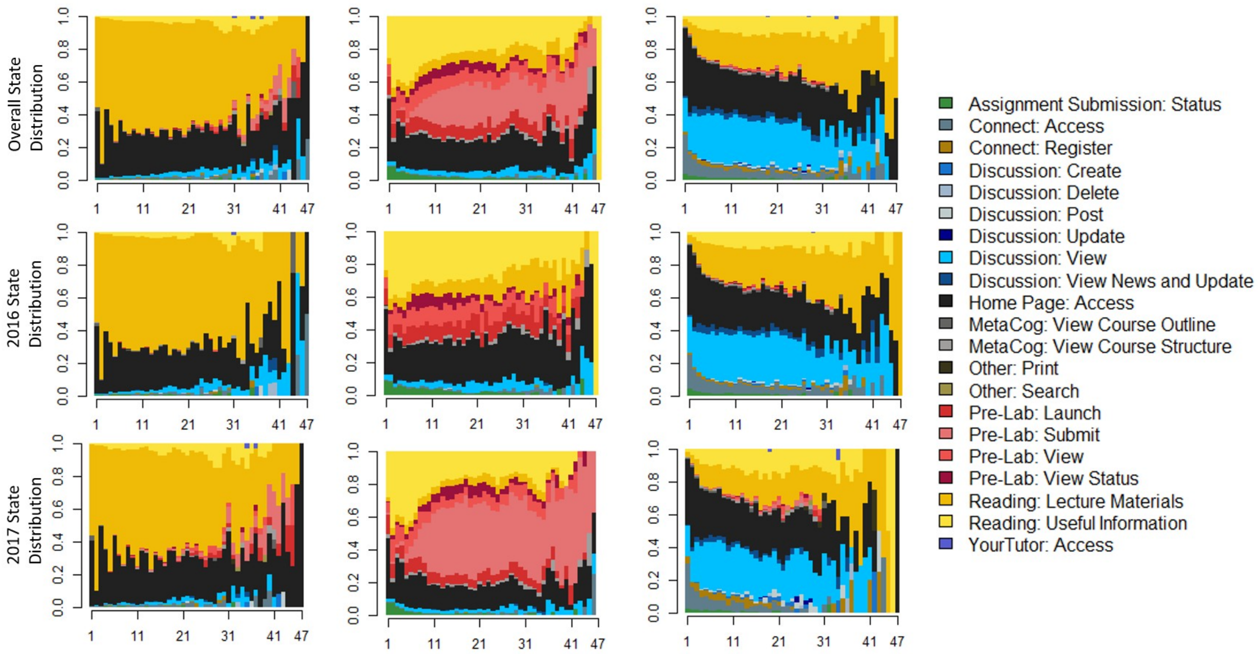


Figure 3. Learning tactics detected from from trace data collected in the the Biology course

homepage, as well as those related to the preparation for the laboratory practice. In 2016, the interactions with pre-lab activities were slightly lower as compared to 2017.

- Reading and Discussing (N = 10,003, 39.00%): This tactic is characterised by very short learning sessions (median = 4 actions). The most frequently observed learning actions in this tactic were accessing the course homepage and other course pages, viewing the discussion forum, and accessing the external revision tool. The patterns of this tactic were similar for both years.

Introduction to Python Course. Four learning tactics were detected from the trace data collected from the Introduction to Python course. Figure 4 shows the state distribution of learning actions for each detected tactic.

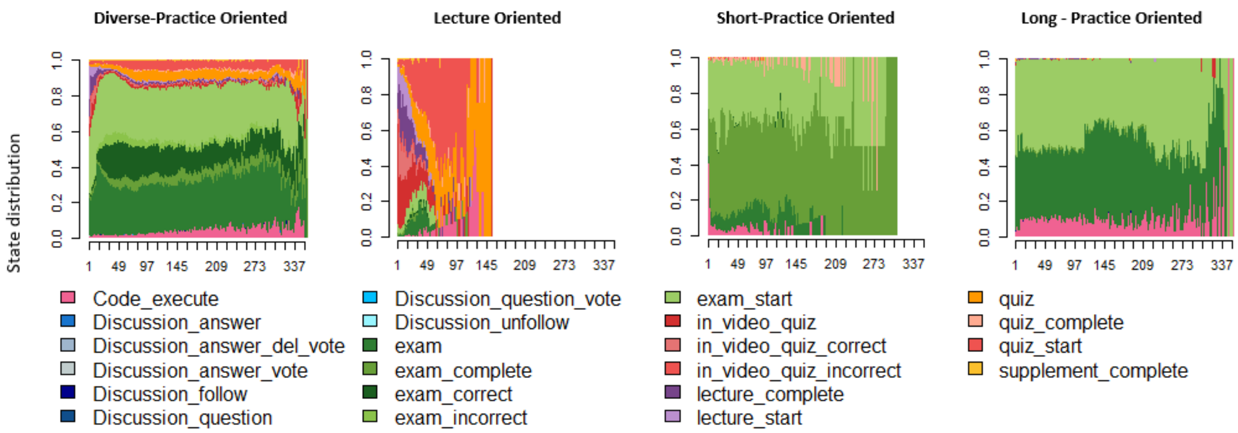


Figure 4. Learning tactics detected from trace data collected in the Introduction to Python course

- Diverse-Practice oriented (N = 2,000, 37.87% of all sequences): Each learning session contained a variety of learning actions. The learning sessions were quite long (median = 105 actions per session). The most dominant learning actions were practical exercises. Other types of learning activities such as quizzes, code execution, and video lectures were also observed.

- Lectures Oriented (N = 1,391, 26.34%): This tactic gathered short learning sessions (median = 7 actions) where students primarily interacted with the video lectures and answered to the embedded quizzes. Interactions with the theoretical quizzes were also observed.
- Short-Practice Oriented (N = 772, 14.62%): This tactic is characterised by short learning sessions (median = 8 actions). The most dominant learning actions were executing the code and performing practical exercises as shown in Figure 5.
- Long-Practice Oriented (N = 1,118, 21.17 %): This tactic exhibited similar pattern as the *Short-Practice Oriented*; that is, code execute and practical exercises (i.e. correctly or incorrectly solved the exam questions) were the most dominant learning actions. However, the learning sessions (median = 31 per session) were longer than those within the *Short-Practical Oriented* tactic. Figure 5 illustrates all learning sequences categorised as *Short-Practical Oriented* and *Long-Practical Oriented* tactics. The alternation between correctly and incorrectly solved exercises, and code execution were commonly observed in the *Long-Practical Oriented* tactic. The *Short-Practical Oriented* tactic contained two different patterns of practice related actions; i) very short sessions of code execution and ii) relatively long sessions of the work on the practical exercises.

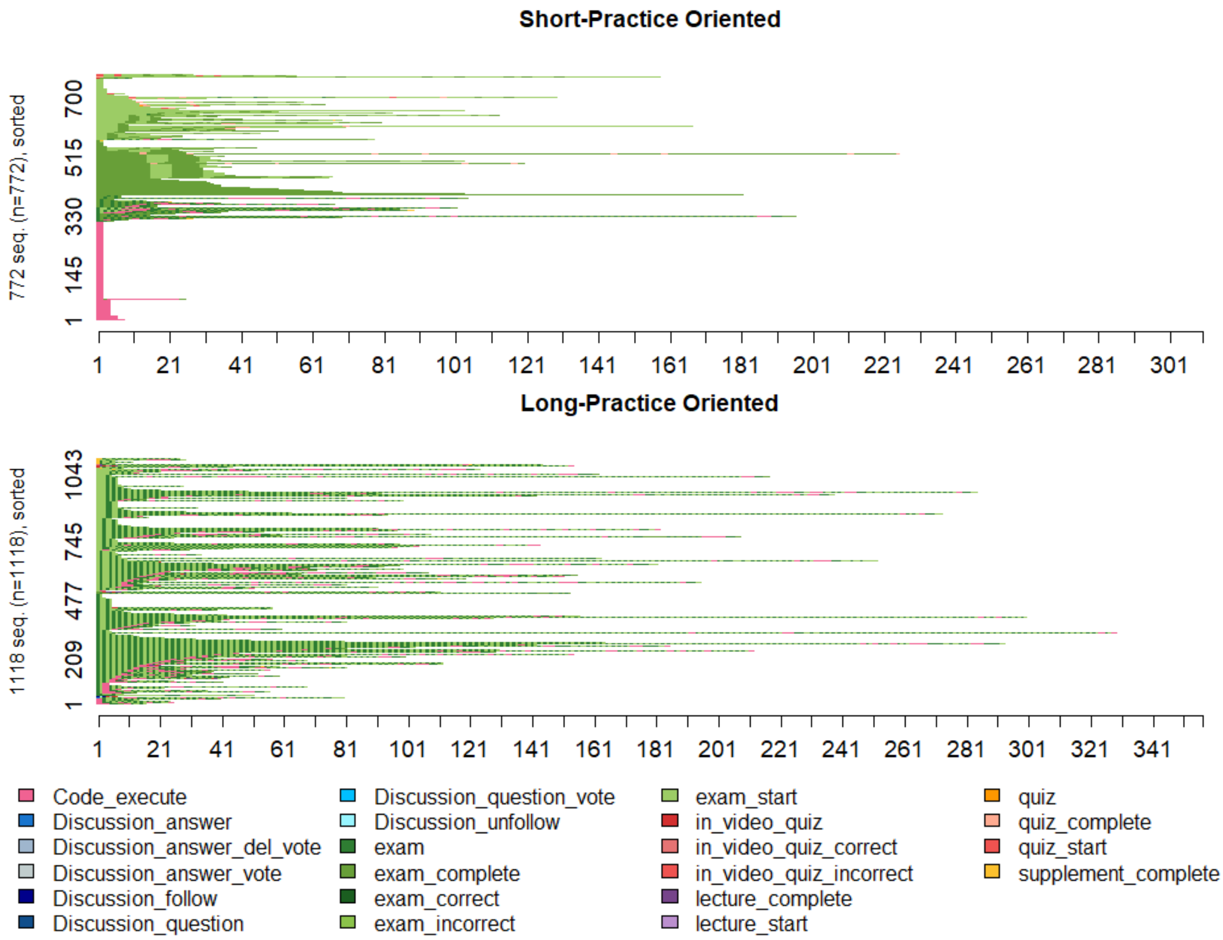


Figure 5. Different Learning Actions between Short- and Long-Practice Oriented Tactics in the Introduction to Python course

4.1.2 Detection of Learning Strategies from the Learning Tactics

Computer Engineering Course. The dendrogram produced by the employed clustering algorithm suggested three strategies as the best solution. Table 4 presents the summary statistics for the variables that served as the input for the clustering: the number of times each of the learning tactics was used as well as the overall number of learning tactics per student. To better understand the detected learning strategies, we also examined, for each strategy, how the use of learning tactics changed throughout the course. Figure 6 illustrates overall and yearly patterns of each detected strategy group, by presenting the median numbers of tactics applied in each week of the course

Table 4. Summary statistics (median, 1st and 3rd quartiles) for variables used for detecting learning strategies in the Computer Engineering course

Tactics	Strategy 1	Strategy 2	Strategy 3
Diverse	8.0 (5.0-11.0)	3.0 (1.0-5.0)	13.0 (8.0-17.0)
Reading	15.0 (11.0-22.5)	7.0 (4.0-11.0)	21.5 (14.0 30.0)
Exercise	15.0 (12.0-18.0)	12.0 (9.0-15.0)	17.0 (13.0-20.0)
MCQ	10.0 (7.0-15.0)	5.0 (2.0-7.0)	16.5 (10.0-27.0)
Video	11.0 (8.0-14.0)	4.0 (2.0-6.0)	23.0 (19.0-27.0)
Total	62.0 (53.0-72.0)	33.0 (25.0-41.0)	94.5 (77.2-111.0)

Strategy Group 1: Strategic -– Moderate Engagement: This was the largest cluster (N = 519, 45.73%). The students in this strategy group tended to use different learning tactics in different weeks of the course. Only the *Exercise* oriented tactic was consistently used throughout the semester across the three years of the courses included in this study. In the first half of the semester (week 2- week 6), in addition to exercises, the students also focused on reading materials and the associated MCQs. These patterns were observed in 2014 and 2015. However, in 2016, students applied lower frequency of reading tactic as compared to the other two years. In the second half of the semester (weeks 7-13), the *Exercise* oriented tactic was combined with the *Video* and *Reading* oriented tactics.

Strategy Group 2: Highly Selective -– Low Engagement: The proportion of the students in this strategy group was relatively high (N=418, 36.83%). The students in this group had low engagement with the preparation activities (Table 3). They chose to focus on specific types of learning tactics, namely, *Exercise* and *Reading* oriented. The *Exercise* oriented tactic was used throughout the semester, whereas the *Reading* oriented tactic was present only up until the midterm exam (week 6). This shallow regulation of learning tactics was consistently observed in all three years of collected data.

Strategy Group 3: Intensive — High Engagement: This was the smallest group (N = 198, 17.44%), comprised of the students with the highest engagement level, especially, in 2014. They applied a variety of learning tactics. The *Reading*, *Video*, *MCQ* and *Exercise* oriented tactics were used each week throughout the semester. The use of the *Diverse* tactic was observed in certain weeks only (week 6 and 10), particularly, in 2016 *Diverse* tactic was the most dominant tactic used.

To further our understanding of the detected learning strategies, first order Markov models were fitted to explore transitions from one learning tactic to another within each learning strategy. Figure 7 presents the process models of each learning strategy based on the transition of tactics.

As the course was designed based on the weekly basis, the process model of learning tactics application was built by using the weekly tactics applied by individual learners as the input. The main focus of the *Strategic — Moderate Engagement* strategy was on the exercises. There was a high probability (p=0.38) for the students in this group to have weeks based on the *Exercise* oriented tactic. The second most frequently use tactics was the *Reading Oriented*, which was observed with a lower probability (p=0.22). When it comes to changing learning tactics from one session to another one within a week, the most notable were the shift from the *Exercise* oriented to the reading-oriented tactics (p=0.24) and the shift from the Reading-oriented to Diverse tactics (p=0.21). The weeks often ended by using the exercise oriented and diverse oriented tactics. The main characteristics of the *Highly Selective — Low Engagement* strategy was its concentration on the *Exercise* oriented tactic. There was a strong probability that the students in this strategy group began and ended their weeks by doing exercises (p= 0.52, p= 0.44, respectively). The students who adopted the *Intensive — high engagement* strategy tended to begin their study weeks with a variety of learning tactics, rather than relying on one specific tactic. More precisely, the beginning of the week was almost equally distributed across the *Reading* oriented (p=0.24), *Exercise* oriented (p=0.28), and *Video* oriented tactics (p=0.23). Transitions between learning tactics were clearly observable and roughly equally distributed (probabilities ranged from 0.15 to 0.23). This indicates that the students in the *Intensive – High Engagement* group used a variety of learning tactics, which can be interpreted as an indication of their ability to self-regulate their learning (Winne, 2013).

Biology. The learning strategy groups were detected based on how students employed the tactics throughout the course. Similar to Figure 6, Figure 8 shows for each strategy group their weekly pattern in the use of learning tactics. Table 5 presents summary statistics (median, 1st and 3rd quartiles) for the learning tactics used by each of the detected learning strategy groups.

- *Strategy 1: Intensive – High Engagement* (N = 79 students, 16.22%): These students were quite active. They employed a variety of learning tactics, especially, the *Reading* and *Reading and Discussion* oriented tactics. However, in 2017 the

JOURNAL OF LEARNING ANALYTICS

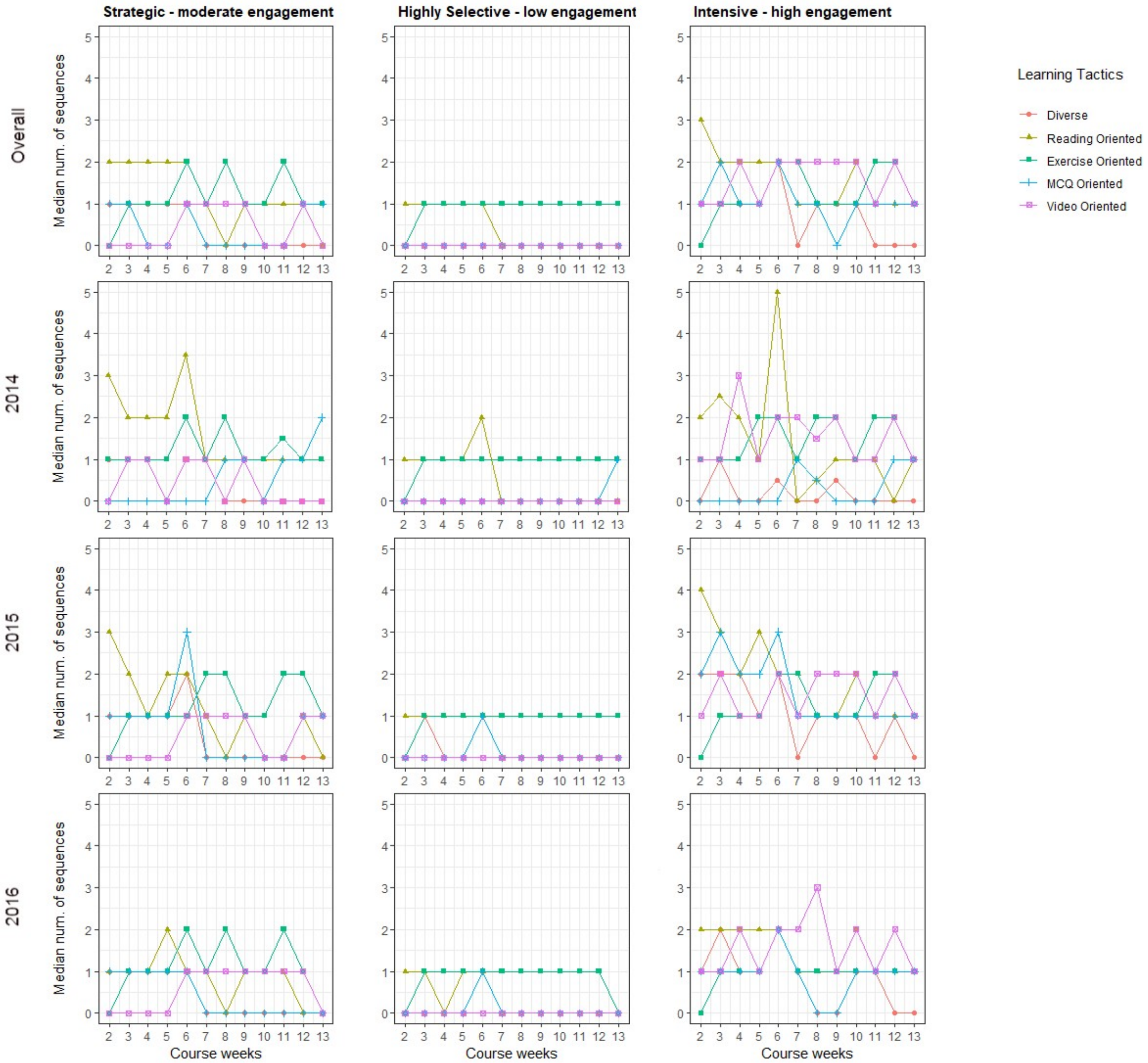


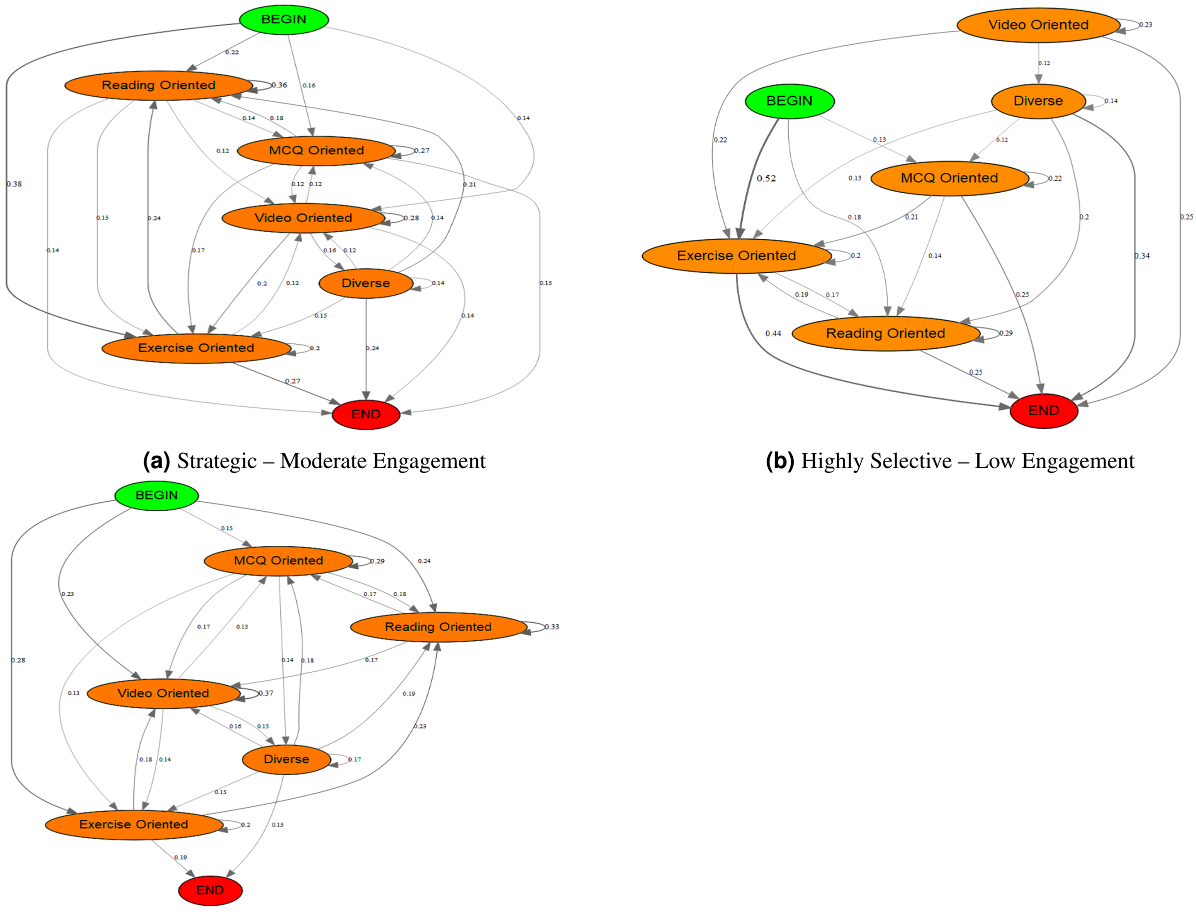
Figure 6. Learning strategies detected from trace data collected in the Computer Engineering course

median number of *Reading* tactics application was lower than 2016. As indicated by the high use of the *Reading and Discussion* tactic, students in this group were highly active in the discussion forum, especially during week 7 and week 13 when the mid-term and final exams were scheduled. The students accessed the pre-laboratory activities throughout the study weeks.

- *Strategy 2: Strategic –Moderate Engagement* (N = 193 students, 39.63%): The students in this group exhibited learning behaviors according to the course design. During weeks 4–6 and 9–12, the students were required to prepare for the face-to-face laboratory by completing a set of activities before the session. Access to laboratory activities was observed particularly during these weeks and it was invisible when it was not required. Meanwhile, the *Reading* tactic was used at least once a week. The *Reading and Discussion* tactic was also observed and its use peaked in week 6 when the mid-term exam was scheduled. The regularity of tactics application was similar for both 2016 and 2017.

- *Strategy 3: Highly Selective – Low Engagement* (N = 215 students, 44.15%): The students employed surface learning strategies in their online learning activities. In general, the learning activities were observed at the beginning of the course (weeks 1-4) and when the exams were scheduled (week 7 and week 13). The *Reading and Discussion* tactic was highly employed during the exam weeks. The preparation for the laboratories was not observed except for the first

JOURNAL OF LEARNING ANALYTICS



(a) Strategic – Moderate Engagement

(b) Highly Selective – Low Engagement

(c) Intensive - High Engagement

Figure 7. The process of tactics application on a weekly basis in the the Computer Engineering course

session in week 4. In 2017, the median number of tactics application was lower as compared to 2016.

The process model based on the transition of learning tactics is presented in Figure 9. In the *Intensive – High Engagement* strategy group, there was a high probability for the students to begin the learning sessions by using the *Reading and Discussing* and *Reading* oriented tactics (p=0.46 and p=0.38, respectively). The probability to start weekly learning sessions by using the *Reading and Pre-Lab* tactic was very low (p=0.16). The process model suggested that the students would apply the *Reading* or *Reading and Discussing* oriented tactic after using the *Reading and Pre-Lab* oriented tactic (p=0.30). There were almost an equal probability of transition between applying the *Reading* and *Reading and Discussing*. Unlike the *Intensive – High Engagement* strategy group, the *Strategic – Moderate Engagement* strategy group showed a strong probability to begin weekly online learning by applying the *Reading* oriented tactic (p=0.45) and repeatedly use this tactic (p=0.45). The second most applied tactic was *Reading and Discussing* (p=0.34). This strategy group had a higher probability of applying the *Reading and Pre-lab* at the beginning of the week (p=0.21) as compared to the *Intensive – High Engagement* strategy group. The *Highly Selective – Low Engagement* strategy group showed similar transitions as the *Intensive – High Engagement* strategy users did at the beginning of the weekly topics. They had a strong probability of beginning a study week by using the *Reading* (p = 0.41) and *Reading and Discussing* oriented tactics (p = 0.38). However, the transitions between tactics were less prominent as compared to the *Intensive – High Engagement* group. The students in this strategy group had a high probability of ending their weekly study after applying one learning tactic only.

Introduction to Python. The dendrogram suggested three learning strategy groups for the Python course. Table 6 shows summary statistics (median, 1st and 3rd quartile) for the number of the learning tactics used by each of the detected strategy group.

Unlike the other two learning modalities where the learning schedules were fixed, the Introduction to Python course, being delivered in a MOOC setting, offered the self-paced learning schedule. That is, students could begin and complete the learning

JOURNAL OF LEARNING ANALYTICS

Table 5. Summary statistics (median, 1st and 3rd quartile) for variables used for detecting learning strategies in the Biology course

Tactics	Strategy 1	Strategy 2	Strategy 3
Reading	40.0 (28.0 - 52.0)	25.0 (17.0 - 34.0)	12.0 (7.0 - 19.0)
Reading and Pre-lab	13.0 (10.0 - 17.0)	11.0 (8.0 - 13.0)	5.0 (3.0 - 7.0)
Reading and Discussion	44.0 (34.0 - 53.0)	20.0 (14.0 - 26.0)	10.0 (6.0 - 16.0)
Total	95.0 (82.5 - 112.0)	57.0 (49.0 - 65.0)	32.0 (21.5 - 40.0)

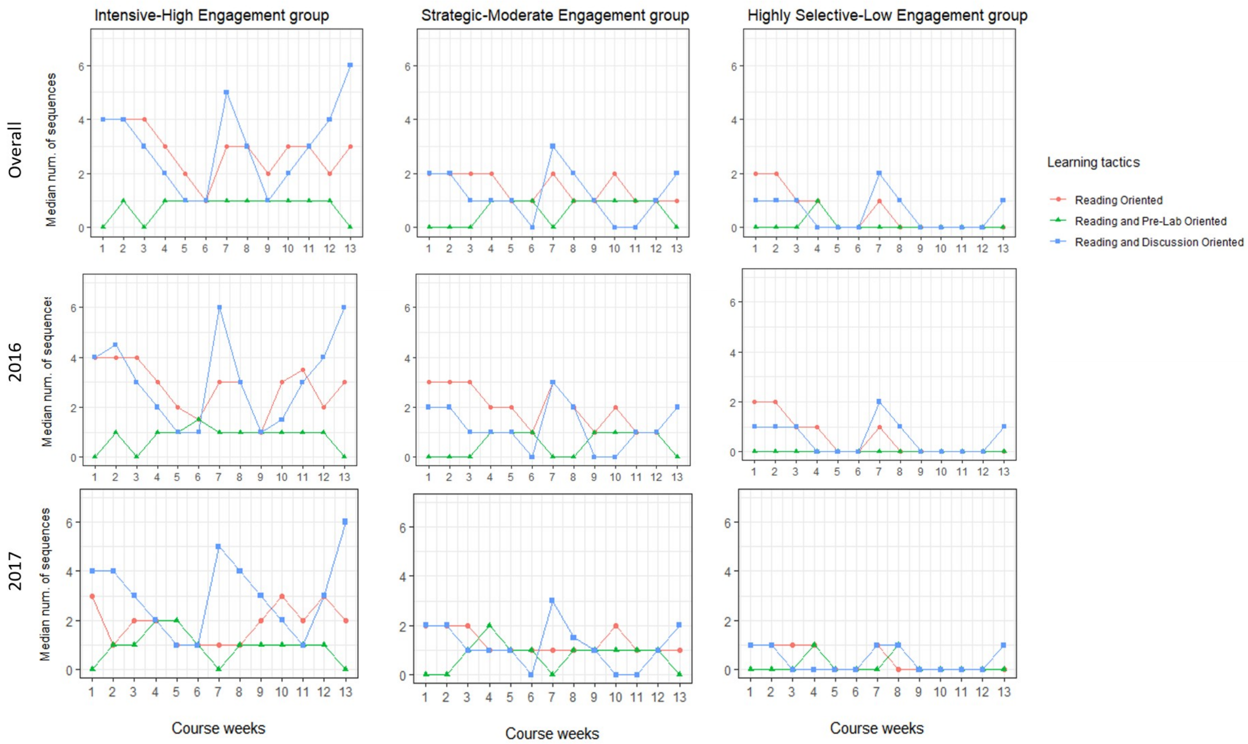


Figure 8. Learning strategies detected from trace collected in the Biology course

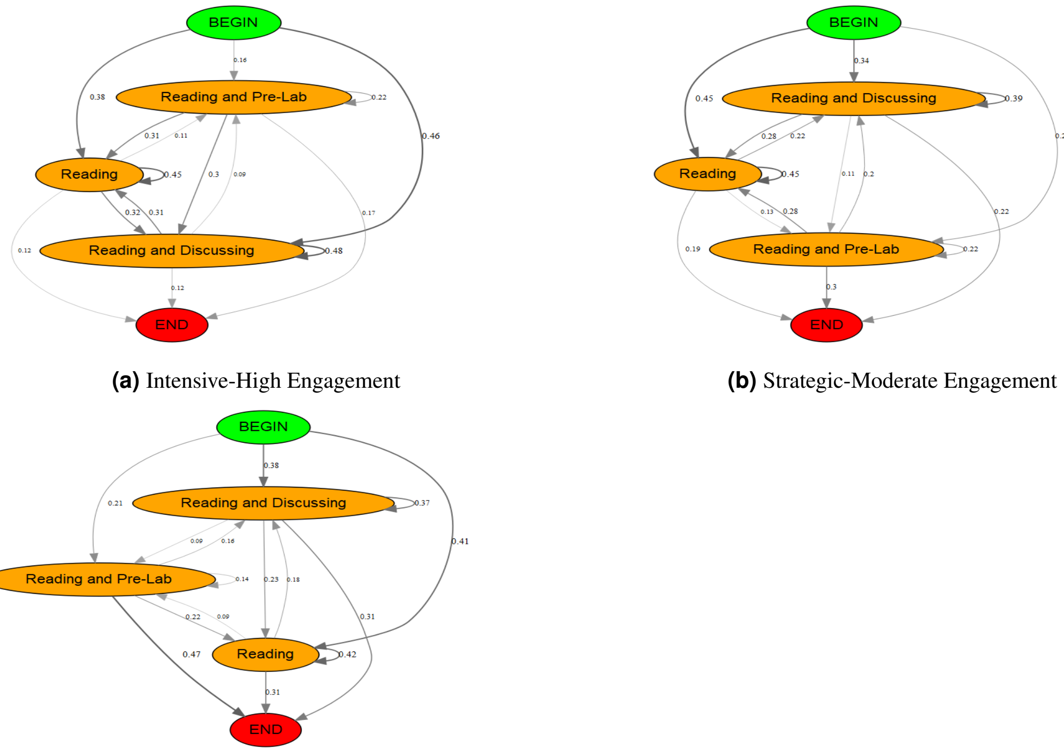
activities at their preferred timing. Therefore, the exploration of the learning strategies was done by focusing on the course topics rather than the course weeks. Figure 10 presents the average frequency of the four detected learning tactics according to the study topics.

- *Strategy 1: Inactive* (N = 215 students, 58.42%): The students employed a low level of learning effort. A low level of engagement was observed in relation to each study topic. The most dominant learning tactic was the *Diverse* tactic.

- *Strategy 2: Highly active at the beginning* (N = 89 students, 24.18%): The students were highly active at the beginning of the course (first two topics). The amount of effort dropped after the second topic. The dominant learning tactics used were the *Lecture* and *Diversity* oriented ones. One possible explanation for the observed behaviour is that the students were not familiar with programming or were inexperienced in learning to program. Even though the course design was strongly oriented towards practical exercises (see Table 3), the students highly used the *Lecture* oriented tactic rather than employing the *Short* or *Long-Practice* oriented tactics.

- *Strategy 3: Highly active* (N = 64 students, 17.39%): The students were highly active. Multiple uses of different tactics were observed. The use of the *Long-Practice* oriented tactic increased as the course progressed. The *Short-Practice* oriented tactic was used at least once when interacted with each course topic but highly used during the fourth topic. Similarly, the *Diverse* tactic was used at least once in each course topic but highly used when interacted with the third and fourth topic. The *Lecture* oriented tactic was also observed at least once for each learning topic.

JOURNAL OF LEARNING ANALYTICS



(a) Intensive-High Engagement

(b) Strategic-Moderate Engagement

(c) Highly Selective-Low Engagement

Figure 9. The process of tactic application on a weekly basis in the Biology course

Table 6. Summary statistics (median, 1st and 3rd quartile) for variables used for the learning strategies detected in the Python course

Tactics	Strategy 1 - Inactive	Strategy 2-Highly Active at the Beginning	Strategy 3 - Highly Active
Diverse	4.0 (2.0 - 7.0)	7.0 (4.0 - 11.0)	14.0 (11.0 - 20.2)
Lecture-Based	2.0 (0.5 - 3.0)	8.0 (7.0 - 11.0)	5.0 (2.0 - 9.2)
Short-Practice	0.0 (0.0 - 1.0)	0.0 (0.0 - 2.0)	1.5 (0.0 - 3.0)
Long-Practice	1.0 (0.0 - 3.0)	2.0 (1.0 - 3.0)	9.0 (7.0 - 13.0)
Total	8.0 (5.0 - 13.0)	20.0 (15.0 - 28.0)	29.5 (25.0 - 44.5)

The process model based on the tactics applied by each strategy group per each course topic of the Python course is presented in Figure 11. The central learning tactic used in *Inactive* strategy group was the *Diverse-Practice* oriented one. The *Inactive* strategy users frequently used the *Diverse-Practice* oriented tactic at the beginning of interaction with each topic (p = 0.59) and at the end of their topic related activities (p = 0.36). There was an observable common practice to begin interaction with each topic by using the *Lecture* oriented tactic, and ended the topic-related activity by applying the *Diverse-Practice* tactics. A transition from applying the *Long-Practice* to *Diverse-Practice* tactics could also be observed (p = 0.35). The *Highly Active at the Beginning* strategy group had a high probability of beginning their learning in each topic by either applying the *Diverse-Practice* (p=0.52) or *Lecture* oriented tactics (p=0.42). They had a strong probability of re-using the *Lecture* oriented tactic (p=0.53). Similar to the *Inactive* strategy, transitions from *Long-Practice* to *Diverse-Practice* tactics could also be observed (p = 0.39). This strategy group had equal transitions between the *Lecture* and *Diverse-Practice* oriented tactics. The *Highly Active* strategy group emphasized the use of the *Diverse-practice* tactic. They had a strong probability of beginning their learning in each topic with a long session of the *Diverse-practice* oriented tactic (p = 0.65). The transition from *Lecture* to *Diverse-Practice* oriented tactics (p=0.42) were observed. The most dominant transition in the *Highly Active* strategy group was the repetition of the *Short-Practice* oriented tactic (p=0.85). This indicates that the students tended to practice by repeatedly doing the exercises designed for each topic covered in the course.

JOURNAL OF LEARNING ANALYTICS
SOLAR SOCIETY for LEARNING ANALYTICS RESEARCH

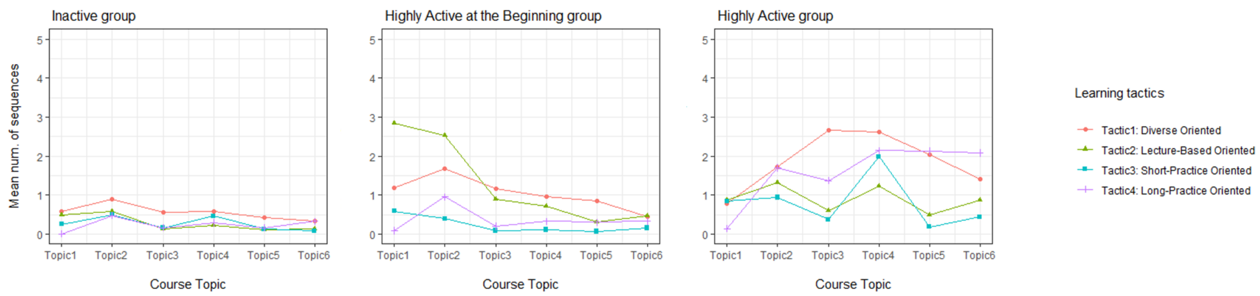


Figure 10. Learning strategies detected in the Introduction to Python course

4.2 RQ2: Associations between Learning Strategies and Academic Performance

Computer Engineering. Summary statistics for the students' course performance per learning strategy group are presented in Table 7. Kruskal Wallis tests showed a significant association between student learning strategy groups and the students' course performance (p-value < 0.0001) for both midterm and final exam scores. To further examine the associations between the detected learning strategies and the students' academic performance, we did pairwise comparisons of learning strategy groups with respect to the midterm and final exam scores.

As shown in Table 8, significant differences with respect to both midterm and final exam scores were present between *Strategy 1: Strategic* and *Strategy 2: Highly Selective* and between *Strategy 3: Intensive* and *Strategy 2: Highly Selective*. However, no statistically significant difference was found between *Strategy 1: Strategic* and *Strategy 3: Intensive* group with respect with midterm and final exam score.

Table 7. Summary statistics (median, 1st and 3rd quartile) for course score for each strategy group

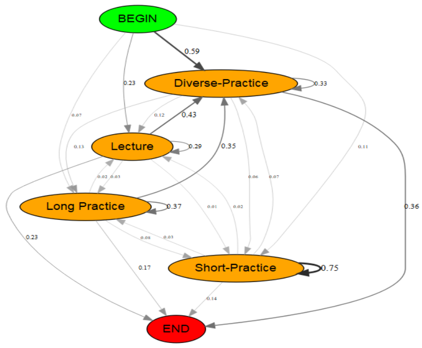
Course	Performance	Highly Selective	Strategic	Intensive
Computer Engineering	Midterm (20)	12 (10 – 15)	15 (12 – 17)	15 (12 – 17)
	Final Exam (40)	15 (10 – 20)	21 (14 – 30)	22 (16 – 32)
		Highly Selective	Strategic	Intensive
Biology	Total Score (100)	63.7 (48.0 - 75.5)	71.3 (59.0 - 78.1)	71.4 (62.1 - 81.2)
		Inactive	Highly Active at the beginning	Highly Active
Python	Total Score (100)	29.3 (7.2 - 58.5)	20.4 (6.9 - 63.8)	82.7 (54.2 - 89.6)
	Passed Graded Item (22)	7.0 (2.0 - 15.0)	5.0 (2.0 - 16.0)	20.0 (13.0 - 22.0)

Table 8. Pairwise comparison of strategy groups with respect to the scores

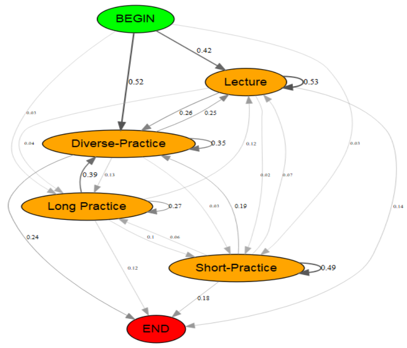
Performance	Strategy	Strategy	Z	p	r
ComEng: Mid-term Score	Strategic-Moderate Engagement	Highly Selective-Low Engagement	9.329	< 0.0001*	0.31
	Strategic-Moderate Engagement	Intensive-High Engagement	-0.1577	0.875	0.01
	Highly Selective-Low Engagement	Intensive-High Engagement	-7.0258	< 0.0001*	0.28
Final Exam Score	Strategic-Moderate Engagement	Highly Selective-Low Engagement	9.338	< 0.0001*	0.311
	Strategic-Moderate Engagement	Intensive-High Engagement	-1.851	0.064	0.077
	Highly Selective-Low Engagement	Intensive-High Engagement	-8.847	< 0.0001*	0.366
Biology: Course Score	Intensive-High Engagement	Strategic-Moderate Engagement	0.841	0.34	0.051
	Intensive-High Engagement	Highly Selective-Low Engagement	3.568	< 0.0001*	0.208
	Strategic-Moderate Engagement	Highly Selective-Low Engagement	3.628	< 0.0001*	0.18
Python: Course Score	Inactive	Highly Active at the Beginning	0.108	0.914	0.006
	Inactive	Highly Active	-6.927	< 0.0001*	0.415
	Highly Active at the Beginning	Highly Active	-5.879	< 0.0001*	0.475

Biology. Table 7 presents the median, first and third quartile of the course mark of the Biology course. Average exam marks

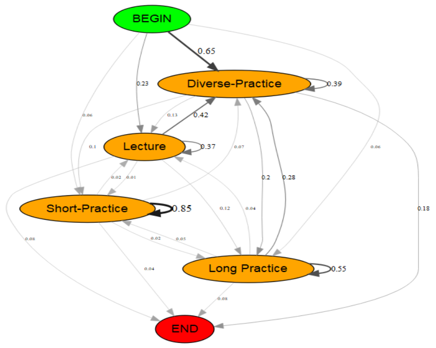
JOURNAL OF LEARNING ANALYTICS



(a) Inactive



(b) Highly Active at the Beginning



(c) Highly Active

Figure 11. The process of tactics application on a related learning topic in the Introduction to Python course

of the *Intensive – High Engagement* and *Strategic – Moderate Engagement* strategy groups were almost equal (median = 71.4 and 71.3, respectively), but the *Highly Selective – Low Engagement* group had lower average course marks (median = 63.7). Kruskal-Wallis tests revealed a statistically significant difference (p-value$<$ 0.0001) among the learning strategy groups with respect to total course score. The results of the follow up pairwise comparisons of the strategy groups are presented in Table 8. Significant differences on total course score were present between *Intensive – High Engagement* and *Highly Selective – Low Engagement* and between *Strategic – Moderate Engagement* and *Highly Selective – Low Engagement*.

Python. The performance indicator in the Introduction to Python course was the final course mark. The median, first and third quartile of the performance indicators for each learning strategy group are presented in Table 7. The *Inactive* and *Highly active at the beginning* strategy groups showed relatively low average course marks (median = 29.3 and 20.4, respectively). The *Highly Active* group achieved high average course marks (median = 82.7), well above the other two strategy groups. The Kruskal-Wallis test confirmed the statistical difference among the strategy groups with respect to the course grade. The pairwise comparisons of the strategy groups are presented in Table 8. Significant differences were observed between *Strategy 2: Highly Active at the Beginning* and *Strategy 3: Highly Active* and between *Strategy 1: Inactive* and *Strategy 3: Highly Active*.

5. Discussion

5.1 RQ1: Tactics and Strategies Detection

5.1.1 The detected tactics and the differences in learning contexts

The results indicate that the course design shaped the selection of learning activities and consequently, led to the adoption of diverse learning tactics (Winne & Hadwin, 1998). For instance, in the Computer Engineering course, a variety of learning activities were available including video lectures, quizzes, practical exercises, and reading materials (Table 3). Even though the dominant learning actions were those related to the practical exercises (42 %), other learning activities shared a relatively high frequency of use (more than 10 % for each activity). The adopted analytic method based on process mining led to the detection of distinct clusters of students' learning sequences (i.e. completed learning sessions), characterised by different dominant actions and different composition of actions in learning sequences (Figure 2 and Tables 4). As such, these clusters are representative of the patterns in students' learning behaviour and indicative of the learning tactics that students used.

Meanwhile, the design of the Introduction to Python course placed primary focus on practice through coding exercises (Table 3). This focus was reflected in the detected learning tactics. In particular, apart from one (*Lecture-based*) tactic, we identified three distinct tactics associated with the practical exercises. For instance, the students interacted with the practical exercises either by continuously working on a problem for a long period of time (i.e. the *Long-Practice* oriented tactic) or by breaking the practical exercises into smaller sessions (i.e., the *Short-Practice* oriented tactic) or by combining the practical exercises with other learning activities in a long session (i.e., the *Diverse-Practical Exam* oriented tactic).

In contrast to the two problem-solving based courses, the online learning activities of the Biology course were designed to support self-revision through reading materials, discussion, and self-preparation of the laboratory activities. The actions in the Biology course contained more than 70% of navigation to the course homepage, which consisted of the general information, structure, and objectives of the course, and the reading materials (28 % accessing the homepage and 44 % accessing the reading materials). Therefore, the three detected learning tactics contained reading activities with different characteristics. One tactic was characterised by actions related to the reading materials and homepage. Another tactic was about reading and pre-laboratory related activities. The last tactic focused on activities related to reading and discussion. Common to all tactics were short learning sessions.

Even though the tactics detected across the three learning contexts were different, the proposed analytic method could be applied to detect learning patterns indicative of learning tactics in these contexts. Moreover, by replicating the method, we observed that the tactics detected in different learning conditions reflected the design of the course. In particular, we recognised two features of the tactics that defined them. The first feature is the type of the learning actions within a tactic. These learning actions reflected how a course was designed. The type of learning action defined the main focus of students in each learning tactic (e.g. focusing on the reading material in the *Reading* oriented tactic). The second feature is the length of sessions in which the tactics were applied. This reflects the technique student used to direct the learning. For instance, in the Python course, students interacted with the practical exercises by either dividing the tasks into *Short* practical exercises or interacted with the tasks for a *Long* practical session. Hence, the pattern of the detected learning tactics reflected the design of the course. This notion is well confirmed by SRL theory that emphasizes the role of instructional and course design on the selection of learning tactics (Winne & Hadwin, 1998).

5.1.2 The strategies detected across the different learning contexts

The detected learning strategies are aligned with the well-defined approaches to learning as described by Entwistle (1991), Marton and Säljö (1976), and Biggs (1987). Approaches to learning are situation, content, and intuition dependent. In other words, the goal and motivation of individual students (intuition), the way in which learning was carried out (the situation), and, the content that students needed to learn (the content) play highly important roles in the selection of the approach to learning (Entwistle, 2007). This notion is well-aligned with SRL theory that emphasizes the influence of task and cognitive conditions on the selection of learning strategies (Winne & Hadwin, 1998). Entwistle (1991) and Biggs (1987) defined three approaches to learning, extended from the initial concept of approaches to learning proposed by Marton and Säljö (1976), namely, surface, deep, and strategic approaches to learning.

- **Surface approach to learning:** This approach to learning is characterized by a superficial method of learning where the focus is mainly on assessment, with the lack of understanding of content (Entwistle, 1991). Students use shortcuts to reduce the time they should devote to the learning activities.

  The students who used the *Highly Selective – Low Engagement* strategy from the Computer Engineering course, and the *Inactive* strategy group from the Python course are good representatives of this approach to learning. The low level of engagement in each week (i.e. in relation to each topic, in case of the Python course) demonstrates the lack of intention and effort towards accomplishing a higher order of thinking. The assessment tactic is the only consistently used tactic of these two strategy groups.

  The surface learner is also characterized by focusing on a portion of learning facts and often jumping to conclusions as demonstrated by the use of a single learning tactic (*Reading oriented*) and directly jumping to the assessment tasks in the *Highly Selective – Low Engagement* group during the first half of the semester of the Computer Engineering course (Chonkar et al., 2018; Jovanovic et al., 2017). The students interacted with the basic learning concept activities slightly. Moreover, the students barely realized the value of the ungraded formative assessment (i.e. quizzes). Even though a consistent use of practical exercises (i.e. the adoption of the practice-oriented technique) is highlighted by many researchers (e.g., Bjork, Dunlosky, and Kornell (2013); Dunlosky, Rawson, Marsh, Nathan, and Willingham (2013)) as effective, when superficially applied, without proper understanding of the basic concepts, it cannot reach its potential effectiveness. Therefore, the students in the *Highly Selective – Low Engagement* strategy group did not fully utilise the learning opportunities that were offered to them.

  In contrast to the two courses focused on problem-solving activities, the Biology course offered a few online practical exercises to the students. Our process mining approach detected one type of learning strategy in this course that represented surface engagement namely the *Highly Selective – Low Engagement* strategy group. The students in this strategy group were active at the beginning of the course, but their level of engagement dropped significantly after the forth week. The exams proved to be the only motivating element, as only the exam periods (weeks 7 and 13) were associated with higher levels of engagement. This behavioral pattern also reflected the students' focus on assessment.

- **Deep approach to learning:** This approach to learning, characterised by active engagement in various learning activities, is considered desirable (Marton & Säljö, 1976) since a number of research studies have reported a positive impact of deep learning approaches on academic performance (Chonkar et al., 2018). In our study, students who used the *Intensive -– High Engagement* strategy from the Computer Engineering course, the *Intensive* strategy from the Biology course, and the *Highly Active* strategy from the Python course reflect the use of the deep approach to learning. The high level of engagement of these strategy groups indicates an active concentration and high amount of effort the students exerted. Moreover, the students used a diverse set of learning tactics, and completed various learning activities, including multiple interactions with practical exercises. The performance of students who employed these strategies tended to be the highest among the three groups of strategies identified in each of the three contexts. This finding is consistent with previous research results suggesting that the deep approach to learning is associated with a high level of performance (Marton & Säljö, 1976; Zeegers, 2001).

- **Strategic approach to learning:** Strategic learning is also referred to as the achieving learning approach (Biggs, 1987). It is the study approach of choice for those students whose primary intention is to achieve high performance (Diseth & Martinsen, 2003). Similar to surface learning, the focus is on assessment, but a considerable amount of effort is also put into understanding the topic under study. Therefore, strategic learners combine surface and deep approaches to learning (Chonkar et al., 2018), and they often do well with time and study management. Among the detected learning strategy groups, *Strategic – Moderate Engagement* from the Computer Engineering course and the Biology course are clear representatives of the strategic approach to learning. The students in these groups put in a moderate amount of effort as compared to the other two groups i.e. *Intensive – High Engagement* and *Highly Selective – Low Engagement*; yet, they achieved a high performance level. The strategic students from the Computer Engineering course were consistently concentrated on the assessment activities, as evident in the extensive use of the *Exercise-Oriented* tactic. In the Biology course, the preparation of the pre-laboratory activities was observed for the strategic students mainly during the weeks when these preparations were a requirement. This demonstrates a strategic planning of the time to devote to completing the course requirements. During weeks 7 and 13, when the exams were scheduled, a higher level of engagement than in other weeks was observed for these students. This is a further confirmation of the strategic approach of this strategy group.

Unlike the flipped classroom and blended learning modalities, in the Python course delivered through a MOOC, we detected the *Highly Active at the Beginning* strategy group. The students strategically relied on a select subset of learning tactics at the beginning of the course (first two course topics); that is, the students heavily employed the *Lecture-Based* and *Diverse – Practice* oriented tactics. The tactics used by this particular strategy group suggest that the students might not have been familiar with or had a minimal background in programming. Contrary to the nature of problem-solving based learning, the students employed the *Lecture-based* tactic in several learning sessions. Being based on re-watching the video lectures, this tactic can be considered a passive and less effective learning technique (Dunlosky et al., 2013). Due to the ineffective learning tactic used, it is reasonable to believe that the students in this strategy group faced some difficulties in understanding the course content. Despite a few attempts to deal with the practical exercises, the level of engagement consistently dropped from the third topic onward. The deep approach to learning applied when interacted with the first two course topic was then replaced by the surface approach to learning.

A possible explanation of this finding is that students with background in different subject domains and lack of prior knowlege of the topics under study may opt for a particular strategy they are familiar with (as reflected by the choices of their tactics). However, the selected strategy may not be suitable for the discipline and/or topic in the given course. This is particularly emphasized in the success of learners when they take non-for-credit courses (MOOC) where the stakes of investment are lower and the backgrounds of learners much more diverse than in for-credit courses.

Overall, we observed that in the learning contexts where online learning was accompanied by the face-to-face activities (i.e. blended learning and flipped classroom), similar learning strategies were observed. That is, students consistently exhibited either deep, surface or strategic approach to learning throughout the course. However, in the MOOC, we detected two learning strategies representing the deep and surface approaches to learning and one learning strategy exhibiting a transition from deep to surface approach, i.e. the *Highly Active at the Beginning* strategy group.

5.2 RQ2: Association between Learning Strategies and Academic Performance

This study found that students who used a variety of learning tactics tended to have higher performance than those who relied on a single tactic. Some of the employed tactics were general, passive, and less effective such as re-reading (Dunlosky, 2013) or re-watching lecture videos (Dunlosky et al., 2013), but some were course-specific and more effective such as practical testing as demonstrated by the frequent use of exercise-oriented tactic and MCQ oriented tactic. Similarly, Fincham et al. (2018); Gašević, Jovanović, Pardo, and Dawson (2017); Nandagopal and Ericsson (2012) found that the students who engaged in various learning tactics had a tendency to perform better. High achieving learners know when, where, and how to apply learning tactics and strategies (Pressley, Borkowski, & Schneider, 1987).

Referring again to the students' approaches to learning, the deep and strategic approaches to learning are the desirable learning approaches, associated with higher performance (Chonkar et al., 2018; Mattick, Dennis, & Bligh, 2004), whereas the surface approach to learning is less desirable and associated with low performance. Consistent with the literature, the current study found that students who used the *Highly Selective — Low Engagement* strategy from the Computer Engineering course, the *Inactive* from the Python course, and the *Highly Selective — Low Engagement* from the Biology course, thus reflecting the surface approach to learning, performed poorer than those who used one of the other two learning strategies (Byrne, Flood, & Willis, 2010). We also observed that the students in the *Highly active at the beginning* strategy group from the Python course showed relatively low level of academic performance. The academic performance was mostly obtained from the first two course topics where they were deeply engaged with the learning content. Later, students exhibited the surface learning when interacted with the rest of the course topics (i.e. from topic 3-6).

We observed no significant association in terms of pairwise comparison between the *Strategic -– Moderate Engagement* strategy and the *Intensive -– High Engagement* strategy in the Computer Engineering course and in the Biology course. This corresponds to the use of strategic and deep approaches to learning, respectively. Considering that these two approaches to learning are known to be associated with high academic performance (Byrne et al., 2010; Zeegers, 2001), it was not surprising that we found no difference in academic performance of these two groups. The amount of effort exerted by the *Strategic -– Moderate Engagement* group declined after the midterm test. Prior research has denoted the role of engagement as one of the success factors in learning (Dabbagh, 2007). Hence, our finding that the *Intensive — High Engagement* strategy group performed better on the final exam than the *Strategic — Moderate Engagement* group corroborates the findings of several empirical studies that demonstrated that a higher level of engagement showed a significant association with course achievement (Lee, 2014; Fincham et al., 2018). Effective learners are those who did not only choose the effective learning strategies but have also realised that learning requires effort (Pressley et al., 1987; Winne, 2013).

To sum up, our exploration across three different learning contexts revealed that the higher level of engagement and application of diverse learning tactics were associated with the higher academic performance. The students who employed the learning strategies that were reflective of the deep or strategic approaches to learning obtained higher learning performance as compared to those who applied the surface approach to learning. These findings are well aligned with several empirical studies on the approaches to learning and association with the academic performance (Dabbagh, 2007; Byrne et al., 2010; Zeegers, 2001; Chonkar et al., 2018; Mattick et al., 2004).

6. Conclusions

6.1 Summary of the Contributions

This study aimed to explore the dynamics of the students' learning tactics across different course designs that were delivered by using different modalities. The main contributions of the study include, first, an analytic method for discovering latent patterns in students' learning behaviour, which are reflective of the students' learning tactics and strategies. As latent constructs, learning tactics and strategies cannot be directly observed in the collected trace data, but need to be extracted using appropriate statistical or machine learning algorithms. We have detected tactics at the level of learning sessions, extracted learning strategies from the patterns of use of the identified tactics, and interpreted the meaning of the detected learning tactics and strategies by considering their temporal and sequential attributes from the perspective of the relevant learning theories (SRL and approaches to learning).

Secondly, we validated the proposed method across several learning contexts in terms of the dependency of the detected tactics and strategies on the mode of the course delivery and the course design. The proposed method proved to provide detailed insights into how student complete online learning activities across three different learning contexts, i.e., blended learning, flipped classroom, and MOOC. The detected tactics and to a lesser extent strategies are context dependent. That is, the specific learning tactics and strategies have to be interpreted in the particular learning context the trace data originated from (Gašević et al., 2016; Gašević, Dawson, & Siemens, 2015; Winne, 2013). The interpretation should consider both the chronological ordering as well as the temporal dimension as learning tactics and strategies are "dynamic constructs" and change over time. We found that the detected tactics were likely to depend on the design of the learning tasks and activities of the courses. That is, the detected tactics represent the techniques used by students to accomplish the learning tasks. For instance, the design of the Python course accentuated practical exercises, and different manners of students' interactions with the practical exercises were

JOURNAL OF LEARNING ANALYTICS

reflected in the detected learning tactics. The Biology course encouraged students to review the reading materials; accordingly, we detected different reading behaviors. The design of the Computer Engineering course included a variety of learning tasks; our method detected distinctive tactics, each tactic representing one dominant learning task. Meanwhile, the detected strategies were less sensitive to the course design, that is, we detected a rather consistent pattern of tactics use across the courses. Across different learning modalities, we detected the behaviors of applying surface, deep, and strategic approaches to learning.

6.2 Limitations

This research provides some novel insight in the field. It showed promising results for tactic and strategy detection from trace data and their associations with academic performance across three different learning contexts. However, the study also has certain limitations. First, the study was based on the data about online activities, which means students' engagement in face-to-face sessions or individual offline learning was not analyzed. Even though online activities play an important role in supporting the development of conceptual understanding and aid the in-class activities, in-class and offline learning activities contribute to deepening of students' understanding and consequently affect their course performance, especially in the case of blended learning and flipped classroom context. Furthermore, the students' demographics and previous knowledge were not considered due to the terms and conditions imposed by the institutional ethics approvals.

The methods used to detect learning tactics and strategies (EM and hierarchical clustering) belong to the group of the unsupervised machine learning methods, which, by their nature, introduce a certain level of subjectivity. The validity of the analytics method is explored in terms of how the results is supported by the learning contexts and educational theory as suggested by several scholars (Gašević et al., 2015; Joksimovic, Kovanovic, & Dawson, 2019; Matcha, Ahmad Uzir, et al., 2019). Therefore, the explanation of the identified tactics and strategies is subjected to the theory as to how tactics and strategies are composed (Maldonado-Mahauad et al., 2018; Pardo et al., 2018). However, the use of self-reports (Zhou & Winne, 2012; Gašević, Jovanović, et al., 2017) or multimodal techniques to capture the students' motivation and goals could help in further validating the study results.

In spite of the stated limitations, our findings are consistent with the literature. The detected learning tactics and strategies are meaningful in the learning context and in accordance with learning theories (Biggs, 1987; Entwistle, 1991), which demonstrates that the findings warrant some merit and can inform future research.

6.3 Implications

To answer the call for generalisation of learning analytics methods as highlighted by Baker (2019), we have demonstrated the use of an analytics method in detection of learning tactics and strategies across three different learning contexts. This study provided a robust evidence for the cross-context applicability of the proposed method. The implications derived from this study are:

Implications for research: Future direction of the research in learning strategies detection should incorporate the offline data to better understand the learning strategies used by students, especially, in the flipped classroom and blended learning contexts. For instance, multimodal learning analytics could be used to triangulate the online and offline data. How offline learning tactics and strategies complement and/or support the online learning tactics and strategies is under-explored.

Furthermore, the replication of the analytics based method across different learning contexts has provided important insights in how students' learning behaviour – i.e. their learning tactics and strategies – tend to be shaped by the course design and the delivery modalities. Hence, through replication, we were able to come up with new insights into the learning process. This result further highlights the call for replication studies in Learning analytics research, as suggested by Baker (2019).

Implications for teachers: Research highlights the importance of learning strategies (Winne & Hadwin, 1998; Lust et al., 2013a). Nonetheless, students rarely received advice on how to choose learning strategies and apply them effectively (Matcha, Ahmad Uzir, et al., 2019). The proposed method could be used to bridge this gap. Our method for detecting learning tactics and strategies could help teachers to get insights into the students' learning behaviors. Hence, necessary support could be provided to those who might need it. For instance, advice could be provided to students on how to apply effective learning tactics and strategies by using customised feedback messages i.e. personalised feedback. (McCabe, 2011) conducted an experiment on the students' awareness of effective learning strategies. They found that the experimental group who received an instruction on effective learning strategies outperformed the group of students who did not receive any strategy-related advice. Therefore, informing students about the value of effective learning strategies can lead to an improvement in academic performance. Similarly, informing students about the benefits of effective learning tactics and strategies can increase the students' intentions to apply them (Clarebout, Elen, Collazo, Lust, & Jiang, 2013). Previous studies showed that by customising the feedback to suit the given learning situation learners could be provided with the help for improving their academic performance (Matcha, Gašević, Ahmad Uzir, Jovanović, & Pardo, 2019; Lim et al., 2019). The provision of personalised feedback, however, requires precise information of a student's learning situation. Hence, employing the proposed method to detect the tactics and strategies used by the students could help the instructors identify students in need and offer them suggestions on the effective tactics and strategies for the given learning situations. This reinforces the role of instructors in tailoring the feedback that corresponds to the requirement of students. It also facilitates the work of instructor in observing the learning progress of their students.

Furthermore, the detected learning tactics and strategies could be used to support the redesign of the course. For instance, the learning strategies showed that in certain weeks and/or course topics, students exhibited high engagement and application of diverse learning tactics. These behaviors inform the importance and/or the difficulty experienced by students in understanding the learning content. Hence, taking the students' behavior when interacting with the course content could inform the instructors on how to redesign their courses.

Acknowledgments

Dirección de Investigación de la Universidad de Cuenca (DIUC), Cuenca-Ecuador, under the project "Analítica del aprendizaje para el estudio de estrategias de aprendizaje autorregulado en un contexto de aprendizaje híbrido" (DIUC_XVIII_2019_54)

Declaration of Conflicting Interest

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

Baker, R. S. (2019). Challenges for the Future of Educational Data Mining: The Baker Learning Analytics Prizes. *Journal of Educational Data Mining*, *11*(1), 1–17.

Bernard, R. M., Borokhovski, E., Schmid, R. F., Tamim, R. M., & Abrami, P. C. (2014, 4). A meta-analysis of blended learning and technology use in higher education: from the general to the applied. *Journal of Computing in Higher Education*, *26*(1), 87–122. doi: 10.1007/s12528-013-9077-3

Biggs. (1987). *Student Approaches to Learning and Studying.* Retrieved from `https://eric.ed.gov/?id=ED308201`

Bjork, R. a., Dunlosky, J., & Kornell, N. (2013). Self-Regulated Learning: Beliefs, Techniques, and Illusions. *Annual Review of Psychology*, *64*(1), 120928131529005. doi: 10.1146/annurev-psych-113011-143823

Broadbent, J. (2017). Comparing online and blended learner's self-regulated learning strategies and academic performance. *The Internet and Higher Education*, *33*, 24–32. doi: 10.1016/j.iheduc.2017.01.004

Broadbent, J., & Poon, W. L. (2015). Self-regulated learning strategies & academic achievement in online higher education learning environments: A systematic review. *Internet and Higher Education*, *27*, 1–13. doi: 10.1016/j.iheduc.2015.04 .007

Byrne, M., Flood, B., & Willis, P. (2010). The relationship between learning approaches and learning outcomes : a study of Irish accounting students The relationship between learning approaches and learning outcomes : a study of Irish. , *9284*. doi: 10.1080/0963928021015325

Chonkar, S. P., Ha, T. C., Chu, S. S. H., Ng, A. X., Lim, M. L. S., Ee, T. X., ... Tan, K. H. (2018). The predominant learning approaches of medical students. *BMC Medical Education*, *18*(1), 1–8. doi: 10.1186/s12909-018-1122-5

Clarebout, G., Elen, J., Collazo, N. A. J., Lust, G., & Jiang, L. (2013, 1). Metacognition and the Use of Tools. In R. Azevedo & V. Aleven (Eds.), *International handbook of metacognition and learning technologies* (pp. 187–195). Springer New York.

Dabbagh, N. (2007). The online learner: Characteristics and pedagogical implications. *Contemporary Issues in Technology and Teacher Education*, *7*(3), 217–226. Retrieved from `http://go.editlib.org/p/22904`

Derry, S. J. (1989). Putting learning strategies to work. *Educational Leadership*, *47*(5), 4–10.

DiFrancesca, D., Nietfeld, J. L., & Cao, L. (2016). A comparison of high and low achieving students on self-regulated learning variables. *Learning and Individual Differences*, *45*, 228–236. doi: 10.1016/j.lindif.2015.11.010

Diseth, A., & Martinsen, (2003). Approaches to Learning, Cognitive Style, and Motives as Predictors of Academic Achievement. *Educational Psychology*, *23*(2), 195–207. doi: 10.1080/01443410303225

Dunlosky, J. (2013). Strengthening the Student Toolbox. *American Educator*, *37*(3), 12–21. Retrieved from `http://www.aft.org/sites/default/files/periodicals/dunlosky.pdf`

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013, 1). Improving Students' Learning With Effective Learning Techniques Promising Directions From Cognitive and Educational Psychology. *Psychological Science in the Public Interest*, *14*(1), 4–58. doi: 10.1177/1529100612453266

Entwistle, N. J. (1991). Approaches to Learning and Perceptions of the Learning Environment : Introduction to the Special Issue. *Higher Eduation*, *22*(3), 201–204. doi: 10.1007/BF00132287

Entwistle, N. J. (2007). Research into student learning and university teaching. *The British Psychological Society*(October), 1–18. doi: 10.1348/000709906X166772

Eriksson, T., Adawi, T., & Stöhr, C. (2017, 4). "Time is the bottleneck": a qualitative study exploring why learners drop out of MOOCs. *Journal of Computing in Higher Education*, *29*(1), 133–146. doi: 10.1007/s12528-016-9127-8

Ferreira, D. R., & Gillblad, D. (2009). Discovering Process Models from Unlabelled Event Logs. *Business Process Management*, *5701*, 143–158. doi: 10.1007/978-3-642-03848-8{_}11

Fincham, O. E., Gasevic, D. V., Jovanovic, J. M., & Pardo, A. (2018). From Study Tactics to Learning Strategies: An Analytical Method for Extracting Interpretable Representations. *IEEE Transactions on Learning Technologies*, *1382*(c), 1–14. doi: 10.1109/TLT.2018.2823317

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences*, *111*(23), 8410–8415. Retrieved from `http://www.pnas.org/content/early/2014/05/08/1319030111.abstract` doi: 10.1073/pnas.1319030111

Gabadinho, A., Ritschard, G., Studer, M., & Muller, N. S. (2008). Mining sequence data in R with the TraMineR package: A user's guide. *Department of Econometrics and Laboratory of Demography, University of Geneva, Switzerland*, *1*, 1–124. Retrieved from `http://mephisto.unige.ch/traminer`

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

Gašević, D., Dawson, S., & Siemens, G. (2015). Let's not forget: Learning analytics are about learning. *TechTrends*, *59*(1). doi: 10.1007/s11528-014-0822-x

Gašević, D., Jovanović, J., Pardo, A., & Dawson, S. (2017). Detecting Learning Strategies with Analytics: Links with Self-Reported Measures and Academic Performance. *Journal of Learning Analytics*, *4*(2), 113–128. Retrieved from `http://www.learning-analytics.info/journals/index.php/JLA/article/view/5085` doi: 10.18608/jla.2017.42.10

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), 63–78. doi: 10.1080/23735082.2017.1286142

Gasevic, D., Mirriahi, N., Dawson, S., & Joksimovic, S. (2017). Effects of instructional conditions and experience on the adoption of a learning tool. *Computers in Human Behavior*, *67*, 207–220. doi: 10.1016/j.chb.2016.10.026

Gatta, R., Lenkowicz, J., Vallati, M., & Stefanini, A. (2017). *pMineR: Processes Mining in Medicine.* Retrieved from `https://cran.r-project.org/package=pMineR`

Hadwin, A. F., Nesbit, J. C., Jamieson-Noel, D., Code, J., & Winne, P. H. (2007). Examining trace data to explore self-regulated learning. *Metacognition and Learning*, *2*(2-3), 107–124. doi: 10.1007/s11409-007-9016-7

Joksimovic, S., Kovanovic, V., & Dawson, S. (2019). The Journey of Learning Analytics. *HERDSA Review of Higher Education*, *6*(January), 37–63. Retrieved from `www.herdsa.org.au/herdsa-review-higher-education-vol-6/37-63`

Jovanovic, J., Gasevic, D., Dawson, S., Pardo, A., & Mirriahi, N. (2017). Learning analytics to unveil learning strategies in a flipped classroom. *The Internet and Higher Education*, *33*, 74–85. Retrieved from `http://linkinghub.elsevier.com/retrieve/pii/S1096751617300684` doi: 10.1016/j.iheduc.2017.02.001

Kizilcec, R. F., Pérez-Sanagustín, M., & Maldonado, J. J. (2017). Self-regulated learning strategies predict learner behavior and goal attainment in Massive Open Online Courses. *Computers & Education*, *104*, 18–33. Retrieved from `http://www.sciencedirect.com/science/article/pii/S0360131516301798` doi: http://dx.doi.org/10.1016/j.compedu.2016.10.001

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

Lai, C. L., & Hwang, G. J. (2016, 9). A self-regulated flipped classroom approach to improving students' learning performance in a mathematics course. *Computers and Education*, *100*, 126–140. doi: 10.1016/j.compedu.2016.05.006

Lee, J.-s. (2014). The Relationship Between Student Engagement and Academic Performance : Is It a Myth or Reality? *The Journal of Educational Research*, *107*(3), 177– 185. doi: 10.1080/00220671.2013.807491

Lim, L. A., Gentili, S., Pardo, A., Kovanović, V., Whitelock-Wainwright, A., Gašević, D., & Dawson, S. (2019). What changes, and for whom? A study of the impact of learning analytics-based process feedback in a large course. *Learning and Instruction*. doi: 10.1016/j.learninstruc.2019.04.003

Lust, G., Elen, J., & Clarebout, G. (2013a). Regulation of tool-use within a blended course: Student differences and performance effects. *Computers and Education*, *60*(1), 385–395. Retrieved from `http://dx.doi.org/10.1016/j.compedu.2012.09.001` doi: 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. *Computer in Human Behavior*, *29*, 2013–2021. doi: 10.1016/j.chb.2013.03.014

Maldonado-Mahauad, J., Pérez-Sanagustín, M., Kizilcec, R. F., Morales, N., & Munoz-Gama, J. (2018). Mining theory-based patterns from Big data: Identifying self-regulated learning strategies in Massive Open Online Courses. *Computers in Human Behavior*, *80*, 179–196. doi: 10.1016/j.chb.2017.11.011

Malmberg, J., Järvelä, S., & Kirschner, P. A. (2014). Elementary school students' strategic learning: does task-type matter? *Metacognition and Learning*, *9*(2), 113–136. doi: 10.1007/s11409-013-9108-5

Malmberg, J., Järvenoja, H., & Järvelä, S. (2010). Tracing elementary school students' study tactic use in gStudy by examining a strategic and self-regulated learning. *Computers in Human Behavior*, *26*(5), 1034–1042. Retrieved from http://dx.doi.org/10.1016/j.chb.2010.03.004 doi: 10.1016/j.chb.2010.03.004

Marton, F., & Säljö, R. (1976). on Qualitative Differences in Learning: I-Outcome and Process*. *British Journal of Educational Psychology*, *46*(1), 4–11. Retrieved from http://doi.wiley.com/10.1111/j.2044-8279.1976.tb02980.x doi: 10.1111/j.2044-8279.1976.tb02980.x

Matcha, W., Ahmad Uzir, N., Gasevic, D., & Pardo, A. (2019). A Systematic Review of Empirical Studies on Learning Analytics Dashboards: A Self-Regulated Learning Perspective. *IEEE Transactions on Learning Technologies*, *1382*(c), 1–1. Retrieved from https://ieeexplore.ieee.org/document/8713912/ doi: 10.1109/TLT.2019.2916802

Matcha, W., Gašević, D., Ahmad Uzir, N., Jovanović, J., & Pardo, A. (2019). Analytics of Learning Strategies: Associations with Academic Performance and Feedback. In *Proceedings of the 9th international conference on learning analytics & knowledge* (pp. 461–470). Retrieved from http://doi.acm.org/10.1145/3303772.3303787 doi: 10.1145/3303772.3303787

Matcha, W., Gašević, D., Ahmad Uzir, N., Jovanović, J., Pardo, A., Maldonado-Mahauad, J., & Pérez-Sanagustín, M. (2019). Detection of Learning Strategies: A Comparison of Process, Sequence and Network Analytic Approaches. In *European conference on technology enhanced learning* (pp. 525–540). Springer. Retrieved from https://link.springer.com/chapter/10.1007/978-3-030-29736-7_39

Mattick, K., Dennis, I., & Bligh, J. (2004). Approaches to learning and studying in medical students: Validation of a revised inventory and its relation to student characteristics and performance. *Medical Education*, *38*(5), 535–543. doi: 10.1111/j.1365-2929.2004.01836.x

McCabe, J. (2011, 4). Metacognitive awareness of learning strategies in undergraduates. *Memory & Cognition*, *39*(3), 462–476. doi: 10.3758/s13421-010-0035-2

Molenaar, I. (2014). Advances in temporal analysis in learning and instruction. *Frontline Learning Research*, *6*, 15–24.

Nandagopal, K., & Ericsson, K. A. (2012). An expert performance approach to the study of individual differences in self-regulated learning activities in upper-level college students. *Learning and Individual Differences*, *22*(5), 597–609. Retrieved from http://dx.doi.org/10.1016/j.lindif.2011.11.018 doi: 10.1016/j.lindif.2011.11.018

Nugent, G., Guru, A., & Namuth-Covert, D. (2018). Students' Approaches to E-Learning: Analyzing Credit/Noncredit and High/Low Performers. *Interdisciplinary Journal of e-Skills and Lifelong Learning*, *14*, 143–158. doi: 10.28945/4133

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

Olney, T., Rienties, B., & Toetenel, L. (2018). Gathering, visualising and interpreting learning design analytics to inform classroom practice and curriculum design: a student-centred approach from the Open University. In *From data and analytics to the classroom: Translating learning analytics for teachers.* London: Routledge. Retrieved from https://www.researchgate.net/publication/328145720

Pardo, A. (2018). A feedback model for data-rich learning experiences. *Assessment & Evaluation in Higher Education*, *43*(3), 428–438. doi: 10.1080/02602938.2017.1356905

Pardo, A., Gasevic, D., Jovanovic, J. M., Dawson, S., & Mirriahi, N. (2018). Exploring Student Interactions with Preparation Activities in a Flipped Classroom Experience. *IEEE Transactions on Learning Technologies*. Retrieved from https://ieeexplore.ieee.org/document/8417959/ doi: 10.1109/TLT.2018.2858790

Pardo, A., & Mirriahi, N. (2017). Design, Deployment and Evaluation of a Flipped Learning First Year Engineering Course. In C. Reidsema, L. Kavanagh, R. Hadgraft, & N. Smith (Eds.), *The flipped classroom: Practice and practices in higher education* (pp. 177–191). Singapore: Springer.

Pérez-Sanagustín, M., Hilliger, I., Alario-Hoyos, C., Kloos, C. D., & Rayyan, S. (2017, 4). H-MOOC framework: reusing MOOCs for hybrid education. *Journal of Computing in Higher Education*, *29*(1), 47–64. doi: 10.1007/s12528-017-9133-5

Pressley, M., Borkowski, J. G., & Schneider, W. (1987). Cognitive Strategies : Good Strategy Users Coordinate Metacognition and Knowledge. *Annals of Child Development*, *4*(2), 89–129.

JOURNAL OF LEARNING ANALYTICS

Proctor, B. E., Prevatt, F. F., Adams, K. S., Reaser, A., & Petscher, Y. (2006). Study Skills Profiles of Normal-Achieving and Academically-Struggling College Students. *Journal of College Student Development*, *47*(1), 37–51.

Rachal, C. K., Daigle, S., & Rachal, W. S. (2007). Learning problems reported by college students: Are they using learning strategies? *Journal of Instructional Psychology*, *34*(4), 191–199. Retrieved from `http://content.ebscohost.com.proxy-remote.galib.uga.edu/ContentServer.asp?T=P&P=AN&K=28349624&S=R&D=slh&EbscoContent=dGJyMNXb4kSep7Y4yNfsOLCmr0qep7BSs6e4Sq6WxWXS&ContentCustomer=dGJyMPPd30m549+B7LHfhOoA`

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. doi: 10.1166/asl.2015.6514

Rodríguez, M. F., Correa, J. H., Pérez-Sanagustín, M., Pertuze, J. A., & Alario-Hoyos, C. (2017). A MOOC-based flipped class: Lessons learned from the orchestration perspective. In *Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics)* (Vol. 10254 LNCS, pp. 102–112). Springer Verlag. doi: 10.1007/978-3-319-59044-8{_}12

Tessmer, M., & Richey, R. C. (1997). The role of context in learning and instructional design. *Educational Technology Research and Development*, *45*(2), 85–115. doi: 10.1007/BF02299526

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

Winne, P. H. (2013). Learning Strategies, Study Skills, and Self-Regulated Learning in Postsecondary Education. *Higher Education: Handbook of Theory and Research*(28), 337–403. Retrieved from `http://www.springerlink.com/index/10.1007/978-90-481-8598-6` doi: 10.1007/978-90-481-8598-6

Winne, P. H. (2017). Learning Analytics for Self-Regulated Learning. *Handbook of learning analytics*, 241–249. doi: 10.18608/hla17.021

Winne, P. H., & Hadwin, A. F. (1998). Studying as Self-Regulated Learning. *Metacognition in educational theory and practice*, *93*, 277–304.

Winne, P. H., Jamieson-Noel, D., & Muis, K. (2002). *Methodological issues and advances in researching tactics, strategies, and self-regulated learning* (Vol. 12).

Yip, M. C. W. (2007). Differences in Learning and Study Strategies between High and Low Achieving University Students: A Hong Kong study. *Educational Psychology*, *27*(4), 597–606. doi: 10.1080/01443410701309126

Zeegers, P. (2001). Approaches to learning in science : A longitudinal study. *British Journal of Educational Psychology*(71), 115–132.

Zhou, M., & Winne, P. H. (2012). Modeling academic achievement by self-reported versus traced goal orientation. *Learning and Instruction*, *22*(6), 413–419. Retrieved from `http://dx.doi.org/10.1016/j.learninstruc.2012.03.004` doi: 10.1016/j.learninstruc.2012.03.004

Zhu, Y., Au, W., & Yates, G. (2016). University students' self-control and self-regulated learning in a blended course. *Internet and Higher Education*, *30*, 54–62. doi: 10.1016/j.iheduc.2016.04.001

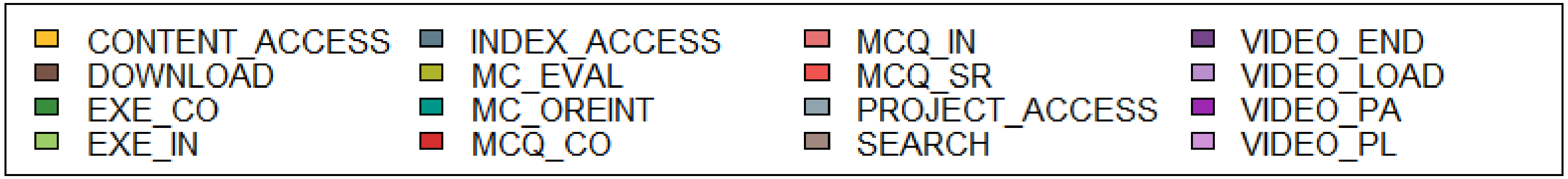
Zimmerman, B. J. (1986). Becoming a self-regulated learner: Which are the key subprocesses? *Contemporary Educational Psychology*, *11*(4), 307–313. doi: 10.1016/0361-476X(86)90027-5

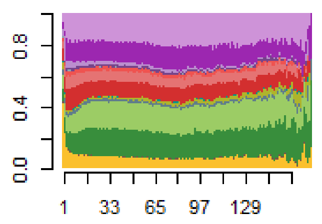
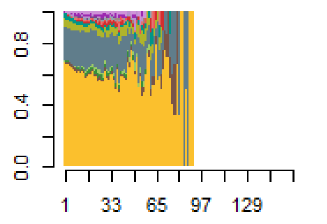
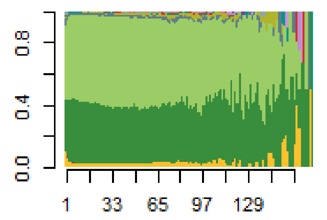
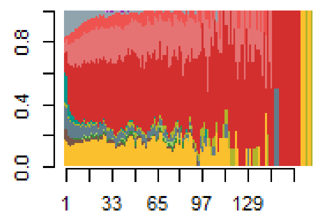
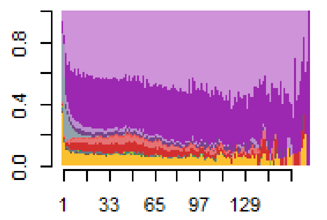
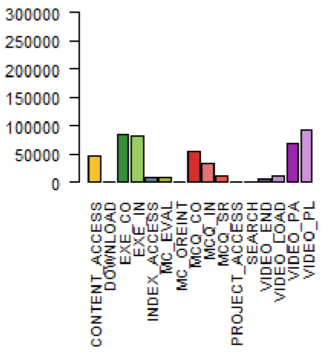
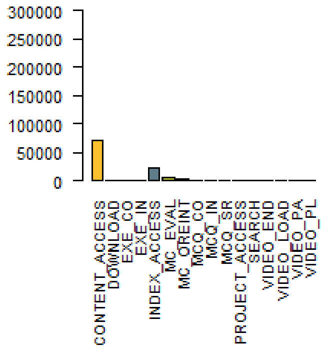
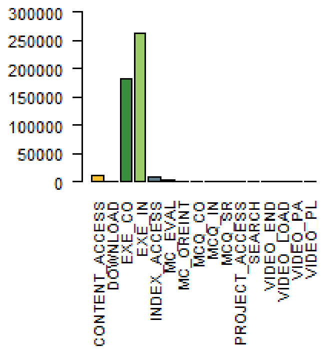
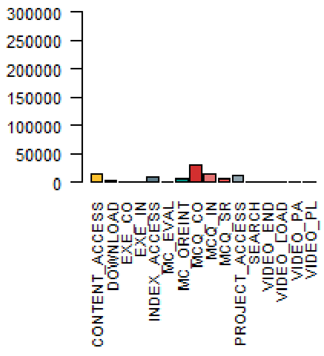
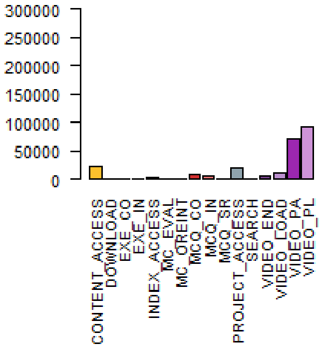
Supplement Material: Analytics of Learning Strategies: Role of Course Design and Delivery Modality

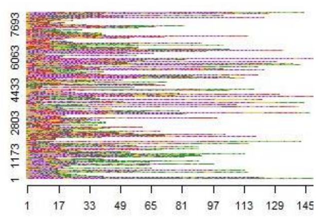
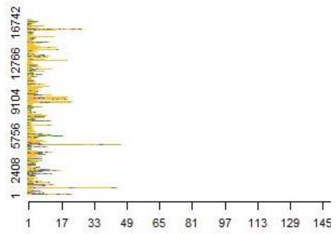
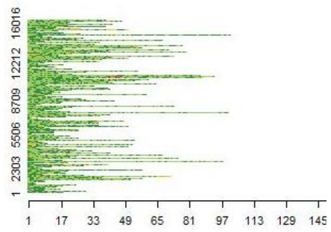
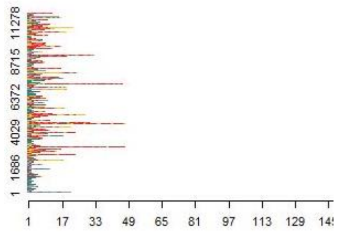
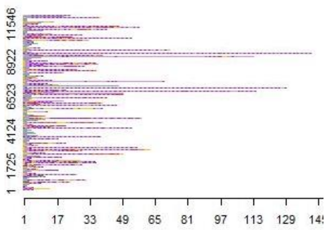
DatasetA: Computer Engineering Course

Table 1 Characteristics of the detected tactics In the Computer Engineering Course

	Diverse	**Reading oriented**	**Exercise oriented**	**MCQ oriented**	**Video oriented**
No.	N= 8288 (12.61% of all learning sessions)	N = 17024 (25.91 % of all learning sessions)	N = 16287 (24.79 % of all learning sessions)	N = 11915 (18.13 % of all learning sessions)	N = 12196 (18.56% of all learning sessions)



	Diverse	Reading oriented	Exercise oriented	MCQ oriented	Video oriented
State distribution					
Frequency of actions					

<table>
<tr><th></th><th>Diverse</th><th>Reading oriented</th><th>Exercise oriented</th><th>MCQ oriented</th><th>Video oriented</th></tr>
<tr>
<td>Order of sequence</td>
<td>
</td>
<td>
</td>
<td>
</td>
<td>
</td>
<td>
</td>
</tr>
</table>

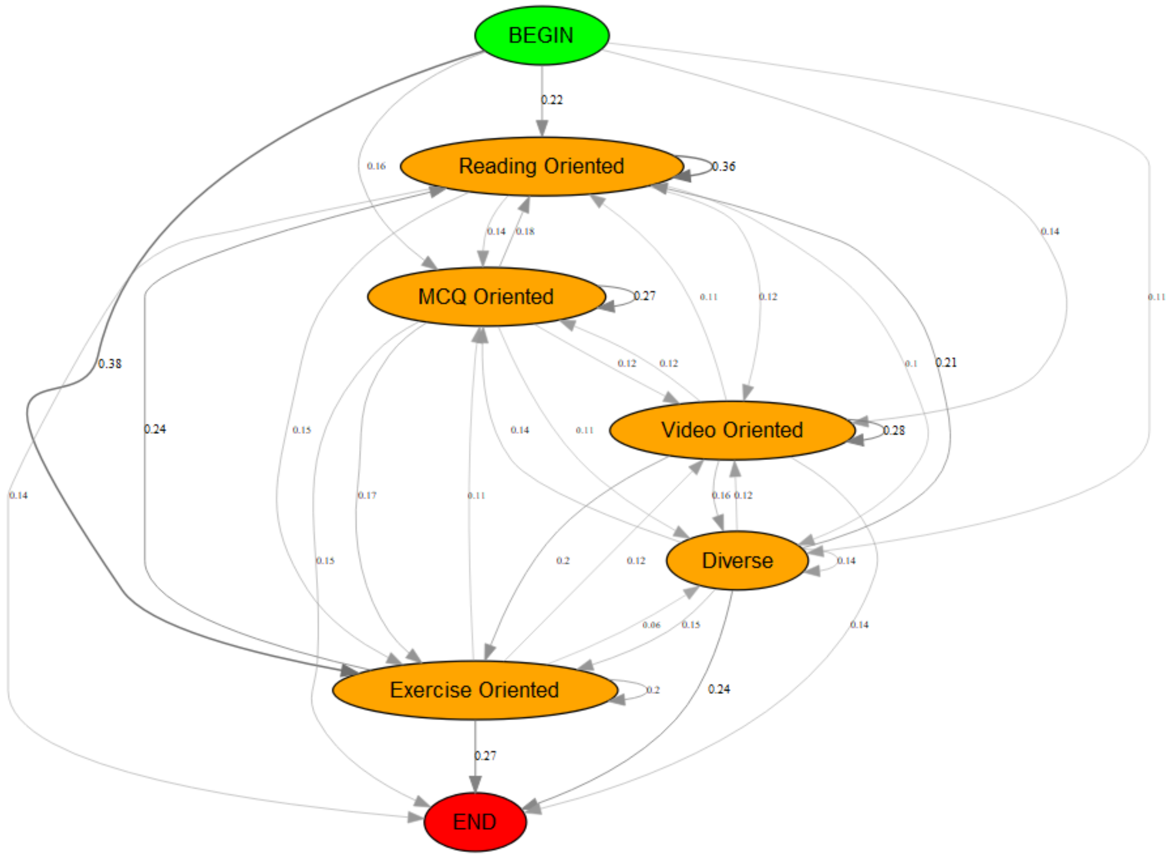


Figure 1 Process Model of Strategic – Moderate Engagement Group in Computer Engineering Course

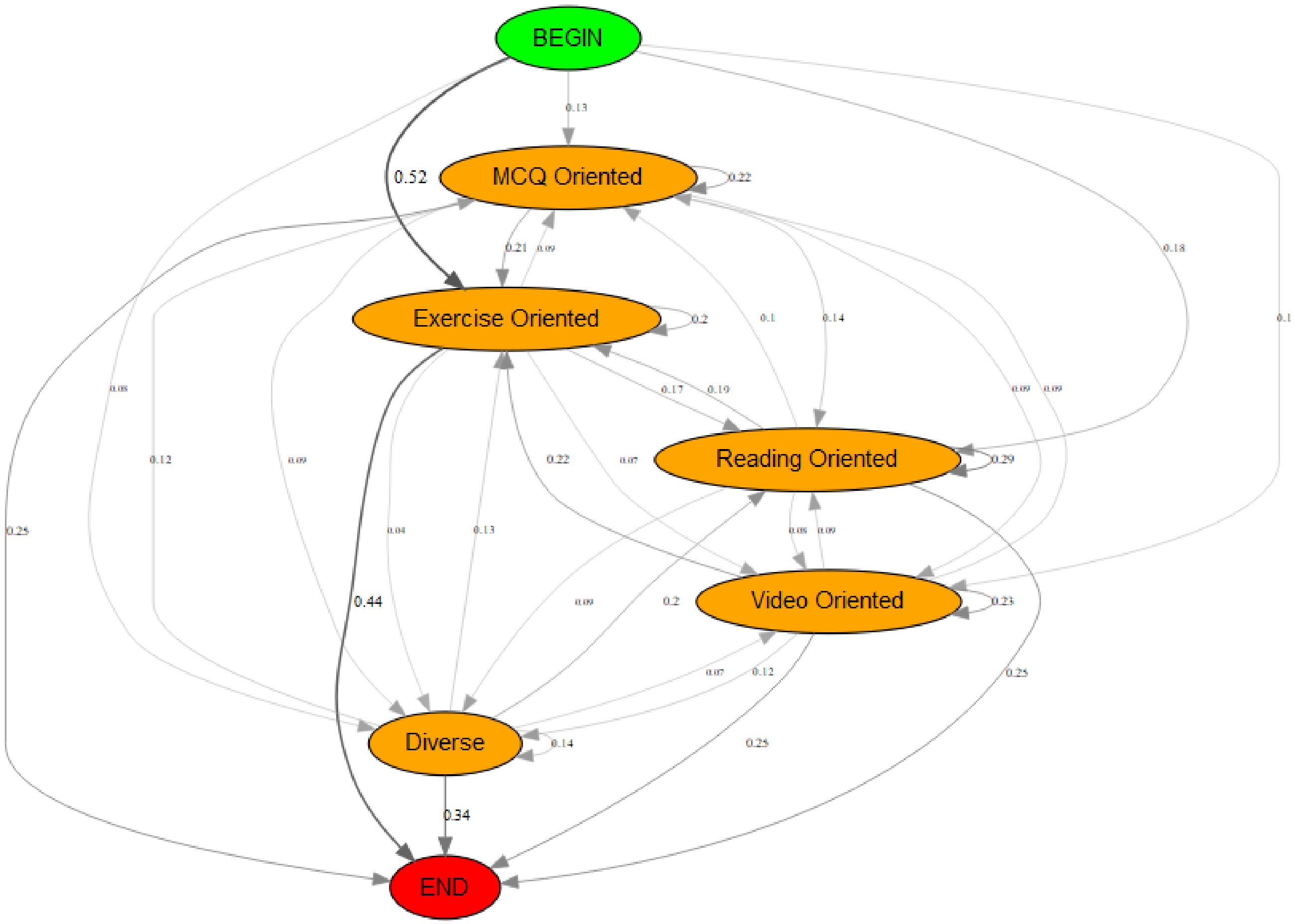


Figure 2 Process Model of Highly Selective – Low Engagement Group in Computer Engineering Course

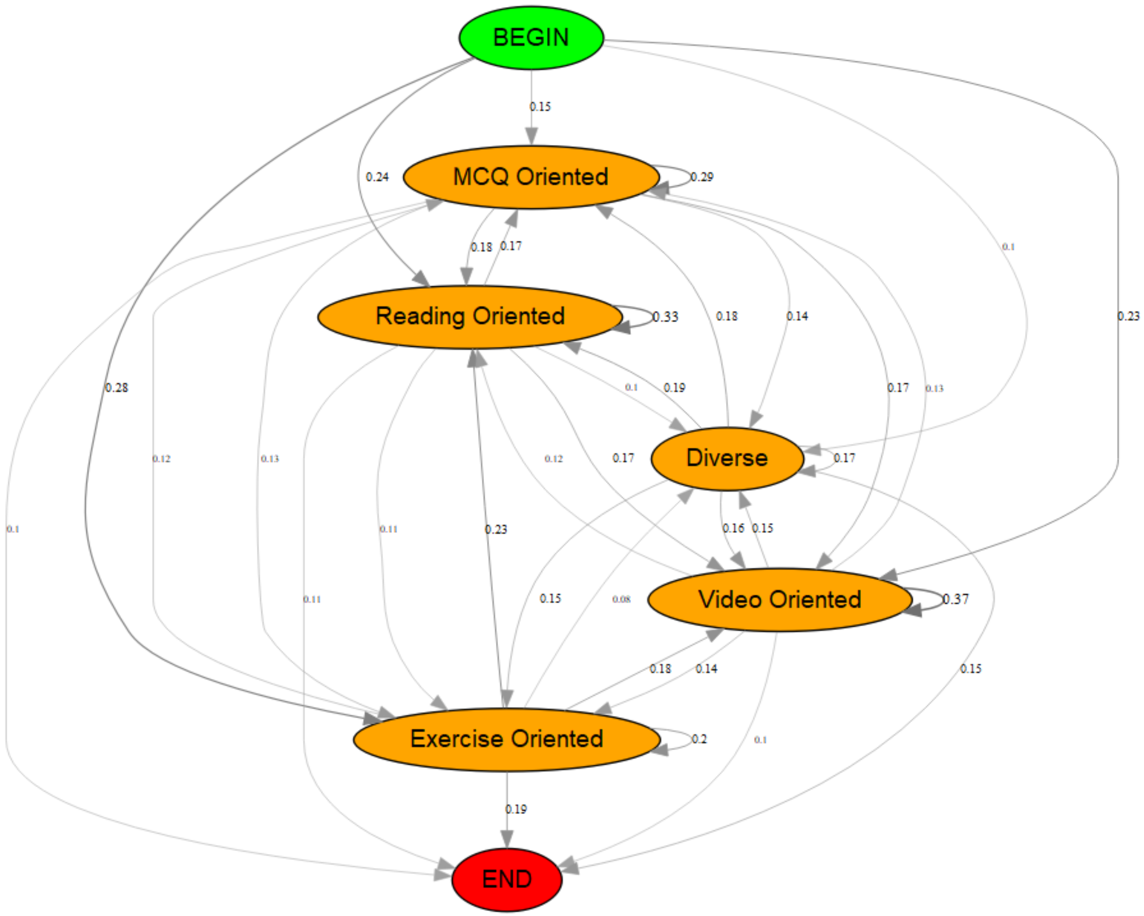
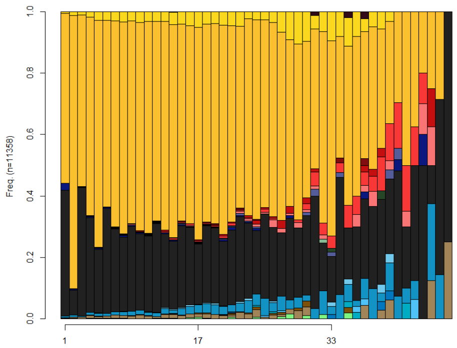
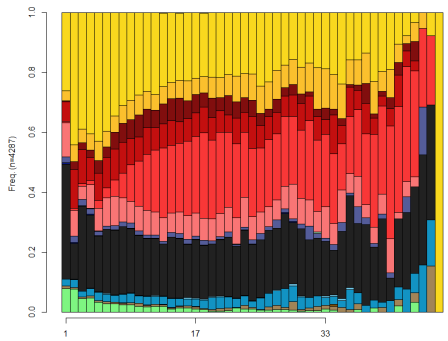
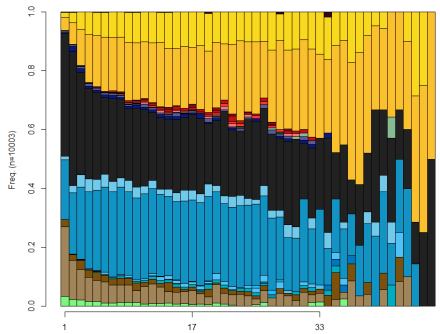
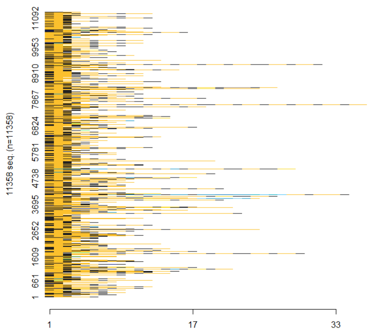
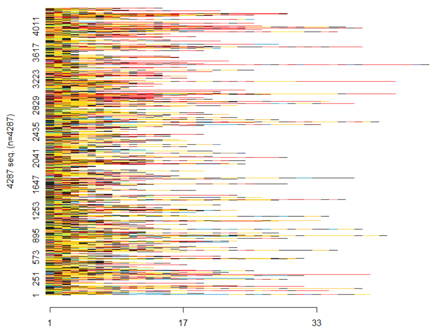
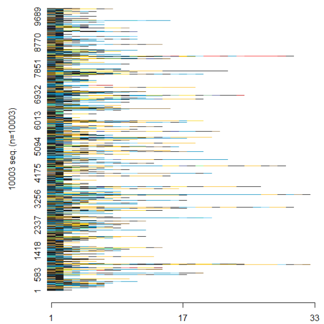


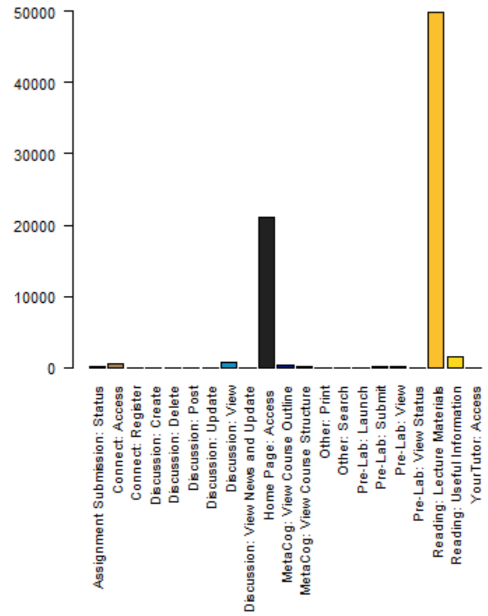
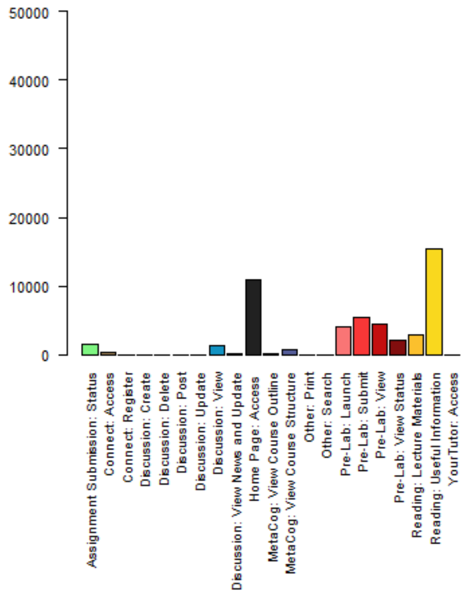
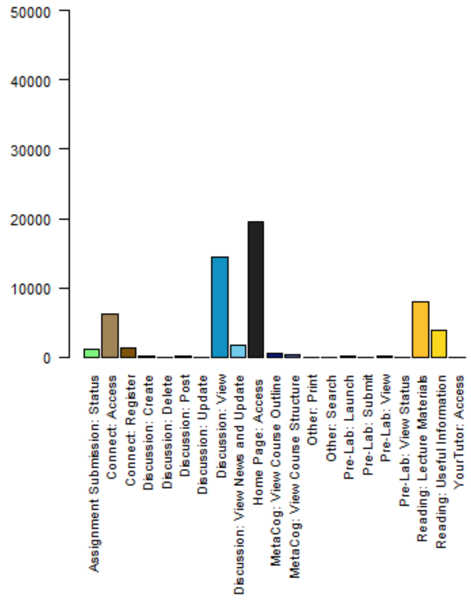
Figure 3 Process Model of Intensive – High Engagement Group in Computer Engineering Course

DatasetB: Biology course

Table 2 Characteristics of the detected tactics In the Biology Course

<table>
<tr><td></td><td>Reading</td><td>Reading and Pre-Lab</td><td>Reading and Discussion</td></tr>
<tr><td>No. Session</td><td>N= 11358
(44.28 % of all learning sessions)</td><td>N= 4287
(16.71% of all learning sessions)</td><td>N= 10003
(39.00% of all learning sessions)</td></tr>
<tr><td>Legend</td><td colspan="3">Assignment Submission: Status
Connect: Access
Connect: Register
Discussion: Create
Discussion: Delete
Discussion: Post
Discussion: Update
Discussion: View
Discussion: View News and Update
Home Page: Access
MetaCog: View Course Outline
MetaCog: View Course Structure
Other: Print
Other: Search
Pre-Lab: Launch
Pre-Lab: Submit
Pre-Lab: View
Pre-Lab: View Status
Reading: Lecture Materials
Reading: Useful Information
YourTutor: Access</td></tr>
<tr><td>State Distribution</td><td></td><td></td><td></td></tr>
</table>

	Reading	**Reading and Pre-Lab**	**Reading and Discussion**
All learning Sequences			

<table>
<tr><th></th><th>Reading</th><th>Reading and Pre-Lab</th><th>Reading and Discussion</th></tr>
<tr><td>Frequency of Actions</td>
<td>
</td>
<td>
</td>
<td>
</td></tr>
</table>

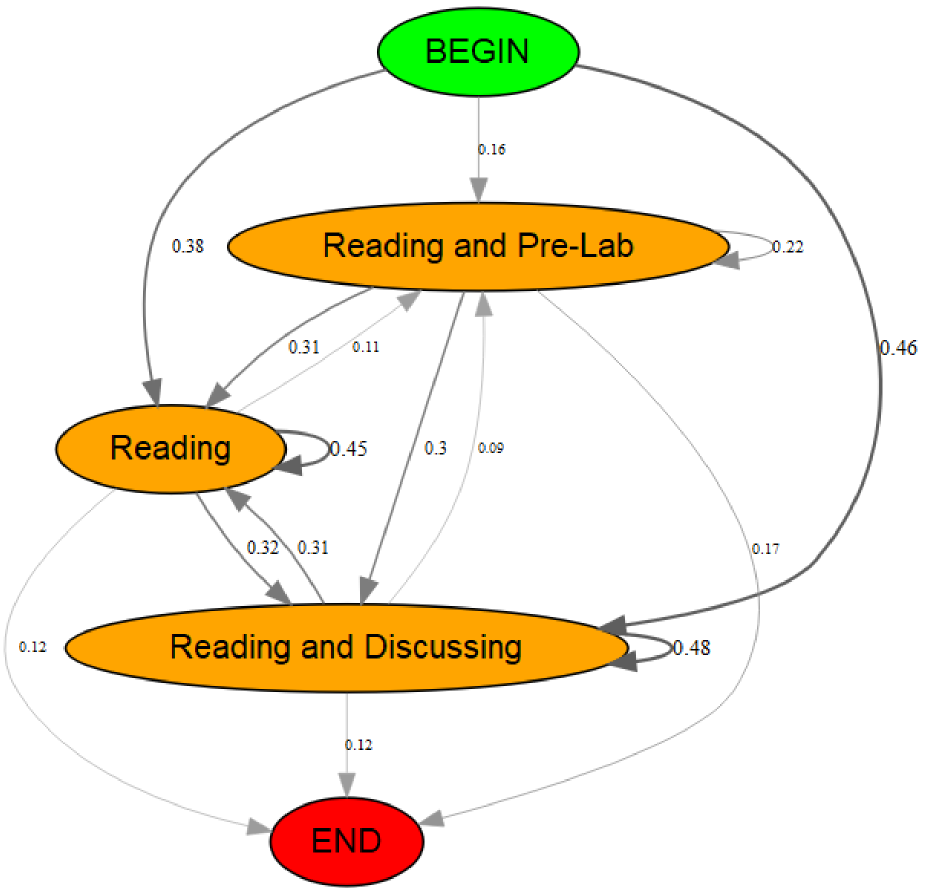


Figure 4 Process Model of Intensive – High Engagement Group in Biology Course

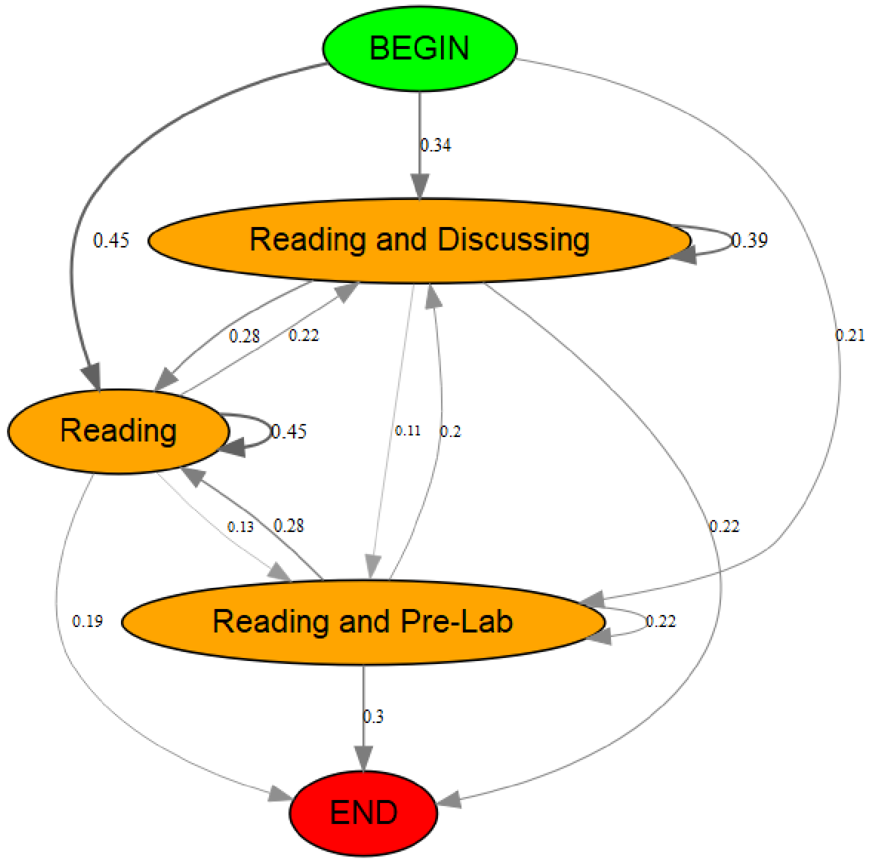


Figure 5 Process Model of Strategic – Moderate Engagement Group in Biology Course

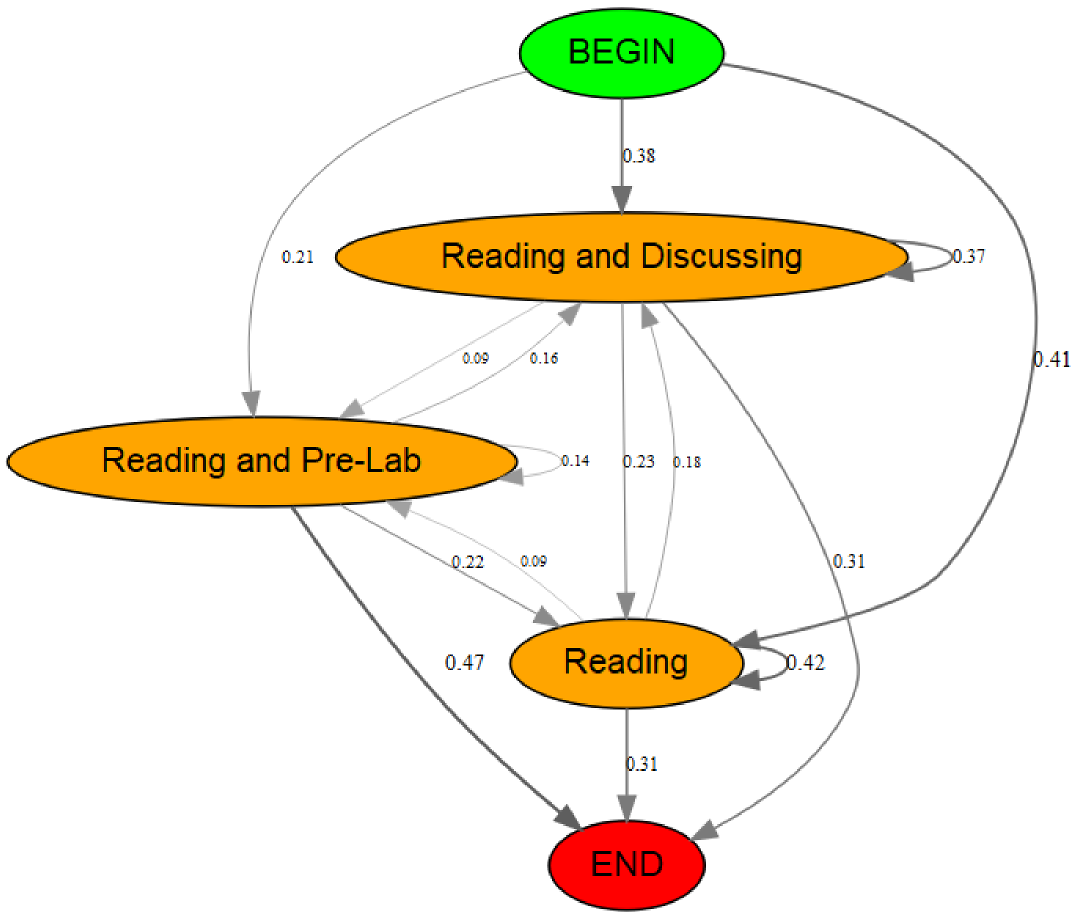
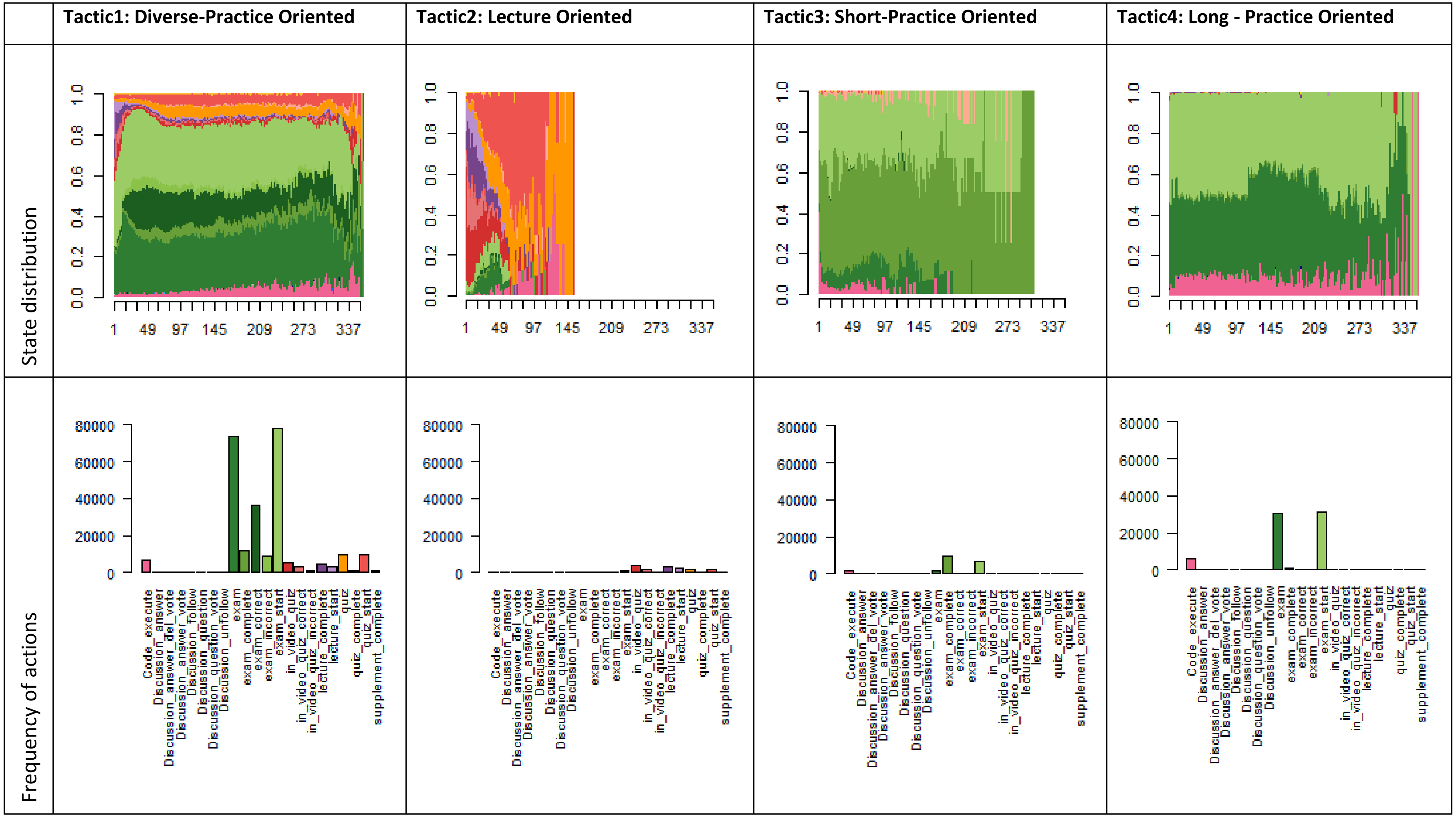


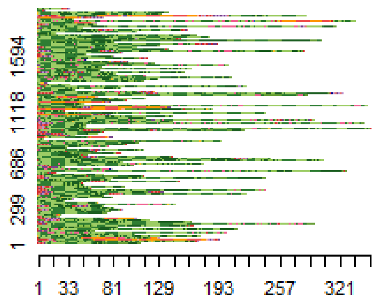
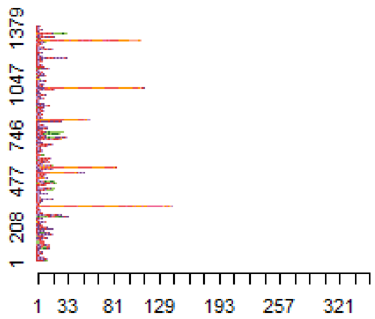
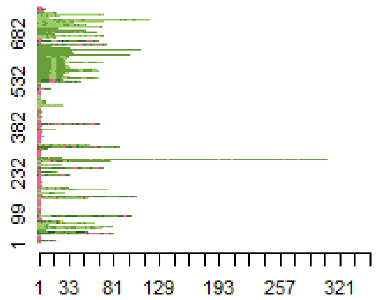
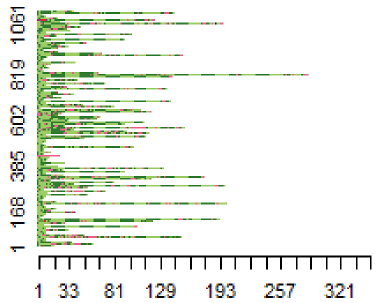
Figure 6 Process Model of Low Engagement Group in Biology Course

DatasetC: Introduction to Python

Table 3 Characteristics of the detected tactics From The Introduction to Python Course

	Tactic1: Diverse-Practice Oriented	**Tactic2: Lecture Oriented**	**Tactic3: Short-Practice Oriented**	**Tactic4: Long - Practice Oriented**
No. Session	N= 1892 Sessions (35.83 % of all learning sessions)	N = 929 Sessions (17.59 % of all learning sessions)	N= 1776 Sessions (33.63 % of all learning sessions)	N= 684 Sessions (12.95 % of all learning sessions)
Legend	Code_execute Discussion_answer Discussion_answer_del_vote Discussion_answer_vote Discussion_follow Discussion_question	Discussion_question_vote Discussion_unfollow exam exam_complete exam_correct exam_incorrect	exam_start in_video_quiz in_video_quiz_correct in_video_quiz_incorrect lecture_complete lecture_start	quiz quiz_complete quiz_start supplement_complete


Tactic1: Diverse-Practice Oriented
Tactic2: Lecture Oriented
Tactic3: Short-Practice Oriented
Tactic4: Long - Practice Oriented
State distribution
Frequency of actions
0.0
0.2
0.4
0.6
0.8
1.0
1
49
97
145
209
273
337
0
20000
40000
60000
80000
Code_execute
Discussion_answer
Discussion_answer_del_vote
Discussion_answer_vote
Discussion_follow
Discussion_question
Discussion_question_vote
Discussion_unfollow
exam
exam_complete
exam_correct
exam_incorrect
exam_start
in_video_quiz
in_video_quiz_correct
in_video_quiz_incorrect
lecture_complete
lecture_start
quiz
quiz_complete
quiz_start
supplement_complete

<table>
<tr><th></th><th>Tactic1: Diverse-Practice Oriented</th><th>Tactic2: Lecture Oriented</th><th>Tactic3: Short-Practice Oriented</th><th>Tactic4: Long - Practice Oriented</th></tr>
<tr><td>Order of sequence</td>
<td>
</td>
<td>
</td>
<td>
</td>
<td>
</td></tr>
</table>

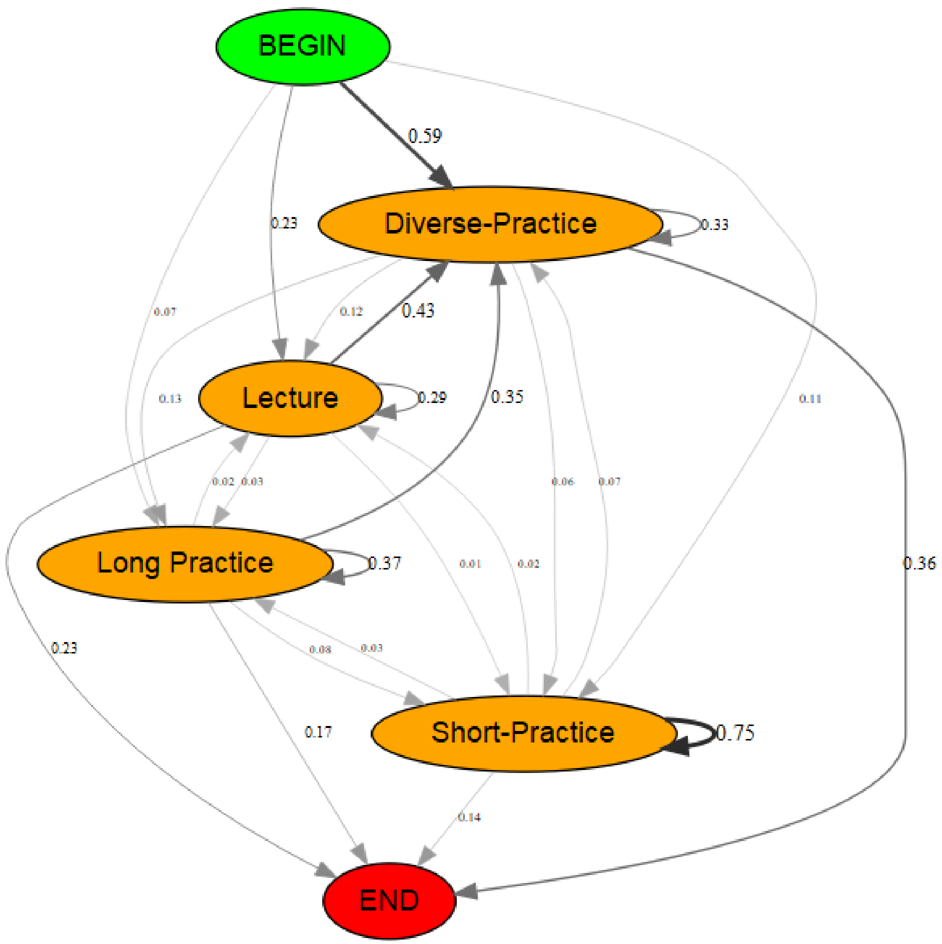


Figure 7 Process Model of Inactive Group in Introduction to Python Course

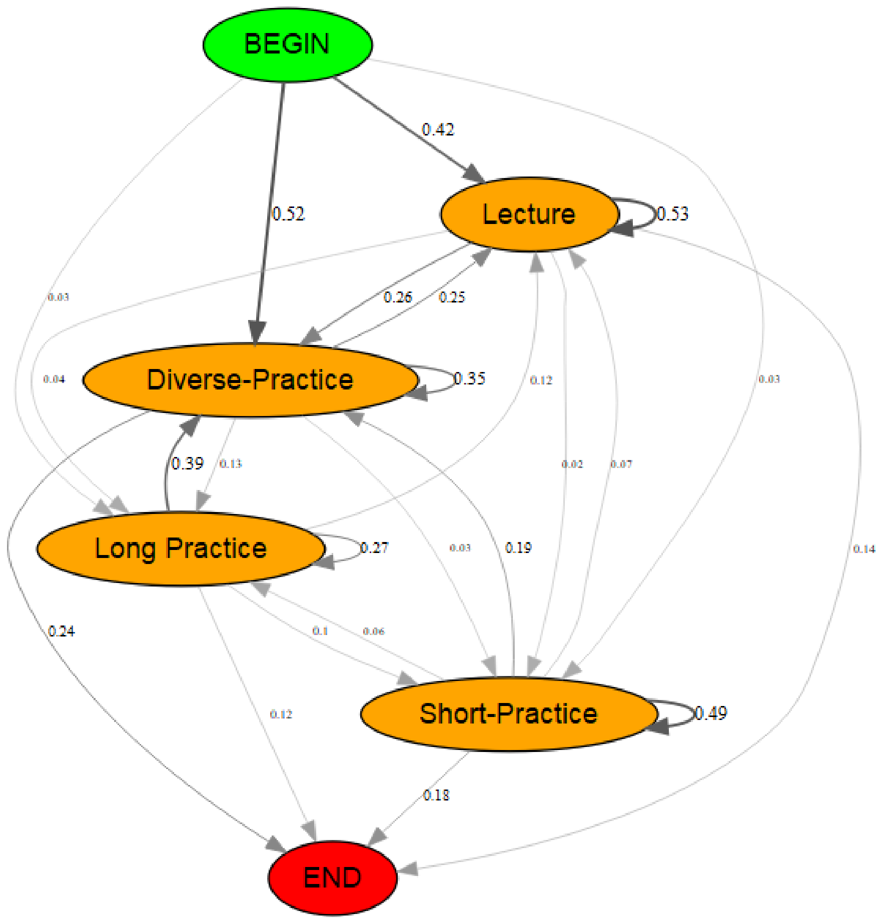


Figure 8 Process Model of Highly Active at the Beginning Group in Introduction to Python Course

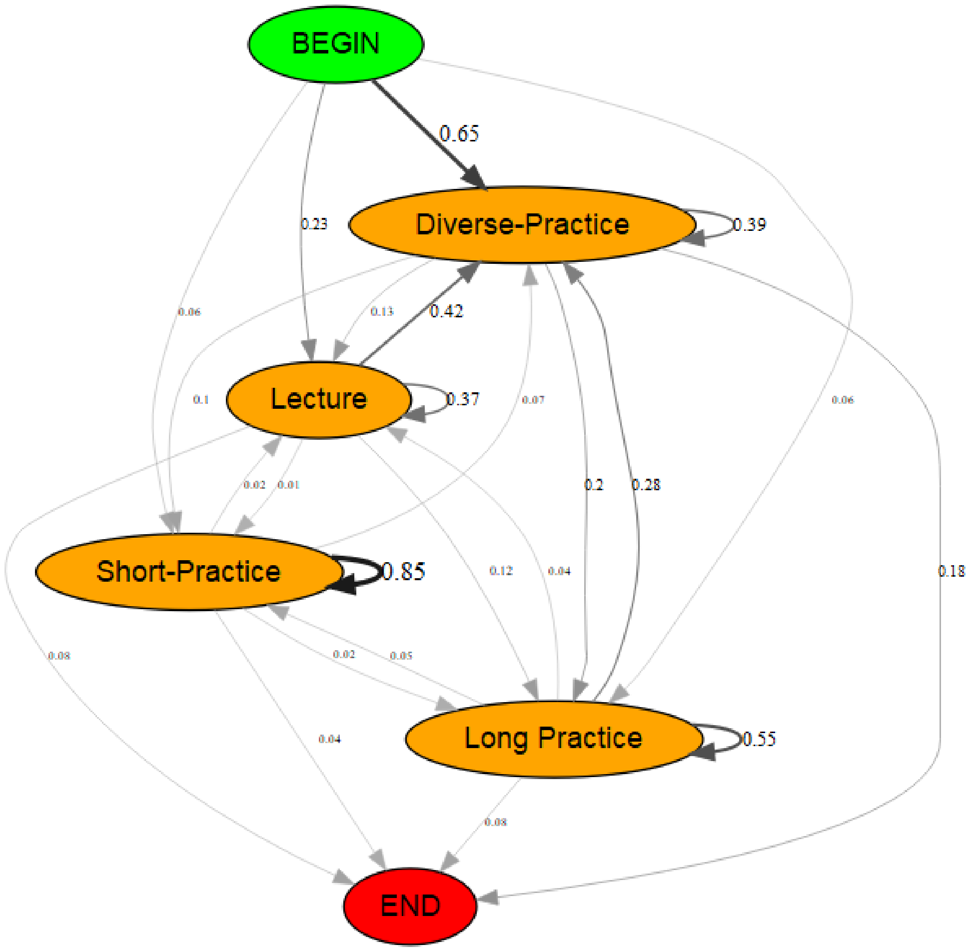


Figure 9 Process Model of Highly Active Group in Introduction to Python Course

5.3 Summary

As indicated by several researchers, the use of learning tactics and strategies are context-dependent (Entwistle, 2007; Lust et al., 2013; Rachal et al., 2007). As such, it is essential to explore how learning tactics and strategies detected from trace data align with the findings from the traditional research. Aligned with the studies carried out by using data analytics-based approach in other chapters, the results presented in this chapter are discussed in response to the approaches to learning (RQ3 in Section 1.1). That is, the surface approach which is characterised by the shallow engagement and assessment-focus (Entwistle, 1991) was detected by the process analytics-based approach across the three learning contexts. The surface approach to learning was represented by the *Inactive* strategy group in the MOOC and the *Highly Selective-Low Engagement* strategy group based on the flipped classroom and blended learning datasets. The deep approach to learning represents a high level of engagement and diverse tactics application (Chonkar et al., 2018; Mattick et al., 2004). The strategies indicative of deep approach to learning were detected across the three different learning contexts. That is, *Highly Active* strategy in the MOOC, and *Intensive-High Engagement* in the flipped classroom and blended learning settings were representatives of the deep approach to learning. However, the strategic approach which is characterised by strategically selecting the tactics, by focusing on the assessment, and by employing a considerable amount of effort (Chonkar et al., 2018; Diseth, 2003), was only detected in the flipped classroom and blended learning settings (i.e. *Strategic-Moderate Engagement* strategy).

Additionally, the examination of the analysed learning tactics and strategies in different learning contexts enables us to explore the research question four (RQ4) which was formulated to examine how SRL constructs are associated the adoption of learning tactics and strategies. Specifically, the study contributes to the insights into the connection of the task conditions to the choice of learning tactics and strategies. That is, the results of data analytics-based detection of learning tactics reveal that learning tactics are influenced by course design, representing how students handle tasks at hand in each learning session. We observed two common features representing the characteristics of learning tactics across different contexts, namely, the composition of learning actions and the length of the sequences. Future research into the detection of learning tactics should consider these features when dealing with the detection of the tactics from the trace data. The detected learning strategies are less sensitive to course design but are rather influenced by delivery modalities. That is, in the learning contexts that combine both face-to-face and online learning activities (i.e., blended learning and flipped classroom), we detected similar learning strategies. In the MOOC, we detected two strategies exhibiting similar behaviours to those found in the blended learning and flipped classroom. However, one different learning strategy exhibiting the transition from using effective learning strategies to the use of less effective ones was detected in the MOOC. That is, the *Highly Active at the Beginning* strategy group showed that learners employed the deep approach to learning but after few weeks, the students changed to the surface approach to learning. This might be due to the difficulty of the learning activities in a particular week or the learners' inexperience in studying computer programming (Matcha, Gašević, Ahmad Uzir, Jovanovic, et al., 2020).

Another significant contribution of this chapter is that by replicating the application of the process analytics-based approach across three different learning contexts, evidence supporting the generalisability of the approach is provided, hence, addressing research question two (RQ2) raised in Section 1.1. That is, the approach can analyse trace data that are collected on different learning platforms as the detected tactics and strategies could be interpreted according to relevant educational theory, i.e., approaches to learning as discussed above.

The association with the SRL constructs explored in this chapter only covered the relationship with the task conditions. The extent to which the approach enables the detection of learning tactics and strategies that are associated with the cognitive conditions is yet unexplored. The next chapter focuses on how cognitive conditions are associated with learning tactics and strategies that can be detected from trace data by using process analytics-based approach.

6 Analytics of Learning Strategy: Associations with Personality Traits

> People are of course influenced by external forces and chance events, but at the end of the day each of us can wave the magic wand of freedom and decide things for ourselves.
>
> — Yuval Noah Harari, *Homo Deus A brief History of Tomorrow*

6.1 Introduction

Cognitive conditions refer to the restrictions that are internal to the learning process of an individual. They can have significant impact on the selection of learning tactics and strategies (Greene & Azevedo, 2007; Winne & Hadwin, 1998). Disposition of individual students is one example of cognitive conditions (Winne & Hadwin, 1998). Dispositions reflect the characteristics of individuals and the characteristics can lead to distinctive forms of behaviours, cognition, and emotions (Bidjerano & Dai, 2007; Roberts & Mroczek, 2008). As such, dispositions of an individual may have an influence on their engagement with learning activities (Tlili et al., 2016). Personality is one of the most studied dispositions (Conard, 2006; Farsides & Woodfield, 2003). Personality is regarded as cognitive and behavioural patterns which individuals demonstrate consistently across time and circumstances (Bozionelos, 2004). One of the well-known personality measurement is by analysing personality traits (John & Srivastava, 1999). Personality is commonly characterised by five traits (John & Srivastava, 1999), including,

- Extraversion – is described as a sociable, talkative and active person.
- Agreeableness – is referred to a person who is generous, compassionate and kind.
- Conscientiousness – is described as being a dependable, responsible, organised, disciplined and achievement-focused individual.
- Emotional Instability (also referred to as neuroticism) – is referred to as a person who tends to be anxious and sensitive.

- Openness – is recognised as a creative person who is willing to be exposed to new experiences.

Personality traits have been found to be associated with learning strategies in educational research that used self-reported instruments (Furnham et al., 2009; Marcela, 2015; Zhang, 2003). However, the relations between automatically detected strategies that were extracted from the trace data and the established psychological constructs such as personality traits are under explored.

6.1.1 Chapter overview

This chapter replicates the use of the process analytics-based approach as applied in Chapter three-Chapter five on a trace dataset offered in an on-demand MOOC. The course was on-demand as learners could enrol it and start studying at any time. Same as with the other chapters, the detected learning tactics and strategies will be examined according to the theory of approaches to learning (RQ3) (Biggs, 1987; Entwistle, 1991).

In addition, another major contribution of this chapter is the examination of personality traits. This is a pioneering study that aimed to investigate the association of a psychological construct (i.e., personality trait) and learning strategies that are automatically detected from trace data in order to address research question four (RQ4) defined in Section 1.1. The Ten-Item Personality Inventory (TIPI) was used as a research instrument to collect data about students' personality traits at the time of course enrolment. A multinomial logistic regression was then applied to examine the relationship between learning strategies detected with the process analytics-based approach and personality traits. The findings unveiled a significant association of deep learning approaches with consciousness and agreeableness. The adoption of learning strategies indicative of surface approaches to learning was found to be associated with emotional instability. These findings are consistent with the previous research on the relationship between learning strategies and personality traits that were conducted with the use of self-report instruments (Chamorro-Premuzic et al., 2007; Duff et al., 2004; Shokri et al., 2007; Woodfield et al., 2006; Zhang, 2003).

6.2 Publication: Analytics of Learning Strategies: the Association with the Personality Traits

The following section includes the verbatim copy of the following publication:

> Matcha, W., Gašević, D., Jovanović, J., Ahmad Uzir, N., Oliver, C. W., Murray, A., & Gasevic, D. (2020). Analytics of Learning Strategies: the Association with the Personality Traits, *Proceedings of the 10th international conference on learning analytics and knowledge (lak '20)*, Frankfurt, Germany, ACM. https://doi.org/10.1145/3375462.3375534

Analytics of Learning Strategies: the Association with the Personality Traits

Wannisa Matcha*
School of Informatics, University of Edinburgh
Edinburgh, United Kingdom
w.matcha@ed.ac.uk

Dragan Gašević
Faculty of Information Technology, Monash University, and University of Edinburgh
Melbourne, Australia
dragan.gasevic@monash.edu

Jelena Jovanović
Faculty of Organisational Sciences, University of Belgrade
Belgrade, Serbia
jelena.jovanovic@fon.bg.ac.rs

Nora'ayu Ahmad Uzir
School of Informatics, University of Edinburgh, and Universiti Teknologi MARA
Edinburgh, United Kingdom
n.uzir@ed.ac.uk

Chris W Oliver
Physical Activity for Health Research Centre, University of Edinburgh
Edinburgh, United Kingdom
chris.oliver@ed.ac.uk

Andrew Murray
Sport and Exercise, University of Edinburgh
Edinburgh, United Kingdom
docandrewmurray@gmail.com

Danijela Gasevic
School of Public Health and Preventive Medicine, Monash University
Melbourne, Australia
danijela.gasevic@monash.edu

ABSTRACT

Studying online requires well-developed self-regulated learning skills to properly manage one's learning strategies. Learning analytics research has proposed novel methods for extracting theoretically meaningful learning strategies from trace data originating from formal learning settings (online, blended, or flipped classroom). Thus identified strategies proved to be associated with academic achievement. However, automated extraction of theoretically meaningful learning strategies from trace data in the context of massive open online courses (MOOCs) is still under-explored. Moreover, there is a lacuna in research on the relations between automatically detected strategies and the established psychological constructs. The paper reports on a study that (a) applied a state-of-the-art analytic method that combines process and sequence mining techniques to detect learning strategies from the trace data collected in a MOOC (N=1,397), and (b) explored associations of the detected strategies with academic performance and personality traits (Big Five). Four learning strategies detected with the adopted analytics method were shown to be theoretically interpretable as the well-known approaches to learning. The results also revealed that the four detected learning strategies were predicted by conscientiousness, emotional instability, and agreeableness and were associated with academic performance. Implications for theoretical validity and practical application of analytics-detected learning strategies are also provided.

CCS CONCEPTS

• **Applied computing** → *Distance learning*.

KEYWORDS

learning strategies, approaches to learning, personality traits, learning analytics

ACM Reference Format:
Wannisa Matcha, Dragan Gašević, Jelena Jovanović, Nora'ayu Ahmad Uzir, Chris W Oliver, Andrew Murray, and Danijela Gasevic. 2020. Analytics of Learning Strategies: the Association with the Personality Traits. 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/10.1145/3375462.3375534

1 INTRODUCTION

Modern technologies afford many opportunities for online learning. Recently, massive open online courses (MOOCs) have attracted much attention and increased considerably the repertoire of courses that are often freely and openly available to learners in a range of academic disciplines. Learners can enrol into courses that suit their needs, preferences, and situations [31]. The structure and course design of MOOCs are diverse [33]. In terms of the mode of delivery, some MOOCs are scheduled in a specific time frame, while others are self-paced and open for free access throughout the year. In

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

ACM ISBN 978-1-4503-7712-6/20/03…$15.00
https://doi.org/10.1145/3375462.3375534

terms of the learning activities, some MOOCs emphasize practical exercises (e.g., programming courses [38]), while others focus on discussion and/or peer assessment.

To be successful in online learning, learners need to have strong skills to self-regulate their learning [1]. This requires that they can adapt their learning strategies to meet the needs of a learning situation [36, 17]. Much research in educational psychology focuses on the strategies used by learners in formal courses (i.e., courses that are offered for credit by a university and are scheduled to run in a specific period of time). Learning strategies used by learners enrolled in courses that offer flexibility, such as MOOCs, are still under-explored.

Existing research in learning analytics has proposed novel approaches to detection of learning strategies from the trace data collected by platforms for online learning. These approaches are typically grounded in the literature on self-regulated learning and are based on a combination of unsupervised machine learning with sequence, process, and/or network analytics [21, 28, 36, 43, 42]. The applications of these approaches have demonstrated that theoretically meaningful learning strategies can be detected and that they usually correspond to the notion of approaches to learning [18, 23].

This paper reports on a study that aimed to address two limitations in the existing literature on the use of learning analytics for detection of learning strategies. *First*, a majority of the published studies on detection of learning strategies are conducted in formal education settings, while there is limited work published in the context of MOOCs. As well-documented in the literature[19, 31], learners in MOOCs come with higher heterogeneity in motivations, prior knowledge, and time management and study skills than it is usually the case in formal education. Therefore, *the first aim of this study was to examine if theoretically meaningful learning strategies could be extracted from trace data collected in a MOOC and if such strategies were associated with academic performance.*

Second, there is a little understanding of the links between automatically detected learning strategies, i.e., the adoption of those strategies by learners, and learners individual differences. Individual differences have long been identified as one of the key factors associated with both learners' behavior and academic achievement [53]. One of the key psychological individual differences that have been identified as predictive of learning outcomes is personality [11, 20]. Personality traits are also found to be associated with self-reported (but not automatically detected from trace data) learning strategies [22, 58]. Therefore, *the second aim of this paper is to explore how learning strategies extracted from trace data are related with personality traits.* This link is important as it can advance theoretical understanding of automatically detected learning strategies with respect to the established constructs in psychology. This can also have practical implications for personlization with learning analytics.

2 BACKGROUND

2.1 Learning Strategies

The ways learners engage in different learning tasks reflect the employed learning tactics and strategies. Rachel et al.[50] refer to learning strategies as any methods or techniques used by learners to perform cognitive operations in order to facilitate knowledge acquisition and integration. Such learning strategies are guided by learning goals and oriented towards enhancing learning outcomes and performance [55]. The terms learning tactic and strategy are often used interchangeably in the literature [39]. Several researchers have highlighted the differences between them though [40, 13, 39, 26]. A learning tactic is an individual technique used by learners to complete a learning task [13]. A tactic can be observed through the sequences of learning actions that were carried out by the learners in a given learning session [26]. A learning strategy involves selecting, combining, coordinating, and utilising tactics and cognitive operations to achieve a set learning goal [40, 13, 39]. In other words, a learning strategy represents a regularity of tactics used by a learner [13].

In an online learning environment, where high-level of self-regulated learning is required, learners need to monitor the enactment and continuously evaluate the effectiveness of their learning strategies. Unlike formal, face-to-face or blended, and for-credit learning environments where learners' goals are relatively homogeneous and oriented towards completing a course, in a MOOC, learners' goals are much more diverse, while learners are expected to possess well developed skills to self-regulate their learning [1, 30]. Therefore, learning strategies in the online learning environment, and especially MOOCs, might be different from those used in the conventional learning setting. For instance, Morehead et al. [46] found that the majority of online learners avoid using the note-taking technique when participating in an online course. Underdeveloped skills to adjust the learning strategy to suit the course design has been identified as one of the factors that can reduce learning success in blended higher education courses and also in MOOCs [36, 19].

Conventionally, educational psychology research has used self-report instruments to study learning strategies [59, 23]. However, limitations of self-reports are well-documented in the literature [59, 56]. The limitations primarily stem from learners' incomplete and biased memories of their learning experience [4, 59]. Furthermore, self-reports capture learners intentions related to the use of a specific learning tactic or strategy. However, research has demonstrated that learners tend to behave differently from what they have described in their self-reports [59, 23]. As an alternative that can mitigate the limitations of self-reports, the use of trace data is suggested [51]. Trace data collect information about learning actions that are temporally proximal to the actual learning situations and that reflect the "realized intentions" of learners [59]. Therefore, exploring learning strategies through the traces of learning activities that learners undertook to accomplish their learning tasks has a strong potential to shed some light on how students actually self-regulate their learning [28].

The literature offers several analytic methodologies for detection of learning patterns from trace data. For instance, Jovanović et al. [28] proposed a sequence mining methodology to extract learning strategies from the trace data collected in a computer engineering course that followed a flipped classroom pedagogy. They found a significant association between learning strategies and academic performance with a moderate effect size. Maldonado-Mahauad et al. [38] used process mining to identify strategies of self-regulated learning (SRL) of MOOC learners. They explored the process model in terms of a) the most frequent learning sequence, b) the process of model of high and low performance learners and c) the process model of learners based on different SRL strategy groups collected from the self-report [38]. Boroujeni and Dillenbourg [5] proposed a sequence mining approach to extracting learning patterns of MOOC learners. Kizilcec et al. [30] captured and examined self-regulated learning strategies of MOOC learners across six different courses through a combined use of trace data and self-reports.

There is currently a lacuna in research on automatic detection of learning tactics and strategies from trace data collected in a MOOC where learning tactics and strategies are theoretically meaningful. Exiting research has predominantly identified theoretically meaningful strategies and tactics from trace data in formal for-credit contexts, e.g., [28]. When similar data analytic methods are applied to analyze trace data in MOOC contexts, such studies are primarily focused on either data-driven learning patterns without much theoretical interpretation [5] or patterns of self-regulated learning [38]. Therefore, in this study, we aimed to explore the following research question:

RQ1: Given a MOOC, can we detect theoretically meaningful learning tactics and strategies? If so, is there any association of the detected strategies and academic performance?

The association with academic performance is tested to check external validity of the automatically detected strategies.

2.2 Personality Traits

Personality traits have long been used to study the characteristics of an individual [27]. Research has identified five (pairs of) personality traits that are referred to as the "Big Five", namely, a) openness to new experiences, b) conscientiousness, c) extraversion, d) agreeableness, and e) neuroticism or emotional instability [27].

Openness refers to proactive behaviors toward seeking and exploring new experiences [27]. Those who have a high score in the openness are more likely to be creative, whereas those with low scores tend to have a conventional thinking method [27]. Conscientiousness is identified as the trait of persistent, responsible, and dependable persons. A high score in conscientiousness reflects a high ability to maintain goal-directed behaviors, as well as high likelihood of being self-dependable and self-disciplined [27]. Extraversion reflects sociable behaviors. Those who have a high score in extraversion are more friendly, energetic, and outgoing than those with a low score. Agreeableness is characterised by being generous, kind, and helpful. Neuroticism is also referred to as emotional instability. A high score in neuroticism reflects emotional instability and anxiety, whereas a lower score is an indicator of calmness and ability to control one's emotions [27].

Existing research has provided evidence for the positive correlation between academic performance and personality traits of conscientiousness and openness to new experiences [48, 41, 14]. Neuroticism or emotional instability showed a negative effect on learning outcomes [16]. Several studies further explored the relationship of the personality traits and approaches to learning [58, 16, 7, 9], a construct well-known in educational psychology and often found closely connected to learning strategies in the learning analytics literature [23, 44]. Three approaches to learning are proposed by Biggs [3], namely, deep, achieving/strategic and surface learning. Deep approach to learning is characterised by the intention of understanding learning content and putting high amount of effort to achieve a learning goal [3]. Surface learners employ a minimal effort to pass the course. Achieving or strategic approach to learning refers to the objective of achieving high performance with a minimum effort [3]. Research found that conscientiousness and openness were predictive of learners adopting a deep approach to learning [58, 16] whereas, emotional instability was predictive of surface approach to learning [58].

The relationship of personality traits and self-regulated learning strategies are also explored in the literature [2]. The self-regulated learning strategy was captured by using the Motivated Strategy for Learning Questionnaire (MSLQ). They found the connection of conscientiousness and the time management, elaboration, critical thinking, and meta-cognitive skills [2]. Cela-Ranilla et al. [6] explored the relationship of the personality traits and the learning patterns captured by using self-reports. Learners behaviors were divided into four learning patterns: sequential (step-by-step learners), precise (learners who aims to understand the learning content in details), technical (highly hand-ons activities focused learners), and confluent (learners who avoid the conventional learning approach). They found the positive relationship between conscientiousness and the sequential and precise learners [6]. Extraversion was negatively connected to the technical learning pattern.

To our knowledge, most of the existing research on personality traits and learning strategies (e.g., [2, 6]) have relied on the use of self-reporting instruments (e.g., questionnaires). Learning analytics has enabled the analysis of the dynamics of learning processes, hence, actual tactics and strategies adopted by learners can be detected. Therefore, in this study, we aimed to explore the connection of the learning strategies detected from the actual learning steps taken by learners and the personality traits as reported by individual learners. Accordingly, the following research question is formed to guide this study:

RQ2: Is there an association between the learning strategies adopted by learners in a MOOC and any of their personality traits?

3 METHODOLOGY

3.1 Study Context

An open online learning course, *Sit Less, Get Active* is offered by a University of Edinburgh through the Coursera learning platform. With its initial offering in June 2016, it was the first MOOC aimed at raising awareness of the relevance of physical activity for health and providing practical guidance for increasing the level of physical activity in everyday life. The MOOC consists of three weeks of core learning material, followed by the weekly physical activity promotional messages and monthly physical activity promotional videos, sent via announcements, for six months to serve as nudges to help people remain active. The core learning activities consisted of videos (each week contained 5 videos that promoted physical activity), optional readings, quizzes, discussions in forums, and assignments. There were two assignments students were asked to complete: Assignment 1: to identify (and write down) a physical activity goal that a learner would like to achieve in the next month; and Assignment 2: to monitor physical activity (measure a step count) for a week. There were five graded activities in this MOOC, and learners were expected to pass all five to complete the course: Assignment 1 (100% passing threshold), Assignment 2 (70% passing threshold i.e. physical activity being monitored on most days of the week) and three quizzes (70% passing threshold for each quiz). Learners were expected to spend approximately one and a half hour each week to complete the weekly learning activities. However, the schedule was flexible, that is, after enrolling into the course, learners could immediately start working on the learning tasks and complete them at their own pace.

Learners included in this study are those who completed at least one of the five graded activities and responded to the baseline questionnaire that contained questions on personality. In this study, the survey with the highest response rate was used. Three questionnaires in total were distributed during the course. One was delivered before the learners engaged with the course material, second right after the completion of the core course, and third six months after course completion, i.e. upon delivery of all promotional messages and videos. In total 1,397 learners were included, their ages ranged from 15 to 80 years (Mean = 45.59, SD = 14.36). More than 78 percent of them were female (N_{male}= 295, N_{female}= 1,102). The majority had a higher education degree (87.47%) and worked full time (59.53%). 53.47 percent managed to pass the course (N_{pass}= 747, N_{fail}= 650).

3.2 Data

Trace data were collected from the Coursera learning platform. The activities recorded in the trace data, including their description and the associated actions, are presented in Table 1.

To capture the personality traits, a self-reporting instrument was used. Several questionnaire for assessing personality traits have been developed. They range from as large as the 240 items questionnaire developed by Costa and McCrae [12], 100 items instrument by Goldberge in 1992 [24], 44-item Big-Five Inventory [27], and 10-item and 5-item questionnaires [25]. In this study, the Ten-Item Personality Inventory (TIPI) [25] was used to capture the personality of students who enrolled and completed at least one assessment in the *Sit Less, Get Active* course. TIPI was part of the baseline survey the learners completed before they engaged with learning materials. Although a shorter questionnaire might offer inferior psychometric perspective than a lengthier one, it also comes with some important advantages: reduced costs of answering to the highly similar questions, reduced boredom and frustration, and often greater response rate [25]. TIPI has proven to be an adequate and valid instrument for capturing the personality traits in terms of the convergent validity, test-retest reliability, and prediction pattern of external correlation as compared to the most frequently adopted 44-item Big-Five Inventory [25]. TIPI uses a seven-item Likert-like scale and contains 10 question items that correspond to the two opposite characteristics of each of the five personality traits. That is, the use of TIPI allowed us to create five variables, each representing one of the Big Five personality traits.

3.3 Data Analysis

To address research question 1, we applied the technique proposed in [44] and as illustrated in Figure 1. Learning actions in the trace

Table 1: Learning actions and descriptions

Activities	Action	Description
Discussion	discussion_answer	Answer to the discussion
	discussion_AUnvote	Undo the vote for the answer to the discussion
	discussion_AVote	Vote for the answer to the discussion
	discussion_QFollow	Follow the discussion question
	discussion_QUnfollow	Unfollow the discussion question
	discussion_QVote	Vote for the question posted in the forum
Exam	exam	Work on practical exam questions
	exam_begin	Begin working on the practical exam
	exam_complete	Complete the practical exam question
	exam_correct	Correctly solve the practical exam question
	exam_incorrect	Incorrectly solve the practical exam question
Quiz	quiz	work on the quiz
	quiz_begin	Begin working on the quiz
	quiz_complete	Complete the quiz
	quiz_correct	Correctly solve the quiz
	quiz	Incorrectly solve the quiz
Supplement	Supplement_start	Access the extra reading materials
	Supplement _complete	Complete the access to the extra reading materials
Video	video_begin	Begin the video
	video_download	Download the video
	video_end	End the video
	video_pause	Pause the video
	video_play	Play the video
	video_playback _rate	Change the playback rate of the video
	video_seek	Seek the video
	video_subtitle _change	Change the subtitle of the video
	video_subtitle _download	Download the subtitle
	video_volume_change	Change the volume
	video_wait	Wait for the video to load
Information	Page_preload	Load the page for the first time
	Page_view	View the page

data were initially processed by grouping them into learning sessions. Learning sessions were created by splitting sequences of learning actions of individual students using the threshold of 30 minutes of inactive time between two consecutive actions [32]. Then, the overly short and overly long sessions were excluded from further analysis [28]. As a result, 8,990 learning sessions were included into the learning tactics and strategies detection process. To identify learning tactics, a First Order Markov Model (FOMM) was created using learning sessions (i.e. sequences of learning actions) as the input. The FOMM produced the overall learning process based on the probability of transition of actions in each learning session according to the timestamps. The transition matrix produced by the FOMM served as the input for the expectation-maximization (EM) algorithm to detect prototypical groups of learning sessions. Thus identified patterns of learning behaviour were considered as a manifestation of the students' learning tactics.

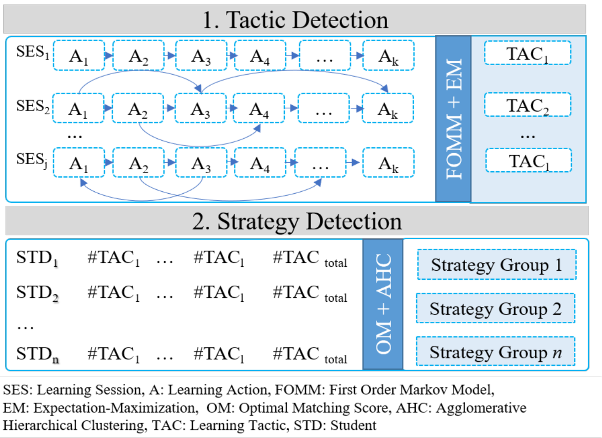


Figure 1: Learning tactics and strategies detection process [43]

The number of each tactic used and the total number of tactics used by individual were taken as the input to detect the strategies adopted by the learners. The agglomerative hierarchical clustering based on Ward's method and the Euclidean similarity measure was used to categorise learners the based on the learning tactic usage patterns. To examine the tactic regulation for each detected strategy, we plot the mean numbers of tactic used across the course topics.

The association between learning strategy group and academic performance was tested by comparing whether there were significant differences between the learning strategy groups on academic performance (i.e. the course final grade obtained). This was done by using the Kruskal Wallis test followed by pair-wise Mann-Whitney U tests. The Bonferroni correction was used to prevent the risk of the 'p-inflation' error.

To address research question 2 and examine the personality traits, first, the descriptive statistics were used to summarise the information reported by the learners. This was followed by multinomial logistic regression (MLR) to examine the likelihood of the students' personality scores to predict the assignment of the learners to the learning strategy groups. The MLR is used to study the discrete unordered categorical dependent variables [34]. Specifically, the dependent variable was the assignment of the learners to the learning strategies and the independent variables were the five variables representing students' scores on the five personality traits as measured by TIPI. The MLR computes the odds of preference event (i.e. chosen learning strategies) against a reference category (i.e., an alternative strategy). We examined the odds of learners adopting a specific learning strategy over the other strategies, using each learning strategy as a reference category.

4 RESULTS

4.1 Research question 1

4.1.1 Learning Tactics. Four tactics were detected as distinctive patterns in sequences of learning actions within the students' learning sessions. The supplementary document[1] provides a detailed insight into each learning tactic in terms of the distribution of actions within learning sessions, chronological order of learning actions within sessions, frequency of individual learning actions, as well as the process model of each learning tactic. Fig. 2 presents, for each detected tactic, the distribution of learning actions within learning sessions that 'belong' to a particular tactic.

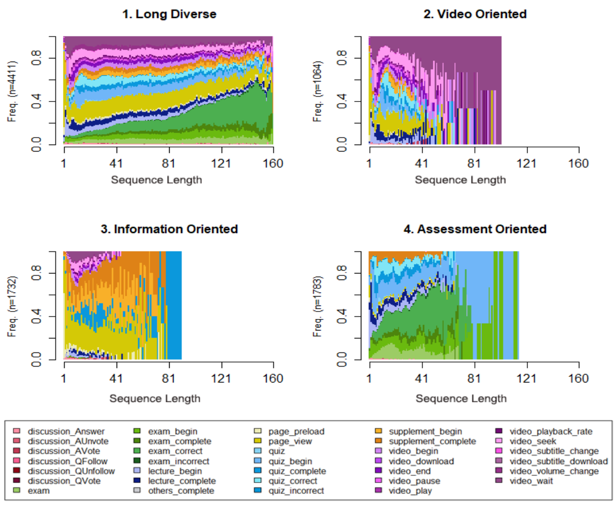


Figure 2: The state distribution of each learning tactic

- Long Diverse (N = 4,411 sessions, 49.07%): This was the most frequently chosen tactic, characterised by long learning sessions (Mdn = 61 actions in a session). Each session consisted of a variety of learning actions. In general, the sessions primarily included access to information pages and interaction with the videos. Interaction with the quizzes and the practical exams were present as well, but to a less extent.
- Video Oriented (N = 1,064 sessions, 11.84%): The length of learning sessions of this tactic was moderate (Mdn = 16 actions per session). The most dominant actions were those associated with the course videos. Access to the information pages and interaction with the quizzes were present, as well.

[1] http://tiny.cc/3cbsdz

- Information Oriented (N = 1,732 sessions, 19.27%): This tactic is characterised by the shortest learning sessions (Mdn = 6 actions per session). The most dominant actions in this tactic included access to the supplementary learning materials and other information pages. Interaction with the course videos and interaction with other students in discussion forums were also observed, but to a less extent.
- Assessment Oriented (N = 1783 sessions, 19.83%): This tactic consisted of moderately long learning sessions (Mdn = 17 actions per session). Its main feature is the dominance of the exam- and quiz-related actions.

4.1.2 Learning Strategies. The four detected learning strategies are depicted in Fig. 3.

- Disengaged (N = 440 learners, 31.50%): This is the largest strategy group and at the same time the group with the lowest level of activity. When interacting with the course content, this group tended to use the *Long Diverse Oriented* and *Assessment Oriented* tactics. However, the number of tactics used decreased as the course progressed. The learners in this group showed the lowest course performance (Mdn (Q1 and Q3) = 40.00 (38.33, 90.00)). The median number of passed graded activities was 2 out of 5, and only 30.68% managed to pass the course. On average, these learners spent 6.24 days interacting with the course content.
- Surface (N = 319 learners, 22.83%): The learners in this group were on average moderately active at the beginning of the course. The number of tactics used gradually decreased with each topic unit. The learners approached their studying by using primarily the *Assessment Oriented* tactic and *Long Diverse Oriented* tactic. In terms of academic performance, the learners obtained relatively low course mark (Mdn (Q1,Q3) = 78.33 (40.00, 98.18)) and the median number of passed graded activities was 4. Only 46.71% managed to complete and pass the course. The median number of days that the learners spent on the course was 20.35 days. These learners spend the median of 14.88 days to complete the final graded activities.
- Active (N = 410 learners, 29.35%): Similar to the *Highly Active* strategy group, the *Active* learners were moderately engaged in learning and extensively used the *Long Diverse Oriented* tactic. The application of the *Long Diverse Oriented* tactic gradually dropped as the course progressed. For the final, third course topic, the learners increased the application of *Information Oriented* tactic. This tactic reflects the concentration on the supplemental reading materials. The use of other learning tactics was hardly observable. In terms of performance, this strategy group obtained relative high scores (Mdn (Q1, Q3) = 91.2 (76.36, 98.18)), the median of passed graded activities was 5, and 66.10% of the learners passed the course. This group spent slightly more time than the *Surface* group (Mdn = 21.28 days). These learners took on average 16.63 days to complete the final graded activities.
- Highly Active (N = 228 learners, 16.32%): This learning strategy group showed a high level of engagement. The learners applied several learning tactics. The most dominant tactic was *Long Diverse Oriented*, though its use gradually decreased as the course progressed. The *Information Oriented* tactic was the second most dominant tactic for this strategy group. The application of this tactic was observed during the interaction with the learning resources of the first and the third course topic. The *Video Oriented* tactic was consistently observed throughout the course. The learners in this group obtained the highest grade (Mdn (Q1, Q3) = 98.18 (90.00, 100.00)). The percentage of learners who passed the course was higher than in the other three learning strategy groups (84.21 % learners passed the course). The learners spent a longer amount of time interacting with the course content (Mdn = 51.51 days) than in the other three groups. However, the learners in this group only took a median of 19.09 days to complete the final graded activities. This indicates the behaviors of revisiting the course resources even after the learners had completed the course.

4.1.3 Association with academic performance. The Kruskal-Wallis test revealed the significant association of learning strategy group and course grade ($\chi^2(3)$=232.27, p < 0.0001). Table 2 presents the results of the pairwise comparisons (done with Mann-Whitney U tests) of learning strategy group with respect to the course mark. Overall, we observed the significant association of learning strategy group and academic performance across each pair of the groups. The effect sizes ranged from small (Surface Disengaged) to large (Disengaged Highly Active).

Table 2: Pairwise comparison of strategy groups with respect to the course grade

Strategy	Strategy	Z	p	r
Surface	Active	-5.61	< 0.0001	0.208
Surface	Disengaged	5.08	< 0.0001	0.185
Surface	Highly Active	-8.87	< 0.0001	0.379
Active	Disengaged	11.32	< 0.0001	0.388
Active	Highly Active	-5.07	0.0003	0.201
Disengaged	Highly Active	-13.11	< 0.0001	0.507

4.2 Research question 2

Table 3 presents the descriptive statistics of the personality traits as reported by the learners. In general, the learners rated themselves as neutral in terms of the extraversion. The median scores of agreeableness, conscientiousness and openness were high (Mdn > 5 out of 7) and emotional instability was relatively high (Mdn = 4.6).

Table 3: Summary of personality traits score

Personality	Median (Q1 and Q3)
Extraversion	3.5 (2.5-5.0)
Agreeableness	5.0 (4.5-6.0)
Conscientiousness	5.5 (4.5-6.5)
Emotional instability	4.5 (3.5-6.0)
Openness	5.5 (4.5-6.0)

The table presents the results of the MLR in terms of the odd ratios of changing the reference learning strategy based on the

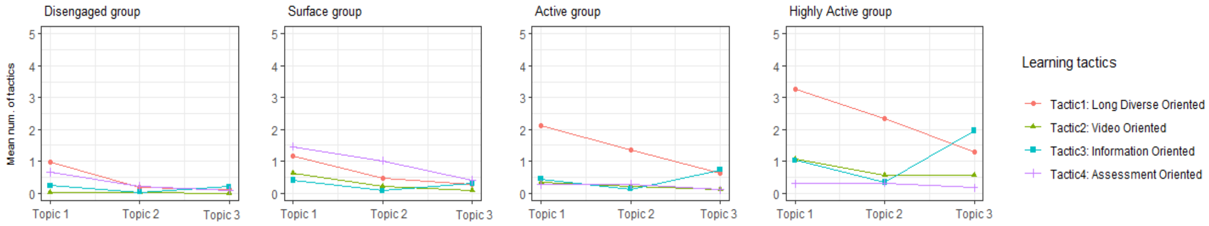


Figure 3: The mean number of learning tactics used by each learning strategy group

personality traits scores can be found from the supplementary document [2]. This table shows the results obtained by taking each strategy group as the reference.

Considering only unique relationships, the significant results obtained from the MLR can be interpreted as follows: a unit increase in the agreeableness was associated with 15.6% greater odds of learners adopting the *Highly Active* learning strategy compared to being Disengaged (OR (95%CI) = 1.156 (0.999, 1.338)). Similarly, an increase in conscientiousness score by one unit was associated with 14% higher odds of adopting the *Highly Active* over Disengaged learning strategy (1.143 (1.007, 1.298)). In contrast, an increase in emotional instability scores was associated with lower odds of opting for *Highly Active* over *Disengaged* learning strategy (0.855 (0.755, 0.967)), i.e. we would expect that the learners would be more likely to remain *Disengaged* rather than to use the *Highly Active* learning strategy. Furthermore, agreeableness was positively associated with the likelihood of adopting the *Surface* strategy over being *Disengaged* (1.223 (1.070, 1.399)), while a one unit increase in extraversion score was associated with lower likelihood of choosing the *Active* Learning strategy compared to being *Disengaged* (0.925 (0.847, 1.011)).

Conscientiousness (1.191 (1.042, 1.362)) and openness (1.150 (0.981, 1.348)) were positively, while emotional instability (0.862 (0.756, 0.983)) was negatively associated with likelihood of using the *Highly Active* compared to the *Surface* strategy. An increase in openness score by one unit was associated with 12.8% higher odds of adopting the *Active* over the *Surface* learning strategy (1.128 (0.984, 1.293)). In contrast, an increase in extraversion (0.897 (0.815, 0.988)) and agreeableness (0.814 (0.710, 0.933)) scores was associated with greater odds of adopting the *Surface* compared to the *Active* learning strategy. In addition, a unit increase in agreeableness score was associated with 16% higher odds of adopting the *Highly Active* over the *Active* learning strategy (1.161 (1.001, 1.347)). Finally, a negative association was observed between emotional instability and likelihood of adopting the *Highly Active* over *Active* learning strategy (0.800 (0.705, 0.907)).

5 DISCUSSION

5.1 Research question 1

5.1.1 Learning Tactics and Strategies. The learning tactics detected based on the analytic approach used in the study depict a pattern of how learners interacted with the learning activities in each learning session [26]. We detected four learning tactics, namely, *Long Diverse, Video, Information , and Assessment Oriented.* In general, the *Long Diverse* tactic was the most frequently used. This pattern showed the longest learning sessions in which the learners interacted with several learning activities. This pattern was also observed by others in analytics-based detection of learning patterns from trace data (i.e. [54, 47, 21, 38]). The *Assessment oriented* tactic represented the pattern of interaction dominantly with the quizzes and exams, this similar pattern was also detected by Maldonado-Mahauad et al. [38]. [6] referred to this pattern as the technical focused learning pattern based on the conventional pattern detection using a self-reporting instrument. The *Video Oriented* tactic demonstrates the dominant interaction pattern to the video were also detected by the previous work [44]. The *Information Oriented* tactic represented the dominant access to the supplement reading materials. Similar tactic indicative of interaction with the reading materials was detected in [21].

Patterns of how learners applied the tactics to study the weekly topics in the MOOC represent learning strategies [13]. Even though the course was short (i.e., only three weeks), meaningful patterns of how learners approached their learning were observed. In particular, similar to [23], we found that the detected learning strategies resemble Bigg's approaches to learning [3]. The approaches to learning adopted by learners are influenced by intuition and motivation, learning situation, and learning content [3, 18]. This notion is well aligned with the self-regulated learning theory. According to Winne and Hadwin[55], how learners operate their learning is influenced by learning goals, tasks conditions, and cognitive conditions.

We found that the *Highly Active* learning strategy group showed a high level of engagement and applied several learning tactics. By considering the trace data as a reflection of the realised intention, this behavior corresponds to deep approach to learning. Deep learners are highly active and aim to understand learning content by applying several tactics [3].

The *Active* strategy exhibited a similar learning pattern to that of the *Highly Active* strategy. However, their level of interaction with the course activities was lower as compared to the *Highly Active* strategy. These two learning strategies (i.e., *Active* and *Highly Active*) were associated with high academic performance. This is well aligned with the characteristics of learners with deep approach to learning. Research found that deep approach to learning is ideal and learners are more likely to perform better than those who exhibit others approaches to learning [10, 45]. We also observed that learners who employed the *Highly Active* learning strategy

[2]http://tiny.cc/3cbsdz

were more likely to spend more time on the course and to revisit the course content after they completed the courses, the latter being indicative of their intention to review the information they had previously studied.

The pattern of learning behaviour associated with the *Disengaged* strategy is indicative of shallow learning. The analysis revealed very low level of activity after the first week of the course (Fig. 3). This strategy is reflective of surface approach to learning [3]. This approach to learning is found to be associated with low academic performance [10], which is in accordance with our findings: the *Disengaged* strategy group showed the highest fail rate and the lowest performance.

The *Surface* strategy group was moderately active at the beginning of the course. However, their level of activity gradually decreased as the course progress. Learners in this group highly relied on the *Assessment Oriented* tactic. This pattern of learning behaviour is well representative of the surface learning approach, which, according to [3], is oriented towards assessment and passing the course.

5.1.2 Analytics-based methods for strategy and tactic detection. Several analytics based methods for the detection of learning tactics and strategies from trace data have been proposed, including, for example, sequence analysis [28], hidden Markov model [21], process mining [44], and network analytic approaches [43]. A common element of these analytical methods is that the detection of learning tactics is done at the level of learning sessions [21, 43] to capture the technique used by learners in each learning task [26]. The detected patterns indicative of learning tactics were usually shaped by the particularities of the course instructional design. This is well recognized in the learning strategy theory [55]. For example, according to Winne and Hadwin [55], the tactics and strategies adopted by learners are influenced by the task (external) and cognitive (internal) conditions. Therefore, the detected tactics should reflect the structure of the course and the tasks students interacted with. For instance, in a flipped classroom course with a design that emphasised summative assessment, a tactic based on the trial-and-error behavior was observed [28], capturing the pattern of learners interaction with the summative assessment activities. Similarly, in a programming MOOC course, two distinct tactics indicative of how learner worked on practical exercises were observed [43]. In the current study, many supplement reading materials were provided, therefore, we detected a tactic strongly characterised by the access to the supplementary materials.

Regularity of the tactics can be observed by inspecting how learners employ the tactics in each learning period and is indicative of the adopted learning strategy [13]. In a well structured course, e.g., blended or flipped classroom, regulation of tactics could be often observed at the level of the course weeks as course curricula tend to have weekly structure [28, 44]. Since MOOCs typically offer self-paced learning and flexible learning schedule, the tactic usage patterns are not necessarily aligned with the study weeks: different learners exhibit different learning pace and have different beginning and completing time. The study in this paper demonstrated that observing the regulation of learning tactics can be done by using the learning topic as a time reference. Even though, the learning strategies were observed by using different time frame (i.e. study weeks in previous studies [29, 44] and topic in this study), we observed similar learning strategy behaviors with the previous studies. The detected learning strategy groups closely represent the characteristics of the approaches to learning as explained in Section 5.1.1. This research added insight into the application of the analytics method that combines process and sequence mining for the detection of learning tactics and strategies in the MOOC setting.

5.2 RQ 2: Personality Traits and the Choice of Learning Strategy

We found that the learners' adoption of learning strategies, as evidenced in the trace data, was associated with their personality traits. For instance, we observed a positive association between learners' scores on the consciousness trait and their adoption of the *Highly Active* learning strategy. This is in line with several other studies that found consciousness (i.e. dependable personality) to be a significant predictor of deep approach to learning [58, 16], since the *Highly Active* and *Active* strategy groups identified in this study can be considered as having deep approach to learning. Furthermore, one of the characteristics of the conscientiousness is a willingness to do tasks well as directed by the goal [15]. Individuals with high scores in conscientiousness tend to have strong motives to complete the learning tasks and to put the effort to understand the learning content [2]. Considering that our study was situated in a MOOC context where intrinsic motivation and deep approach to learning are required for success, it is not surprising that conscientiousness was associated with the choice of the *Highly Active* learning strategy.

Agreeableness reflects the kindness and generosity. This personality trait is common among females [57]. In this study, more than 78% of the learners were females. We found inconsistent results when examining the association of agreeableness and the detected learning strategies. Agreeableness was associated with higher odds of choosing *Surface* or *Highly Active* strategies over the *Disengaged* one, as well as choosing *Highly Active* over *Active* strategy. However, it was also associated with higher odds of opting for the *Surface* over *Active* strategy. Similar inconsistent findings related to the agreeableness trait have been reported in the literature, namely, agreeableness was found to be predictive of both surface and deep approaches to learning [58, 8] and this requires future research.

The emotional instability or neuroticism trait, in our study, proved to be associated with higher odds of choosing *Active*, *Surface*, or *Disengaged* strategy over the *Highly Active* strategy. Simply put, learners with high score for the neuroticism trait rarely adopted the *Highly Active* strategy. This is consistent with the extant research that has found emotional instability as being predictive of adopting surface approach to learning [58, 16] and this approach being represented in our study by the *Surface* and *Disengaged* strategies. [16] suggests that learners with a high score in emotional instability tend to exhibit anxiety over their academic outcomes, which, in turn, motivate them to highly target assessment activities in order to pass the course. This assessment-focus is well represented in the tactics adopted by *Surface* strategy group.

In contrast to several findings on the association of extraversion and deep learning approach [9, 16], in this study, we observed that extraversion was associated only with higher odds of applying the *Surface* or *Disengaged* learning strategies over the *Active* strategy. In other words, our results indicate a positive association of extraversion and the surface approach to learning. The research reported in the literature found both positive and negative association of the extraversion and the surface learning approach. For instance, Zhang [58] found a negative association between extraversion and the surface learning approach, while Shokri et al. [52] found a positive relationship of extraversion and surface approach to learning.

In contrast to several findings (e.g. [8, 16]), we found no significant association between the openness to experience trait and the applied learning strategies.

6 CONCLUSION, IMPLICATIONS, AND LIMITATIONS

Conclusion. This research explored the association of learning strategies and the personality traits. The learning strategies were detected from the trace data originating from learners interactions with the learning activities in a MOOC. Thus identified learning strategies reflect the "realised intention" of learners [23]. We observed four learning strategies adopted by the students, i.e. *Highly Active, Active, Surface* and *Disengaged* strategies. In terms of the relationship of the detected learning strategies and personality traits, we found associations between some of the personality traits and the strategies adopted by the learners. Conscientiousness and agreeableness were predictive of adopting the *Highly Active* strategy, which is reflective of deep approach to learning. The neuroticism or emotional instability was found to have a positive association with the use of the *Surface* and *Disengaged* strategies, which are reflective of the surface approach to learning. We found no statically significant association of openness and any choice of learning strategies.

Implications. We found that some of the personality traits were predictive of adopting certain learning strategies. Hence, the personality traits might be used in order to inform instructional interventions early in a course. Such interventions can be in the form of personalized suggestions about how to approach learning and what learning strategies to apply in the course. However, personality trait informed interventions should be used only as a starting point and the learners should not be treated throughout the entire course purely on the merit of their personality traits [2]. Instead, as learners engage in learning and start generating trace data, learning strategies could be automatically detected by using the analytic approach presented in this study. Thus, identified insights into the students learning behaviour should complement the personality data to inform instructional suggestions (i.e., feedback) to be offered to learners. The benefits of analytics-based personalized feedback on the choice of learning strategies [44, 21], student satisfaction with feedback [49], and academic performance have already been reported in the literature [35]. However, the novelty of this study is that the 'cold start' in the student support at the very beginning of a course can be overcome by the self-reported personality traits.

Future research, however, should explore the extent to which and how long (in weeks) personality scores can also be used together with automatically detected learning strategies to provide analytic-based personalized feedback. It should also be investigated whether there is a temporal dimension in the association of the personality traits and properties of learning strategies. For example, it would be worth exploring whether the personality traits are predictive of the choices of learning strategies primarily in the early weeks of a course, and whether and how their effect fades away. This association should also be investigated in the context of different instructional interventions targeting the use of effective learning strategies.

Limitations. In this study, the majority of those who voluntarily answered the personality survey were females. The self-report, therefore, might be biased to gender, but this bias is often observed in educational research with self-reports [37]. Moreover, the personality detection instrument applied in this study was based on the 10-items TIPI which cannot capture the subscales of the personality traits as accurately as the longer instruments. Our choice of the 10-item instrument was to reduce the workload of the participants and thus maximize participation. In terms of tactics and strategy detection from the learning trace data, we applied an unsupervised machine learning algorithm. This is subject to the interpretation bias of the researcher. To minimize this bias, we took into consideration the course design and and relevant educational theory (self-regulated learning and approaches to learning).

REFERENCES

[1] Philip C. Abrami, Robert M. Bernard, Eva M. Bures, Eugene Borokhovski, and Rana M. Tamim. 2011. Interaction in distance education and online learning: using evidence and theory to improve practice. *Journal of Computing in Higher Education*, 23, 2-3, (December 2011), 82–103. ISSN: 1042-1726.

[2] Temi Bidjerano and David Yun Dai. 2007. The relationship between the big-five model of personality and self-regulated learning strategies. *Learning and Individual Differences*, 17, 1, 69–81. ISSN: 10416080.

[3] Biggs. 1987. *Student Approaches to Learning and Studying*. 153. ISBN: 0855634162.

[4] Robert a. Bjork, John Dunlosky, and Nate Kornell. 2013. Self-Regulated Learning: Beliefs, Techniques, and Illusions. *Annual Review of Psychology*, 64, 1, 120928131529005. ISSN: 0066-4308.

[5] Mina Shirvani Boroujeni and Pierre Dillenbourg. 2018. Discovery and Temporal Analysis of Latent Study Patterns in MOOC Interaction Sequences. In *Proceedings of the 8th International Conference on Learning Analytics and Knowledge* (LAK '18). ACM, New York, 206âĂŞ215. ISBN: 978-1-4503-6400-3.

[6] Jose Maria Cela-Ranilla, Mercé Gisbert, and Janaina Minelli de Oliveira. 2011. Exploring the relationship among learning patterns, personality traits, and academic performance in freshmen. *Educational Research and Evaluation*, 17, 3, (June 2011), 175–192.

[7] Tomas Chamorro-Premuzic and Adrian Furnham. 2009. Mainly Openness: The relationship between the Big Five personality traits and learning approaches. *Learning and Individual Differences*, 19, 4, (December 2009), 524–529. ISSN: 10416080.

[8] Tomas Chamorro-Premuzic and Adrian Furnham. 2008. Personality, intelligence and approaches to learning as predictors of academic performance. *Personality and Individual Differences*, 44, 7, (May 2008), 1596–1603. ISSN: 01918869.

[9] Tomas Chamorro-Premuzic, Adrian Furnham, and Martin Lewis. 2007. Personality and approaches to learning predict preference for different teaching methods. *Learning and Individual Differences*, 17, 3, 241–250. ISSN: 10416080.

[10] Sonali Prashant Chonkar, Tam Cam Ha, Sarah Shan Hang Chu, Ada Xinhui Ng, Melissa Li Shan Lim, Tat Xin Ee, Mor Jack Ng, and Kok Hian Tan. 2018. The predominant learning approaches of medical students. *BMC Medical Education*, 18, 1, 1–8. ISSN: 14726920.

[11] Maureen A. Conard. 2006. Aptitude is not enough: How personality and behavior predict academic performance. *Journal of Research in Personality*, 40, 3, (June 2006), 339–346. ISSN: 00926566.

[12] Paul T. Costa, Robert R. Mccrae, and Gary G. Kay. 1995. Persons, Places, and Personality: Career Assessment Using the Revised NEO Personality Inventory. *Journal of Career Assessment*, 3, 2, 123–139. ISSN: 10690727.

[13] Sharon J. Derry. 1989. Putting learning strategies to work. *Educational Leadership*, 47, 5, 4–10.

[14] Age Diseth. 2003. Personality and Approaches to Learning as Predictors of Academic Achievement. *European Journal of Personality*, 17, 2, (March 2003), 143–155.

[15] Age Diseth and Ÿyvind Martinsen. 2003. Approaches to Learning, Cognitive Style, and Motives as Predictors of Academic Achievement. *Educational Psychology*, 23, 2, 195–207.

[16] Angus Duff, Elizabeth Boyle, Karen Dunleavy, and John Ferguson. 2004. The relationship between personality, approach to learning and academic performance. *Personality and Individual Differences*, 36, 8, 1907–1920. ISSN: 01918869.

[17] John Dunlosky, Katherine A. Rawson, Elizabeth J. Marsh, Mitchell J. Nathan, and Daniel T. Willingham. 2013. Improving Students' Learning With Effective Learning Techniques Promising Directions From Cognitive and Educational Psychology. en. *Psychological Science in the Public Interest*, 14, 1, (January 2013), 4–58.

[18] Noel Entwistle. 2007. Research into student learning and university teaching. *The British Psychological Society*, October, 1–18.

[19] Thommy Eriksson, Tom Adawi, and Christian Stöhr. 2017. "Time is the bottleneck": a qualitative study exploring why learners drop out of MOOCs. *Journal of Computing in Higher Education*, 29, 1, (April 2017), 133–146.

[20] Tom Farsides and Ruth Woodfield. 2003. Individual differences and undergraduate academic success: the roles of personality, intelligence, and application. *Personality and Individual Differences*, 34, 1225–1243.

[21] Oliver Edmund Fincham, Dragan V. Gasevic, Jelena M. Jovanovic, and Abelardo Pardo. 2018. From Study Tactics to Learning Strategies: An Analytical Method for Extracting Interpretable Representations. *IEEE Transactions on Learning Technologies*, 1382, c, 1–14. ISSN: 19391382.

[22] Adrian Furnham, Jeremy Monsen, and Gorkan Ahmetoglu. 2009. Typical intellectual engagement, big five personality traits, approaches to learning and cognitive ability predictors of academic performance. *British Journal of Educational Psychology*, 79, 4, (December 2009), 769–782. ISSN: 00070998.

[23] Dragan Gašević, Jelena Jovanović, Abelardo Pardo, and Shane Dawson. 2017. Detecting Learning Strategies with Analytics: Links with Self-Reported Measures and Academic Performance. *Journal of Learning Analytics*, 4, 2, 113–128. ISSN: 19297750.

[24] Lewis R Goldberg. 1992. The development of markers for the Big-Five factor structure. *Psychological Assessment*, 4, 1, 26–42. ISSN: 1939-134X(Electronic),1040-3590(Print).

[25] Samuel D. Gosling, Peter J. Rentfrow, and William B. Swann. 2003. A very brief measure of the Big-Five personality domains. *Journal of Research in Personality*, 37, 6, 504–528. ISSN: 00926566.

[26] Allyson F. Hadwin, John C. Nesbit, Dianne Jamieson-Noel, Jillianne Code, and Philip H. Winne. 2007. Examining trace data to explore self-regulated learning. *Metacognition and Learning*, 2, 2-3, 107–124. ISSN: 15561623.

[27] Oliver P John and Sanjay Srivastava. 1999. The Big-Five Trait Taxonomy: History, Measurement, and Theoretical Perspectives. *Handbook of personality: Theory and research*, 2, 102–138.

[28] Jelena Jovanovic, Dragan Gasevic, Shane Dawson, Abelardo Pardo, and Negin Mirriahi. 2017. Learning analytics to unveil learning strategies in a flipped classroom. *The Internet and Higher Education*, 33, 74–85. ISSN: 10967516.

[29] Jelena Jovanović, Dragan Gašević, Abelardo Pardo, Shane Dawson, and Alexander Whitelock-Wainwright. 2019. Introducing meaning to clicks: towards traced-measures of self-efficacy and cognitive load. In *Learning Analytics and Knowledge* number March, 511–520. ISBN: 9781450362566.

[30] René F Kizilcec, Mar Pérez-Sanagustín, and Jorge J Maldonado. 2017. Self-regulated learning strategies predict learner behavior and goal attainment in Massive Open Online Courses. *Computers & Education*, 104, 18–33.

[31] Rene F. Kizilcec and Emily Schneider. 2015. Motivation as a lens to understand online learners: Toward data-driven design with the OLEI scale. *ACM Transactions on Computer-Human Interaction*, 22, 2, (March 2015). ISSN: 15577325.

[32] Vitomir Kovanović, Dragan Gašević, Shane Dawson, Srećko Joksimović, Ryan S. Baker, and Marek Hatala. 2015. Penetrating the black box of time-on-task estimation. In *the Fifth International Conference on Learning Analytics And Knowledge*, 184–193. ISBN: 9781450326643.

[33] Vitomir Kovanović, Srećko Joksimović, Oleksandra Poquet, Thieme Hennis, Pieter de Vries, Marek Hatala, Shane Dawson, George Siemens, and Dragan Gašević. 2019. Examining communities of inquiry in Massive Open Online Courses: The role of study strategies. *Internet and Higher Education*, 40, (January 2019), 20–43.

[34] Chanyeong Kwak and Alan Clayton-Matthews. 2002. Multinomial Logitistic Regression. *Nursing Research*, 51, 6, 404–410.

[35] Lisa Angelique Lim, Sheridan Gentili, Abelardo Pardo, Vitomir Kovanović, Alexander Whitelock-Wainwright, Dragan Gašević, and Shane Dawson. 2019. What changes, and for whom? A study of the impact of learning analytics-based process feedback in a large course. *Learning and Instruction*. ISSN: 09594752.

[36] Griet Lust, Jan Elen, and Geraldine Clarebout. 2013. Regulation of tool-use within a blended course: Student differences and performance effects. *Computers and Education*, 60, 1, 385–395. ISSN: 03601315.

[37] Leah P. Macfadyen, Shane Dawson, Stewart Prest, and Dragan Gašević. 2016. Whose feedback? A multilevel analysis of student completion of end-of-term teaching evaluations. *Assessment and Evaluation in Higher Education*, 41, 6, (August 2016), 821–839. ISSN: 1469297X.

[38] Jorge Maldonado-Mahauad, Mar Pérez-Sanagustín, René F. Kizilcec, Nicolás Morales, and Jorge Munoz-Gama. 2018. Mining theory-based patterns from Big data: Identifying self-regulated learning strategies in Massive Open Online Courses. *Computers in Human Behavior*, 80, 179–196.

[39] Jonna Malmberg, Hanna Järvenoja, and Sanna Järvelä. 2010. Tracing elementary school students' study tactic use in gStudy by examining a strategic and self-regulated learning. *Computers in Human Behavior*, 26, 5, 1034–1042. ISSN: 07475632.

[40] Jonna Malmberg, Jarvela Sanna, and Paul A. Kirschner. 2014. Elementary school students' strategic learning: Does task-type matter? *Metacognition and Learning*, 9, 2, 113–136.

[41] Vereąová Marcela. 2015. Learning Strategy, Personality Traits and Academic Achievement of University Students. *Procedia - Social and Behavioral Sciences*, 174, (February 2015), 3473–3478.

[42] Wannisa Matcha, Noraayu Ahmad Uzir, Dragan Gasevic, and Abelardo Pardo. 2019. A Systematic Review of Empirical Studies on Learning Analytics Dashboards: A Self-Regulated Learning Perspective. *IEEE Transactions on Learning Technologies*, 1382, c, 1–1. ISSN: 1939-1382.

[43] Wannisa Matcha, Dragan Gašević, Jelena Jovanović, Abelardo Pardo, Jorge Maldonado-Mahauad, and Mar Pérez-Sanagustín. 2019. Detection of Learning Strategies: A Comparison of Process, Sequence and Network Analytic Approaches. In *European Conference on Technology Enhanced Learning*. Springer, 525–540.

[44] Wannisa Matcha, Dragan Gašević, Nora'Ayu Ahmad Uzir, Jelena Jovanović, and Abelardo Pardo. 2019. Analytics of Learning Strategies: Associations with Academic Performance and Feedback. In *Proceedings of the 9th International Conference on Learning Analytics & Knowledge*, 461–470. ISBN: 978-1-4503-6256-6.

[45] Karen Mattick, Ian Dennis, and John Bligh. 2004. Approaches to learning and studying in medical students: Validation of a revised inventory and its relation to student characteristics and performance. *Medical Education*, 38, 5, 535–543. ISSN: 03080110.

[46] Kayla Morehead, John Dunlosky, Katherine A. Rawson, Rachael Blasiman, and R. Benjamin Hollis. 2019. Note-taking habits of 21st Century college students: implications for student learning, memory, and achievement. *Memory*, 27, 6, (July 2019), 807–819. ISSN: 14640686.

[47] Gwen Nugent, Ashu Guru, and Deana Namuth-Covert. 2018. Students' Approaches to E-Learning: Analyzing Credit/Noncredit and High/Low Performers. *Interdisciplinary Journal of e-Skills and Lifelong Learning*, 14, 143–158.

[48] Melissa C. O'Connor and Sampo V. Paunonen. 2007. Big Five personality predictors of post-secondary academic performance. (October 2007).

[49] Abelardo Pardo, Dragan Gasevic, Jelena M. Jovanovic, Shane Dawson, and Negin Mirriahi. 2018. Exploring Student Interactions with Preparation Activities in a Flipped Classroom Experience. *IEEE Transactions on Learning Technologies*. ISSN: 1939-1382.

[50] K Rachal, S Daigle, and W Rachal. 2007. Learning problems reported by college students: Are they using learning strategies? *Journal of Instructional Psychology*, 34, 4, 191–199. ISSN: 00941956.

[51] Ido Roll and Philip H. Winne. 2015. Understanding, evaluating, and supporting self regulated learning using learning analytics. *Journal of Learning Analytics*, 1, 2, 7–12.

[52] O Shokri, P Kadivar, Vali Elah Farzad, and A A Sangari. 2007. Role of personality traits and learning approaches on academic achievement of university students. English. *Psychological Research*, 9, 3-4, 65–84.

[53] Ahmed Tlili, Fathi Essalmi, Mohamed Jemni, Kinshuk, and Nian Shing Chen. 2016. Role of personality in computer based learning. *Computers in Human Behavior*, 64, (November 2016), 805–813. ISSN: 07475632.

[54] Miaomiao Wen and Carolyn Penstein Rosé. 2014. Identifying latent study habits by mining learner behavior patterns in massive open online courses. In *CIKM 2014 - Proceedings of the 2014 ACM International Conference on Information and Knowledge Management*. Association for Computing Machinery, Inc, (November 2014), 1983–1986. ISBN: 9781450325981.

[55] Philip H Winne and Allyson F Hadwin. 1998. Studying as Self-Regulated Learning. *Metacognition in educational theory and practice*, 93, 277–304.

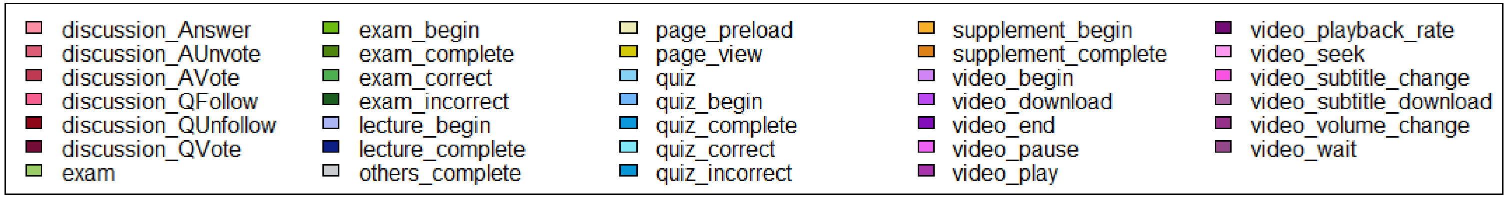
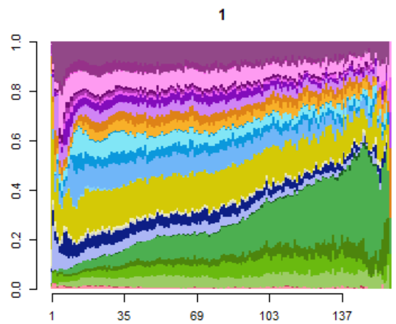
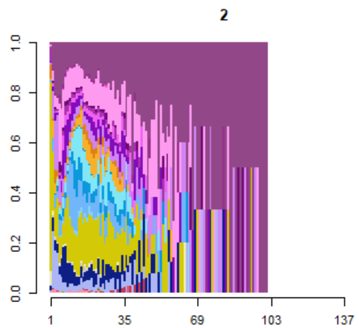
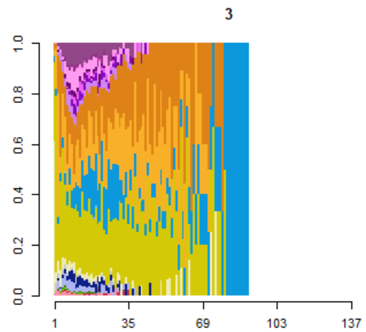
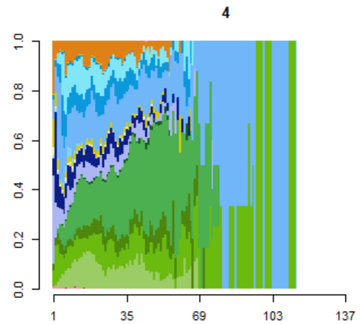
[56] Philip H. Winne, Dianne Jamieson-Noel, and Krista Muis. 2002. *Methodological issues and advances in researching tactics, strategies, and self-regulated learning*. Volume 12, 121–155. ISBN: 0762308192.

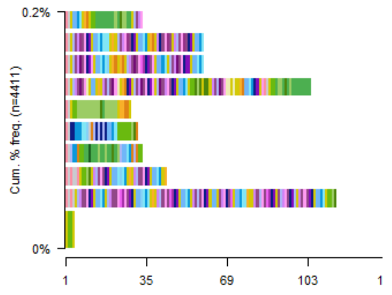
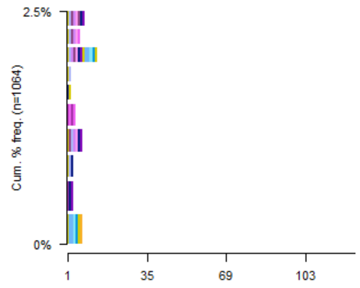
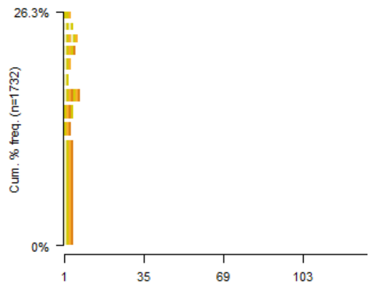
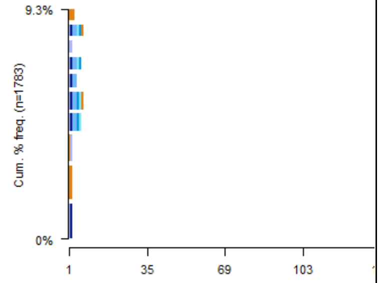
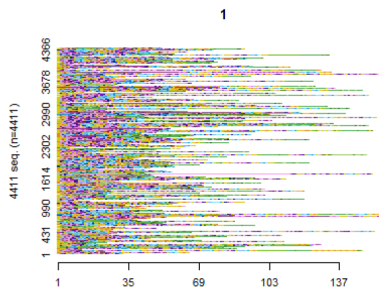
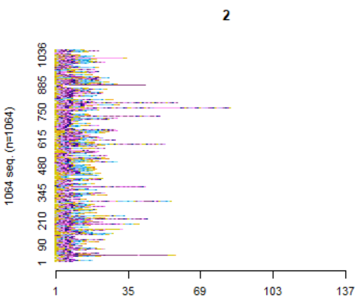
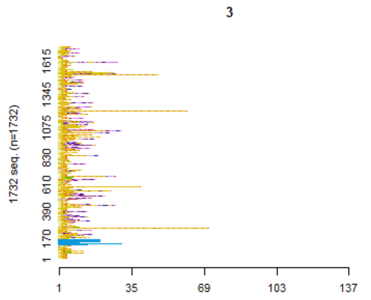
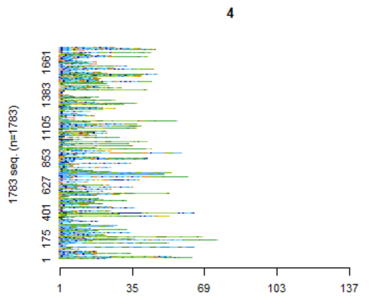
[57] Ruth Woodfield, Donna Jessop, and Lesley McMillan. 2006. Gender differences in undergraduate attendance rates. *Studies in Higher Education*, 31, 1, (February 2006), 1–22. ISSN: 03075079.

[58] Li-Fang Zhang. 2003. Does the big five predict learning approaches? *Personality and Individual Differences*, 34, 1431âĂŞ1446.

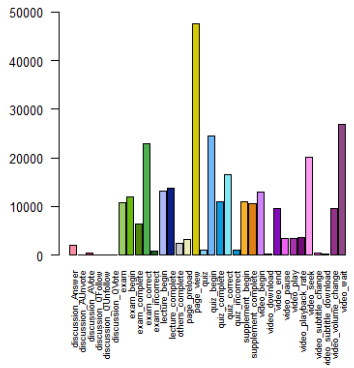
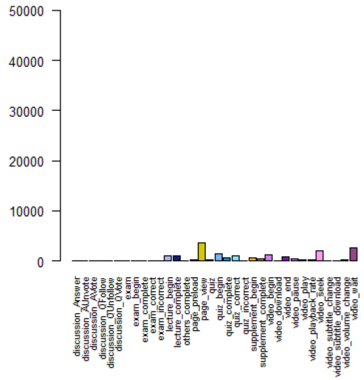
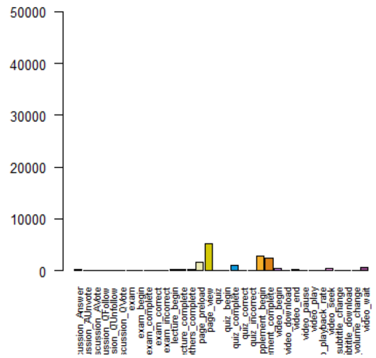
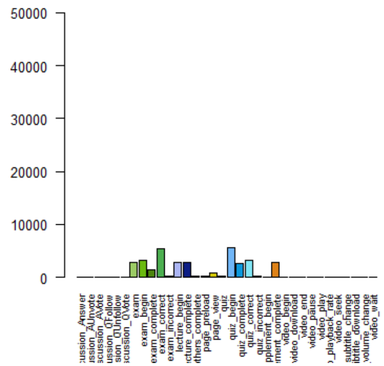
[59] Mingming Zhou and Philip H. Winne. 2012. Modeling academic achievement by self-reported versus traced goal orientation. *Learning and Instruction*, 22, 6, 413–419. ISSN: 09594752.

Supplementary Document for Analytics of Learning strategies: the association with the personality traits

<table>
<tr><th></th><th>Long Diverse</th><th>Video Oriented</th><th>Information Oriented</th><th>Assessment Oriented</th></tr>
<tr><td># Sessions</td><td>N = 4411 sessions (49.07%)</td><td>N = 1064 sessions (11.84%)</td><td>N = 1732 sessions (19.27 %)</td><td>N = 1783 sessions (19.83 %)</td></tr>
<tr><td>Legend</td><td colspan="4">
</td></tr>
<tr><td>State Distribution</td><td>
</td><td>
</td><td>
</td><td>
</td></tr>
</table>

	Long Diverse	Video Oriented	Information Oriented	Assessment Oriented
Most Frequent 10 sequence				
All Sequences				

2 | P a g e

	Long Diverse	Video Oriented	Information Oriented	Assessment Oriented
Action Frequency				
Sequence Length	Min. : 3.00 1st Qu.: 38.00 Median : 61.00 Mean : 68.38 3rd Qu.: 94.00 Max. :160.00	Min. : 2.0 1st Qu.: 11.0 Median : 16.0 Mean : 17.6 3rd Qu.: 22.0 Max. :101.0	Min. : 2.000 1st Qu.: 4.000 Median : 6.000 Mean : 9.411 3rd Qu.:11.000 Max. :90.000	Min. : 2.00 1st Qu.: 8.00 Median : 17.00 Mean : 19.68 3rd Qu.: 28.00 Max. :114.00

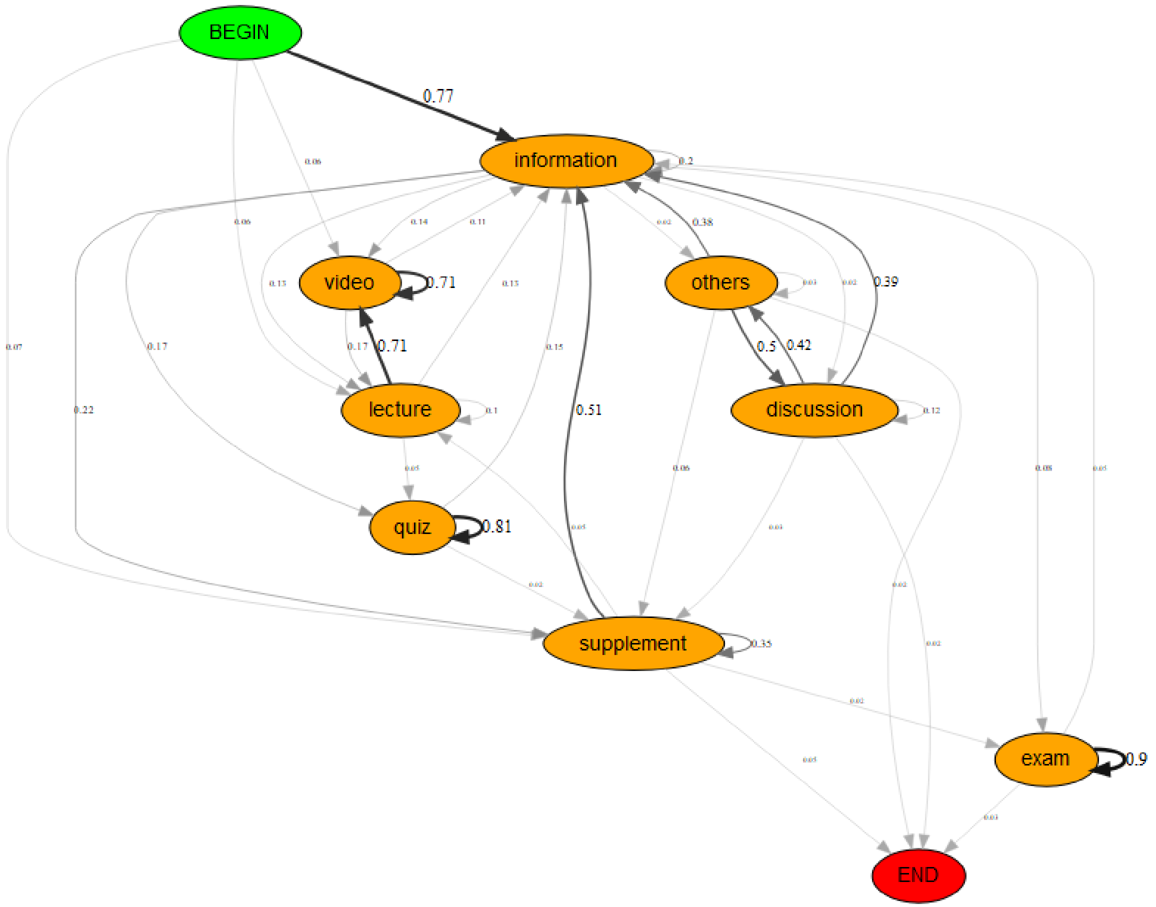


Figure 1: Process of learning tactic1: Long Diverse Oriented tactics

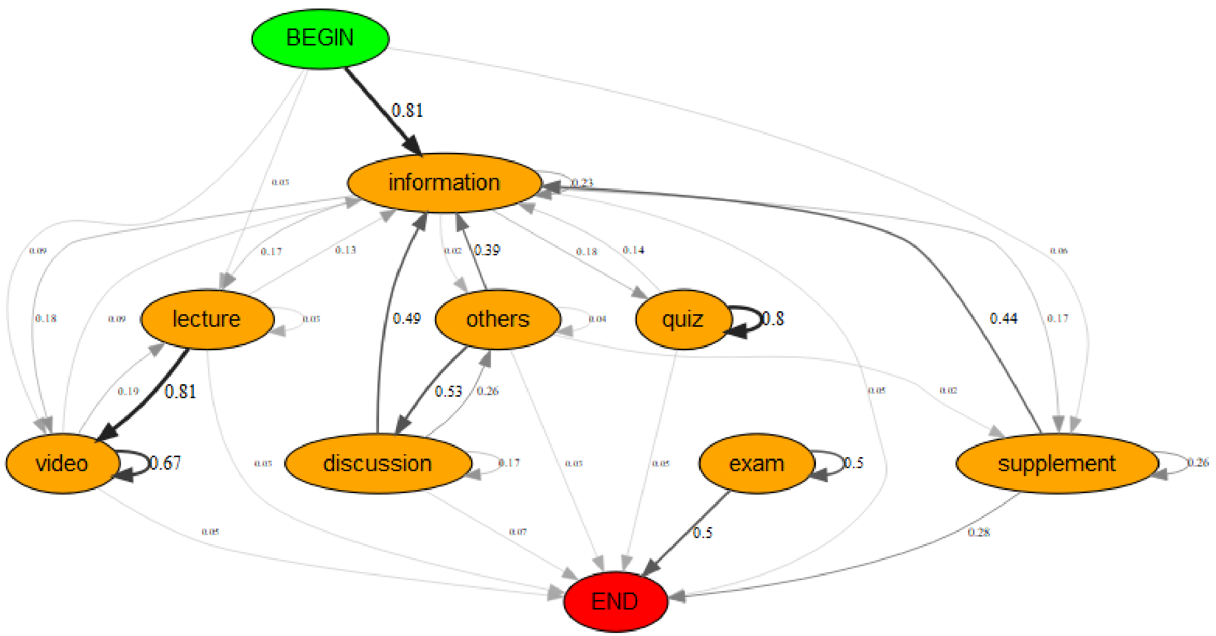


Figure 2: Process of learning tactic2: Video Oriented tactics

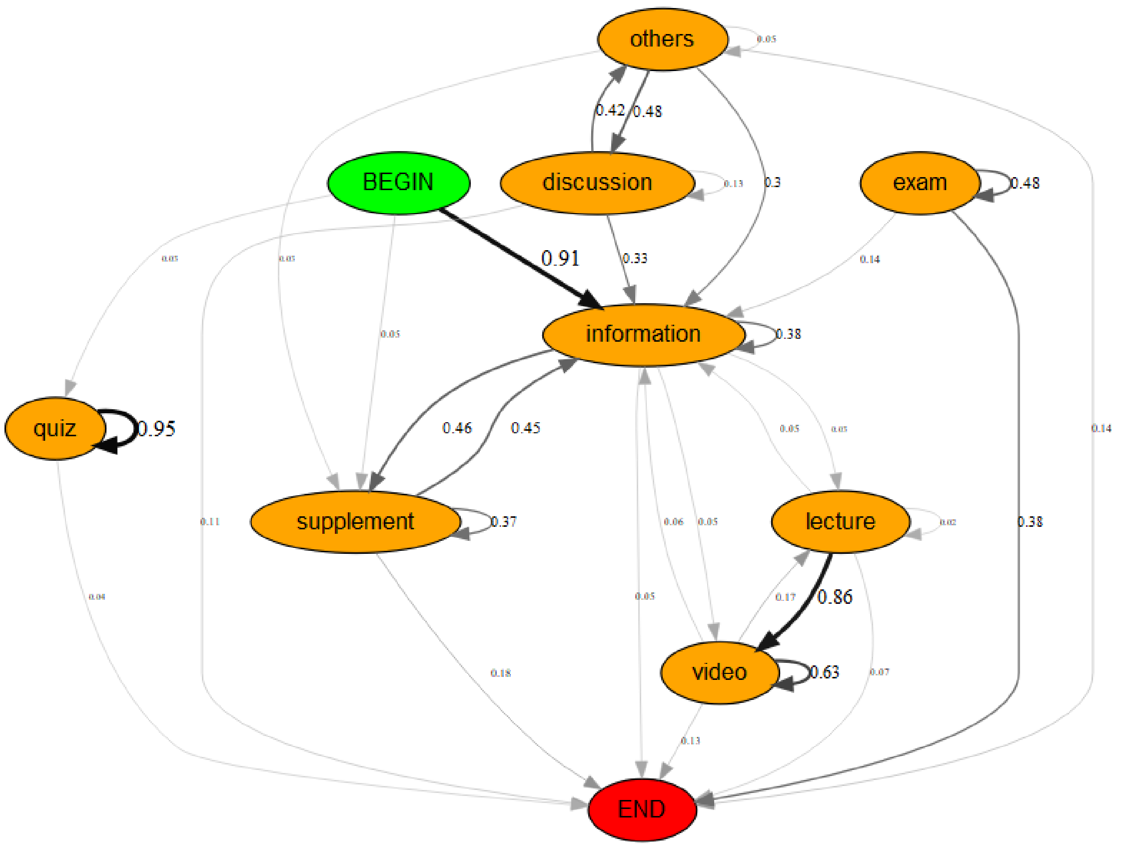


Figure 3: Process of learning tactic 3: Information Oriented tactics

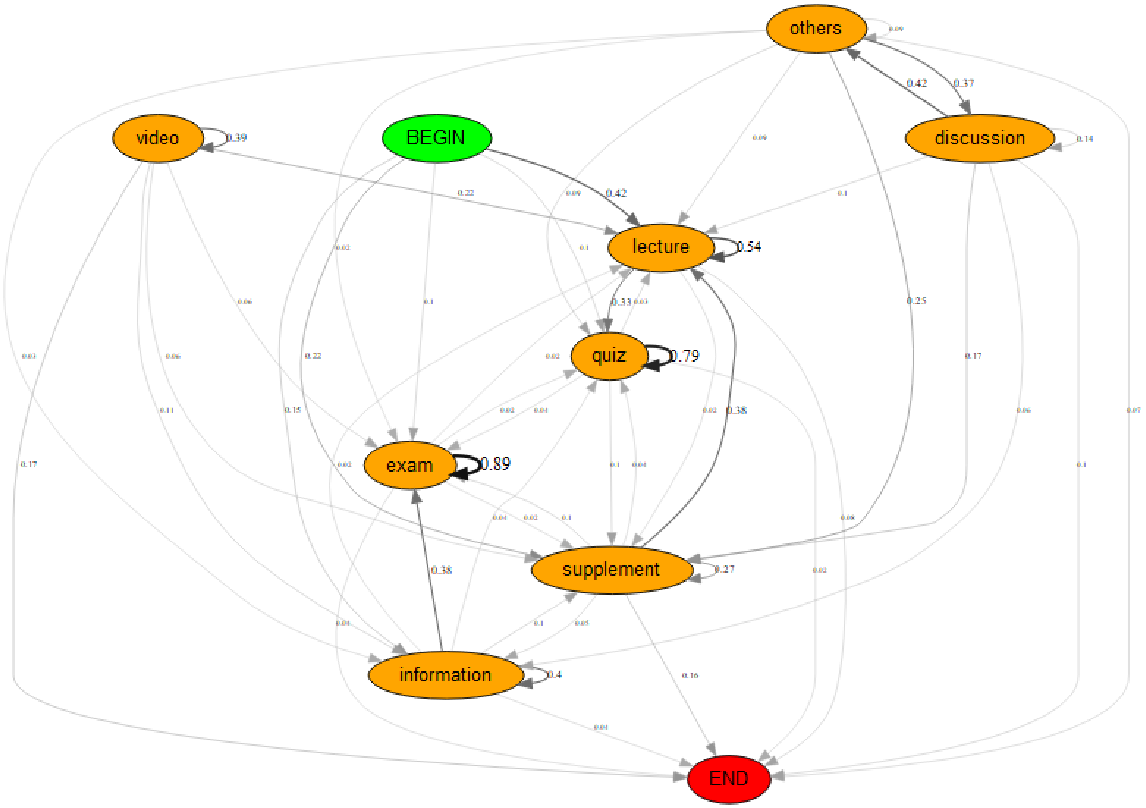


Figure 4: Process of learning tactic4: Assessment Oriented tactic

Table 2: The Association between personality and learning strategy based on the multinomial logistic regression

Reference	Strategy Group	Personality	Exp(B)	CI(95%) of Exp(B)		p	B
Disengage	Highly Active	Extraversion	1.002	0.902	1.113	0.97	0.002
		Agreeableness	1.156*	0.999	1.338	0.052	0.145*
		Conscientiousness	1.143**	1.007	1.298	0.039	0.134**
		Emotional instability	0.855**	0.755	0.967	0.013	-0.157**
		Openness	1.088	0.936	1.264	0.271	0.084
	Active	Extraversion	0.925*	0.847	1.011	0.086	-0.078*
		Agreeableness	0.996	0.882	1.125	0.949	-0.004
		Conscientiousness	1.039	0.935	1.154	0.477	0.038
		Emotional instability	1.069	0.962	1.187	0.217	0.066
		Openness	1.067	0.941	1.210	0.31	0.065
	Surface	Extraversion	1.031	0.938	1.133	0.523	0.031
		Agreeableness	1.223***	1.070	1.399	0.003	0.201***
		Conscientiousness	0.959	0.859	1.072	0.462	-0.042
		Emotional instability	0.992	0.886	1.110	0.886	-0.008
		Openness	0.946	0.828	1.081	0.415	-0.055
Surface	Highly Active	Extraversion	0.972	0.869	1.086	0.611	-0.029
		Agreeableness	0.945	0.807	1.107	0.485	-0.056
		Conscientiousness	1.191**	1.042	1.362	0.010	0.175**
		Emotional instability	0.862**	0.756	0.983	0.027	-0.149**
		Openness	1.150*	0.981	1.348	0.084	0.140*
	Active	Extraversion	0.897**	0.815	0.988	0.027	-0.108**
		Agreeableness	0.814***	0.710	0.933	0.003	-0.205***
		Conscientiousness	1.083	0.967	1.214	0.170	0.08
		Emotional instability	1.078	0.960	1.209	0.204	0.075
		Openness	1.128*	0.984	1.293	0.083	0.121*

Reference	Strategy Group	Personality	Exp(B)	CI(95%) of Exp(B)		p	B
	Disengage	Extraversion	0.97	0.882	1.066	0.523	-0.031
		Agreeableness	0.818***	0.715	0.935	0.003	-0.201***
		Conscientiousness	1.042	0.933	1.165	0.462	0.042
		Emotional instability	1.008	0.901	1.128	0.886	0.008
		Openness	1.057	0.925	1.208	0.415	0.055
Active	Highly Active	Extraversion	1.083	0.974	1.204	0.142	0.08
		Agreeableness	1.161**	1.001	1.347	0.049	0.149**
		Conscientiousness	1.1	0.966	1.252	0.149	0.095
		Emotional instability	0.800***	0.705	0.907	0.001	-0.223***
		Openness	1.019	0.875	1.187	0.805	0.019
	Surface	Extraversion	1.115**	1.012	1.227	0.027	0.108**
		Agreeableness	1.228***	1.071	1.408	0.003	0.205***
		Conscientiousness	0.923	0.824	1.035	0.170	-0.08
		Emotional instability	0.928	0.827	1.041	0.204	-0.075
		Openness	0.886*	0.773	1.016	0.083	-0.121*
	Disengage	Extraversion	1.081*	0.989	1.181	0.086	0.078*
		Agreeableness	1.004	0.889	1.134	0.949	0.004
		Conscientiousness	0.963	0.866	1.069	0.477	-0.038
		Emotional instability	0.936	0.842	1.040	0.217	-0.066
		Openness	0.937	0.826	1.062	0.310	-0.065
Highly Active	Active	Extraversion	0.923	0.830	1.027	0.0262	-0.08
		Agreeableness	0.861**	0.743	0.999	0.1421	-0.149**
		Conscientiousness	0.909	0.799	1.035	0.049	-0.095
		Emotional instability	1.250***	1.102	1.418	0.1488	0.223***
		Openness	0.981	0.842	1.143	0.0005	-0.019
	Surface	Extraversion	1.029	0.921	1.150	0.1194	0.029
		Agreeableness	1.058	0.903	1.239	0.6113	0.056
		Conscientiousness	0.839**	0.734	0.959	0.4851	-0.175**

Reference	Strategy Group	Personality	Exp(B)	CI(95%) of Exp(B)		p	B
		Emotional instability	1.160**	1.017	1.324	0.0103	0.149**
		Openness	0.870*	0.742	1.019	0.027	-0.140*
	Disengage	Extraversion	0.998	0.899	1.108	0.0007	-0.002
		Agreeableness	0.865*	0.747	1.001	0.9705	-0.145*
		Conscientiousness	0.875**	0.771	0.993	0.052	-0.134**
		Emotional instability	1.170**	1.034	1.324	0.039	0.157**
		Openness	0.919	0.791	1.068	0.013	-0.084

Note: *p<0.01; **p<0.05; ***p<0.001

6.3 Summary

In this chapter, the process analytics-based approach is applied to detect learning tactics and strategies of MOOC learners. Given the flexibility afforded by a MOOC in terms of free and open enrolment, it was challenging to examine patterns of learning tactics that would correspond to particular temporal units (weeks) of the course. Hence, the course topics were used as a time unit of analysis. To address research question three (RQ3), which aimed to ensure that the analysed results are consistent with educational theory, this chapter demonstrated that the process analytics-based approach is capable of detecting study patterns indicative of learning tactics and strategies. Aligned with the findings presented in other chapters, the learning tactics detected using the process analytics-based approach are reflective of the course design. These learning tactics are *Long Diverse, Video Oriented, Information Oriented*, and *Assessment Oriented* tactics. This approach extracted four strategies of how learners employed the tactics, including *Disengaged, Surface, Active*, and *Highly Active* strategies. Two of the strategies (*Disengaged* and *Surface*) were reflective of the surface approach to learning, which is characterised by a low level of engagement, high focus on assessment without understanding the learning content, and is associated with poor academic performance (Mattick et al., 2004). The other two detected learning strategies (*Active* and *Highly Active* strategies) were indicative of the deep approach to learning. The students who employed these two strategies showed a high level of engagement, a tendency to apply multiple learning tactics, and an association with high academic performance (Chonkar et al., 2018).

This study serves as one of the first examples that investigated the relationship between personality traits and learning strategies extracted from trace data. The findings from this study, thus, answer research question four (RQ4) by observing how the SRL construct (i.e. personality traits) contributes to the adoption of learning strategies. That is, in this study, we found that the conscientiousness trait was predictive of deep approaches to learning (i.e., *Active* and *Highly Active* strategies based on the automatically detected learning strategies). This finding is aligned with the current research that explore the relationship between conscientiousness trait and deep approaches to learning by using self-reports to capture learning strategies (Duff et al., 2004; Zhang, 2003). The literature reports that agreeableness is predictive of both surface and deep approaches to learning (Chamorro-Premuzic & Furnham, 2008; Zhang, 2003). Similarly, in this study, significant associations were found between agreeableness and the surface strategy as well as agreeableness and the *Active* strategy group. The students who scored high on their self-reports about the emotional instability trait were found to be less likely to adopt *Highly Active* strategy. The extraversion trait was found to be associated with high odds of using the surface approach to learning. There was no significant association between the openness and any of the detected learning strategies.

The results presented in this chapter informed that personality traits are predictive of the adoption of learning strategies. The alignment of detected results with the well-accepted educational research on approaches to learning (Diseth, 2003; Duff et al., 2004; Zhang, 2003) lends some support for the generalizability of the results. That is, conscientiousness is predictive of deep approach application. Meanwhile, emotional instability is connected to the adaption of the surface approach to learning. However, the link between agreeableness requires further research. This alignment of the results in this thesis and the results in the literature on approaches to learning also offers some support for the validity of the proposed analytics-based approach for the detection of theoretically-meaningful learning strategies.

This study has shed light on the understanding of the link between psychological constructs and automatically detected learning strategies. The findings suggest that personality traits are predictive of learning strategies adopted by learners in the MOOC. Hence, personality traits can be used as an initial step to observe the learning strategies that learners are inclined to use, especially when trace data are not available. Self-reports about personality traits can be used to generate personalised feedback on the selection of learning tactics and strategies in the initial phases of a course. However, personality traits should not be used as an assessment of one's learning process nor should they be used throughout the entire duration of a course. Instead, trace data should be used as they are shown to offer more reliable accounts about how students learn (Cicchinelli et al., 2018; Hadwin et al., 2007; Zhou & Winne, 2012).

7 Conclusions and future directions

> Everything is possible. The impossible just takes longer.
>
> — Dan Brown, *Digital Fortress*

THE main intention of this thesis is to propose an approach for automatic detection of learning tactics and strategies in order to enable the process of analytics-based feedback provision and evaluation. This thesis has presented the findings of multiple studies that applied analytics approaches to the datasets collected from several learning environments.

In this chapter, a summary of the main findings and contributions of the work presented in this thesis are discussed with respect to the key research goals and questions identified in Section 1.1. Given that a key purpose of learning analytics is to understand and optimise human learning, this chapter focuses on the implication of work for both research and practice. The potential directions for future work are also discussed.

7.1 Impact of the present work

7.1.1 RQ 1: Feedback on self-regulated learning

In Chapter two, we have presented a systematic review of the literature on learning analytics-based feedback tools from the SRL perspective. Specifically, the review offered an overview of the research on LADs. Serving also as a background study for this PhD thesis, the systematic literature review highlights the need to support the understanding and enhancement of SRL processes with learning analytics.

This systematic review has several key strengths. Firstly, the review was conducted by following a well-known theoretical model of SRL. Previous systematic literature reviews in this area focused on the design and development aspects of LADs with weak grounding in educational theories. The review shows that the current generation of LADs lacks the ability to support all the key elements and phases of self-regulation according to SRL theories (Matcha, Ahmad Uzir, et al., 2020). The study, thus, has identified a need for future research and development on LADs; that is, to develop mechanisms that can include support for learning tactics, learning strategies, knowledge of tasks, standards, and evaluations.

Secondly, the findings of the systematic literature review suggest promising directions that can be followed to overcome the challenges in existing LADs. For instance, if LADs are considered as a form of feedback, their design should be driven by the relevant feedback literature (Hattie & Timperley, 2007). Moreover, LADs should be supported with the theory-informed use of data science approaches to analyse relevant constructs such as learning tactics and strategies.

7.1.2 RQ 2: Detection of learning tactics and strategies

The central and most substantial part of the thesis is dedicated to the development of a learning analytics-based approach that enables the detection of theoretically sound learning tactics and strategies from trace data. Learning tactics and strategies detected from trace data conceptually represent "realised intentions" of students, while learning tactics and strategies identified with self-report instruments can be referred to as "perceived intentions" (Zhou & Winne, 2012). In this thesis, learning tactics are differentiated from learning strategies. That is, a tactic is a sequence of actions performed to complete a given task (Hadwin et al., 2007). A strategy is viewed as regulation or pattern of tactic applications (Derry, 1989; Malmberg et al., 2014; Rachal et al., 2007). By distinguishing learning tactics from strategies, the thesis proposes a data-analytics approach composes of two main steps, including tactic detection and strategy detection. Chapter three has examined three different learning analytics-based approaches for the detection of learning tactics and strategies. The following chapters (Chapter four to Chapter six) investigated the ability of process analytics-based approach to detect learning tactics and strategies. The primary goal of these studies was to validate the applicability of the proposed approach.

Validity is a critical property of a learning analytics-based approach. Validity can be viewed as the extent to which the analysed findings can be explained by existing theories (Gašević et al., 2015; Joksimovic et al., 2019). The extent to which the approach is applicable in a different situation can also be considered as one of the dimension to assess validity. The validity of the data analytics approach proposed in this thesis was tested by replicating the application of the approach to datasets collected in different learning environments including flipped classroom (in Chapter four and Chapter five), blended learning (in Chapter five), session-focused MOOC (in Chapter five) and on-demand MOOC (in Chapter six).

From the practical perspective, the most significant implication of the work presented in this thesis is the development of the data analytics approach that can be used by practitioners and instructors as a background algorithm to automatically detect learning tactics and strategies employed by students. The detected tactics and strategies can inform personalised feedback that supports each student according to their individual needs. Moreover, the results from the application of the proposed data analytics approach can be used to inform the design of interventions (Cicchinelli et al., 2018). For instance, learning tactics are reflective of course instructional design; that is, it allows practitioners to use learning analytics to observe process models, actions, and effort of students. With these learning analytics, practitioners can check if students' activities echo with their pedagogical indent and can get insights that can inform redesign of existing and design of new learning tasks. Learning strategies present patterns of utilising learning tactics. The observation of patterns of learning tactics and strategies allows the practitioners to identify whether particular types of learning tactics tend to be employed in relation to particular weeks/topics. In this way, instructors can better understand how students proceed with learning activities and if they act according to the course designs. This observation, hence, can help instructors to detect students who need support and require attention, as demonstrated in Chapter five.

7.1.3 RQ 3: theory informed learning tactics and strategies

From the research perspective, the absence of theory has been identified as a critical challenge in the current learning analytics research (Gašević, Kovanović, et al., 2017; Wise & Shaffer, 2015). In this regard, the focal point of the thesis is to ensure that the studies conducted were informed by relevant theories. Approaches to learning (Biggs, 1987; Entwistle, 1991) is a popular theory on learning strategies and is widely used to explain learning strategies in previous research (Zeegers, 2001). Therefore, in this thesis, the findings are discussed with relation to approaches to learning as demonstrated in Chapter three to Chapter six.

From the practical perspective, the proposed data analytics approach for the detection of learning tactics and strategies has been extensively evaluated to check its validity. The findings suggest that the approach can robustly be used across different learning contexts. As such, the thesis has demonstrated evidence that supports the generalisability of the results produced by the use of the analytics approach.

7.1.4 RQ 4: association of the self-regulated learning constructs

Enactment of learning tactics and strategies is identified as the core process of the SRL (Zimmerman, 2011). SRL is a recursive learning process that is influenced by many factors (Greene & Azevedo, 2007; Winne & Hadwin, 1998). One of the key research questions in this thesis is to examine how learning tactics and strategies detected automatically from trace data with the proposed data analytics approach are associated with relevant SRL constructs. Four of the learning constructs are examined in this thesis, including, products, external evaluation (feedback), cognitive conditions, and task conditions. As shown in the analyses presented in Chapter three to Chapter six, high academic performance as a proxy of learning products tends associate with the application of particular types of learning strategies that indicative of deep and strategic approaches to learning. In Chapter four, the association of learning strategies and feedback is examined. The results indicate that personalised feedback based on learning analytics is significantly and positively associated with the application of effective learning strategies. Chapter five examined the role of task conditions in the selection of learning tactics and strategies. The results suggested that learning tactics are reflective of the course instructional design, whereas learning strategies are sensitive to the delivery modality. In Chapter six, one of the cognitive conditions (disposition) is examined. Disposition is considered a psychological construct. In this study, dispositions are determined by the personality traits of an individual. The results indicate that learning strategies, detected by the proposed data analytics approach, show associations with personality traits, similar as findings found in studies that use the traditional self-report methods.

Several implications can be drawn from the findings related to this research question. First of all, its association with the selection of an effective learning strategy suggests that the provision of learning analytics-based personalised feedback, such as a customised elaborated textual message, is a promising approach for the use of learning analytics. Second, not only can data analytics approaches be used to detect learning tactics and strategies from trace data, but they can also be used to understand the efficacy of the opportunities provided to the students. That is, the patterns of learning tactics and strategies can be used to obtain insights into student interactions with learning materials. Together with the students' performance, practitioners can observe if the level of efforts, types of actions, and activities are sufficient to support learning. Specifically, we can obtain insights to inform the design of courses. Finally, the self-reports of personality traits can be used as a starting point when trace data is not available to overcome the problems potentially associated with the 'cold start' (e.g., at the beginning of a course).

7.2 Directions for future work

Several avenues of research are left open to be explored further in the future. The main goal of this research is to propose a data analytics approach that enables the detection of learning tactics and strategies to support the provision of personalised feedback to optimise learning. Hence, an important direction of future work lies in exploring ways to integrate the proposed data analytics approach into systems for feedback provision. In Chapter four, we have shown that learning analytics-based personalised feedback is a promising approach to providing a 'skeleton' for instructors towards promoting dialogic feedback (Yang & Carless, 2013). That is, rather than providing feedback that is automatically generated by a system which has been proven as ineffective (Gašević et al., 2016; Kizilcec et al., 2016), a personalised feedback system such as 'OnTask' developed by Pardo et al. (2017) offers templates for instructors to create rules that can personalise messages sent to each learner. The rules give the flexibility to the instructor to make use of relevant measures of learning. Therefore, the proposed data analytics approach for the detection of learning tactics and strategies can be employed in such a system to enable personalised feedback on learning tactics and strategies. Future research should investigate the efficacy of the use of the proposed data analytics approach for learning tactic and strategy detection in the provision of personalised feedback.

The use of learning analytics-based feedback is promising to scale support for personalised feedback. However, several elements that can have significant impact on the effectiveness of feedback are yet under-explored. For instance, Hattie and Timperley (2007) posit that in order for feedback to be effective, feedback needs to address the "process" and "regulation" of learning. Additionally, in Chapter six, it is suggested that questionnaires such as the one of assessment of personality traits can be used to generate personalised feedback on the selection of learning tactics and strategies in the initial phases when trace data are still unavailable. However, the way to provide feedback to support the adoption of effective learning tactics and strategies is yet unexplored. We argue that individual differences, as identified by the personality traits, need to be considered when developing feedback for this purpose. For example, the findings in Chapter six suggest that those who have high emotional instability scores tended to use the surface approach to learning. The surface approach is characterised by low level of engagement and intention to focus mainly on assessment (Biggs, 1987). Hence, feedback for this group of students might need to highlight the importance of an in-depth understanding of a concept that the students need to learn. Moreover, the feedback can include suggestions about suitable tactics to apply in order to understand the concept. Furthermore, motivational factors should be considered and embedded in feedback in order to encourage the students to be persistent in the application of appropriate learning tactics and strategies as well as to motivate them to engage with the learning activities (Moos, 2014).

Given that the research in this thesis is highly focused on trace data, future research should investigate how other data sources such as self-reports and eye-tracking might add more insights into learning tactics and strategies. By combining trace data with other data streams, future research may deeper our understanding of the reasons why students employ particular tactics and strategies in different situations.

Although trace data offer accounts on how learning actually happened in real learning situations, many learning activities cannot be captured by digital technologies. Especially, in the formal learning settings where face-to-face classes are mandatory, physical learning activities play large parts in the learning process. Hence, capturing data about learning activities that happen outside digital spaces is one of the most challenging but important research direction. Combining multimodal data with trace data as used in this thesis is an important research direction to expand our understanding of learning tactics and strategies.

Finally, SRL skills have been identified as essential skills for learning in the digital era (Kizilcec et al., 2017). The COPES model proposed by Winne and Hadwin (1998) is a well-established theoretical model of SRL that offers a holistic view of the learning process. Several dimensions of SRL are left under-explored in this thesis, including the phases of the SRL (i.e., goal setting and planning, task identification, and adaptation phases). Precisely, how to provide support for goal setting, assessment of standards, support for meta-cognitive, control and monitoring of the learning process are important directions for future research.

Bibliography

Abbeel, P., & Ng, A. Y. (2005). Learning first-order Markov models for control, In *Advances in neural information processing systems*.

Abrami, P. C., Bernard, R. M., Bures, E. M., Borokhovski, E., & Tamim, R. M. (2011). Interaction in distance education and online learning: using evidence and theory to improve practice. *Journal of Computing in Higher Education*, *23*(2-3), 82–103. https://doi.org/10.1007/s12528-011-9043-x

Armstrong, R., Hall, B. J., Doyle, J., & Waters, E. (2011). 'Scoping the scope' of a cochrane review. Oxford University Press. https://doi.org/10.1093/pubmed/fdr015

Baker, R. S. (2019). Challenges for the Future of Educational Data Mining: The Baker Learning Analytics Prizes. *Journal of Educational Data Mining*, *11*(1), 1–17.

Bannert, M., Reimann, P., & Sonnenberg, C. (2014). Process mining techniques for analysing patterns and strategies in students' self-regulated learning. *Metacognition and Learning*, *9*(2), 161–185. https://doi.org/10.1007/s11409-013-9107-6

Bidjerano, T., & Dai, D. Y. (2007). The relationship between the big-five model of personality and self-regulated learning strategies. *Learning and Individual Differences*, *17*(1), 69–81. https://doi.org/10.1016/j.lindif.2007.02.001

Biggs. (1987). *Student Approaches to Learning and Studying*. https://eric.ed.gov/?id=ED308201

Bjork, R. a., Dunlosky, J., & Kornell, N. (2013). Self-Regulated Learning: Beliefs, Techniques, and Illusions. *Annual Review of Psychology*, *64*(1), 120928131529005. https://doi.org/10.1146/annurev-psych-113011-143823

Bodily, R., & Verbert, K. [K.]. (2017). Review of Research on Student-Facing Learning Analytics Dashboards and Educational Recommender Systems. *IEEE Transactions on Learning Technologies*, *10*(4), 405–418. https://doi.org/10.1109/TLT.2017.2740172

Bodily, R., & Verbert, K. [Katrien]. (2017). Trends and issues in student-facing learning analytics reporting systems research. *Proceedings of the Seventh International Learning Analytics & Knowledge Conference on - LAK '17*, *1*(212), 309–318. https://doi.org/10.1145/3027385.3027403

Boroujeni, M. S., & Dillenbourg, P. (2018). Discovery and Temporal Analysis of Latent Study Patterns in MOOC Interaction Sequences, In *Proceedings of the 8th international conference on learning analytics and knowledge*, New York, ACM. https://doi.org/10.1145/3170358.3170388

Bozionelos, N. (2004). The big five of personality and work involvement. *Journal of Managerial Psychology*, *19*(1), 69–81. https://doi.org/10.1108/02683940410520664

Broadbent, J. (2017). Comparing online and blended learner's self-regulated learning strategies and academic performance. *The Internet and Higher Education*, *33*, 24–32. https://doi.org/10.1016/j.iheduc.2017.01.004

Byrne, M., Flood, B., & Willis, P. (2010). The relationship between learning approaches and learning outcomes : a study of Irish accounting students The relationship between learning approaches and learning outcomes : a study of Irish, *9284*. https://doi.org/10.1080/0963928021015325

Carpenter, S. K., Witherby, A. E., & Tauber, S. K. (2020). On Students' (Mis)judgments of Learning and Teaching Effectiveness. *Journal of Applied Research in Memory and Cognition*. https://doi.org/10.1016/j.jarmac.2019.12.009

Chamorro-Premuzic, T., & Furnham, A. (2008). Personality, intelligence and approaches to learning as predictors of academic performance. *Personality and Individual Differences*, *44*(7), 1596–1603. https://doi.org/10.1016/j.paid.2008.01.003

Chamorro-Premuzic, T., Furnham, A., & Lewis, M. (2007). Personality and approaches to learning predict preference for different teaching methods. *Learning and Individual Differences*, *17*(3), 241–250. https://doi.org/10.1016/j.lindif.2006.12.001

Chonkar, S. P., Ha, T. C., Chu, S. S. H., Ng, A. X., Lim, M. L. S., Ee, T. X., Ng, M. J., & Tan, K. H. (2018). The predominant learning approaches of medical students. *BMC Medical Education*, *18*(1), 1–8. https://doi.org/10.1186/s12909-018-1122-5

Cicchinelli, A., Veas, E., Pardo, A., Pammer-Schindler, V., Fessl, A., Barreiros, C., & Lindstädt, S. (2018). Finding traces of self-regulated learning in activity streams. *Proceedings of the 8th International Conference on Learning Analytics and Knowledge - LAK '18*, 191–200. https://doi.org/10.1145/3170358.3170381

Conard, M. A. (2006). Aptitude is not enough: How personality and behavior predict academic performance. *Journal of Research in Personality*, *40*(3), 339–346. https://doi.org/10.1016/j.jrp.2004.10.003

Corrin, L., & de Barba, P. (2015). How Do Students Interpret Feedback Delivered via Dashboards? *International Conference on Learning Analytics and Knowledge*, 430–431. https://doi.org/10.1145/2723576.2723662

Derry, S. J. (1989). Putting learning strategies to work. *Educational Leadership*, *47*(5), 4–10.

DiFrancesca, D., Nietfeld, J. L., & Cao, L. (2016). A comparison of high and low achieving students on self-regulated learning variables. *Learning and Individual Differences*, *45*, 228–236. https://doi.org/10.1016/j.lindif.2015.11.010

Diseth, A. (2003). Personality and Approaches to Learning as Predictors of Academic Achievement. *European Journal of Personality*, *17*(2), 143–155. https://doi.org/10.1002/per.469

Duff, A., Boyle, E., Dunleavy, K., & Ferguson, J. (2004). The relationship between personality, approach to learning and academic performance. *Personality and Individual Differences*, *36*(8), 1907–1920. https://doi.org/10.1016/j.paid.2003.08.020

Dunlosky, J. (2013). Strengthening the Student Toolbox. *American Educator*, *37*(3), 12–21. http://www.aft.org/sites/default/files/periodicals/dunlosky.pdf

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving Students' Learning With Effective Learning Techniques Promising Directions From Cognitive and Educational Psychology. *Psychological Science in the Public Interest*, *14*(1), 4–58. https://doi.org/10.1177/1529100612453266

Elen, J., & Clarebout, G. (2005). Support in Learning Environments: Touching the Limits of Instructional Design. *Educational Technology*, *45*(5), 44–47.

Emilia, O., Bloomfield, L., & Rotem, A. (2012). Measuring students ' approaches to learning in different clinical rotations. *BMC Medical Education*, *12*(1), 1. https://doi.org/10.1186/1472-6920-12-114

Entwistle, N. J. (1991). Approaches to Learning and Perceptions of the Learning Environment : Introduction to the Special Issue. *Higher Eduation*, *22*(3), 201–204. https://doi.org/10.1007/BF00132287

Entwistle, N. J. (2007). Research into student learning and university teaching. *The British Psychological Society*, (October), 1–18. https://doi.org/10.1348/000709906X166772

Farsides, T., & Woodfield, R. (2003). Individual differences and undergraduate academic success: the roles of personality, intelligence, and application. *Personality and Individual Differences*, (34), 1225–1243. www.elsevier.com/locate/paid

Ferreira, D., Zacarias, M., Malheiros, M., & Ferreira, P. (2007). Approaching Process Mining with Sequence Clustering: Experiments and Findings. *Business Process Management*, (1), 360–374. https://doi.org/10.1007/978-3-540-75183-0{_}26

Fincham, O. E., Gasevic, D. V., Jovanovic, J. M., & Pardo, A. (2018). From Study Tactics to Learning Strategies: An Analytical Method for Extracting Interpretable Representations. *IEEE Transactions on Learning Technologies*, *1382*(100), 1–14. https://doi.org/10.1109/TLT.2018.2823317

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences*, *111*(23), 8410–8415. https://doi.org/10.1073/pnas.1319030111

Furnham, A., Monsen, J., & Ahmetoglu, G. (2009). Typical intellectual engagement, big five personality traits, approaches to learning and cognitive ability predictors of academic performance. *British Journal of Educational Psychology*, *79*(4), 769–782. https://doi.org/10.1348/978185409X412147

Gabadinho, A., Ritschard, G., Mueller, N. S., & Studer, M. (2011). Analyzing and Visualizing State Sequences in R with TraMineR. *Journal of Statistical Software*, *40*(4), 1–37. https://doi.org/10.18637/jss.v040.i04

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

Gašević, D., Dawson, S., & Siemens, G. (2015). Let's not forget: Learning analytics are about learning. *TechTrends*, *59*(1). https://doi.org/10.1007/s11528-014-0822-x

Gašević, D., Jovanović, J., Pardo, A., & Dawson, S. (2017). Detecting Learning Strategies with Analytics: Links with Self-Reported Measures and Academic Performance. *Journal of Learning Analytics*, *4*(2), 113–128. https://doi.org/10.18608/jla.2017.42.10

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), 63–78. https://doi.org/10.1080/23735082.2017.1286142

Gatta, R., Lenkowicz, J., Vallati, M., & Stefanini, A. (2017). pMineR: Processes Mining in Medicine. https://cran.r-project.org/package=pMineR

Ghadiri, K., Qayoumi, M., Junn, E., Hsu, P., & Sujitparapitaya, S. (2013). The transformative potential of blended learning using MIT edX's 6.002 x online MOOC content combined with student team-based learning in class. *Environment*, *8*, 14.

Greene, J. A., & Azevedo, R. (2007). A Theoretical Review of Winne and Hadwin's Model of Self-Regulated Learning: New Perspectives and Directions. *Review of Educational Research*, *77*(3), 334–372. https://doi.org/10.3102/003465430303953

Hadwin, A. F., Nesbit, J. C., Jamieson-Noel, D., Code, J., & Winne, P. H. (2007). Examining trace data to explore self-regulated learning. *Metacognition and Learning*, *2*(2-3), 107–124. https://doi.org/10.1007/s11409-007-9016-7

Hassani, M., van Zelst, S. J., & van der Aalst, W. M. (2019). On the application of sequential pattern mining primitives to process discovery: Overview, outlook and opportunity identification. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *9*(6). https://doi.org/10.1002/widm.1315

Hattie, J., & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research*, *77*(1), 81–112. https://doi.org/10.3102/003465430298487

Jahan, N., Naveed, S., Zeshan, M., & Tahir, M. A. (2016). How to Conduct a Systematic Review: A Narrative Literature Review. *Cureus*. https://doi.org/10.7759/cureus.864

Jivet, I., Scheffel, M., Drachsler, H., & Specht, M. (2017). Awareness is not enough. Pitfalls of learning analytics dashboards in the educational practice.

John, O. P., & Srivastava, S. (1999). The Big-Five Trait Taxonomy: History, Measurement, and Theoretical Perspectives. *Handbook of personality: Theory and research*, *2*, 102–138.

Joksimovic, S., Kovanovic, V., & Dawson, S. (2019). The Journey of Learning Analytics. *HERDSA Review of Higher Education*, *6*(January), 37–63. www.herdsa.org.au/herdsa-review-higher-education-vol-6/37-63

Jovanovic, J., Gasevic, D., Dawson, S., Pardo, A., & Mirriahi, N. (2017). Learning analytics to unveil learning strategies in a flipped classroom. *The Internet and Higher Education*, *33*, 74–85. https://doi.org/10.1016/j.iheduc.2017.02.001

Jovanović, J., Gašević, D., Pardo, A., Dawson, S., & Whitelock-Wainwright, A. (2019). Introducing meaning to clicks: towards traced-measures of self-efficacy and cognitive load, In *Learning analytics and knowledge*. https://doi.org/10.1145/3303772.3303782

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist*, *41*(2), 75–86. https://doi.org/10.1207/s15326985ep4102{_}1

Kitchenham, B., & Charters, S. (2007). Guidelines for performing Systematic Literature reviews in Software Engineering Version 2.3. *Engineering*, *45*(4ve), 1051. https://doi.org/10.1145/1134285.1134500

Kizilcec, R. F., Pérez-Sanagustín, M., & Maldonado, J. (2016). Recommending self-regulated learning strategies does not work (in MOOC context), In *The third (2016) acm conference on learning @scale*.

Kizilcec, R. F., Pérez-Sanagustín, M., & Maldonado, J. J. (2017). Self-regulated learning strategies predict learner behavior and goal attainment in Massive Open Online Courses. *Computers & Education*, *104*, 18–33. https://doi.org/http://dx.doi.org/10.1016/j.compedu.2016.10.001

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

Kovanović, V., Joksimović, S., Poquet, O., Hennis, T., de Vries, P., Hatala, M., Dawson, S., Siemens, G., & Gašević, D. (2019). Examining communities of inquiry in Massive Open Online Courses: The role of study strategies. *Internet and Higher Education*, *40*, 20–43. https://doi.org/10.1016/j.iheduc.2018.09.001

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

Maldonado, J., Pérez-Sanagustín, M., Manuel Moreno-Marcos, P., Alario-Hoyos, C., Merino, P., & Delgado-Kloos, C. (2018). Predicting Learners' Success in a Self-paced MOOC Through Sequence Patterns of Self-regulated Learning: 13th European Conference on Technology Enhanced Learning, EC-TEL 2018, Leeds, UK, September 3-5, 2018, Proceedings. https://doi.org/10.1007/978-3-319-98572-5{_}27

Maldonado-Mahauad, J., Pérez-Sanagustín, M., Kizilcec, R. F., Morales, N., & Munoz-Gama, J. (2018). Mining theory-based patterns from Big data: Identifying self-regulated learning strategies in Massive Open Online Courses. *Computers in Human Behavior*, *80*, 179–196. https://doi.org/10.1016/j.chb.2017.11.011

Malmberg, J., Järvelä, S., & Kirschner, P. A. (2014). Elementary school students' strategic learning: does task-type matter? *Metacognition and Learning*, *9*(2), 113–136. https://doi.org/10.1007/s11409-013-9108-5

Marcela, V. (2015). Learning Strategy, Personality Traits and Academic Achievement of University Students. *Procedia - Social and Behavioral Sciences*, *174*, 3473–3478. https://doi.org/10.1016/j.sbspro.2015.01.1021

Marton, F., & Säljö, R. (1976). on Qualitative Differences in Learning: I-Outcome and Process*. *British Journal of Educational Psychology*, *46*(1), 4–11. https://doi.org/10.1111/j.2044-8279.1976.tb02980.x

Matcha, W., Ahmad Uzir, N., Gasevic, D., & Pardo, A. (2020). A Systematic Review of Empirical Studies on Learning Analytics Dashboards: A Self-Regulated Learning Perspective. *IEEE Transactions on Learning Technologies*, *13*(2), 226–245. https://doi.org/10.1109/TLT.2019.2916802

Matcha, W., Gašević, D., Ahmad Uzir, N., Jovanovic, J., Pardo, A., Lim, L., Maldonado-Mahauad, J., Gentili, S., Perez-Sanagustín, M., & Tsai, Y.-S. (2020). Analytics of Learning Strategies: Role of Course Design and Delivery Modality. *Journal of Learning Analytics*.

Matcha, W., Gašević, D., Ahmad Uzir, N., Jovanović, J., & Pardo, A. (2019). Analytics of Learning Strategies: Associations with Academic Performance and Feedback, In *Proceedings of the 9th international conference on learning analytics & knowledge*. https://doi.org/10.1145/3303772.3303787

Matcha, W., Gašević, D., Ahmad Uzir, N., Jovanović, J., Pardo, A., Maldonado-Mahauad, J., & Pérez-Sanagustín, M. (2019). Detection of Learning Strategies: A Comparison of Process, Sequence and Network Analytic Approaches, In *European conference on technology enhanced learning*, Springer. https://link.springer.com/chapter/10.1007/978-3-030-29736-7_39

Matcha, W., Gašević, D., Jovanović, J., Ahmad Uzir, N., Oliver, C. W., Murray, A., & Gasevic, D. (2020). Analytics of Learning Strategies: the Association with the Personality Traits, In *Proceedings of the 10th international conference on learning analytics and knowledge (lak '20)*, Frankfurt, Germany, ACM. https://doi.org/10.1145/3375462.3375534

Mattick, K., Dennis, I., & Bligh, J. (2004). Approaches to learning and studying in medical students: Validation of a revised inventory and its relation to student characteristics and performance. *Medical Education*, *38*(5), 535–543. https://doi.org/10.1111/j.1365-2929.2004.01836.x

Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., Shekelle, P., Stewart, L. A., Estarli, M., Barrera, E. S., Martínez-Rodríguez, R., Baladia, E., Agüero, S. D., Camacho, S., Buhring, K., Herrero-López, A., Gil-González, D. M., Altman, D. G., Booth, A., Chan, A. W., Chang, S., Clifford, T., Dickersin, K., Egger, M., Gøtzsche, P. C., Grimshaw, J. M., Groves, T., Helfand, M., Higgins, J., Lasserson, T., Lau, J., Lohr, K., McGowan, J., Mulrow, C., Norton, M., Page, M., Sampson, M., Schünemann, H., Simera, I., Summerskill, W., Tetzlaff, J., Trikalinos, T. A., Tovey, D., Turner, L., & Whitlock, E. (2016). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Revista Espanola de Nutricion Humana y Dietetica*, *20*(2), 148–160. https://doi.org/10.1186/2046-4053-4-1

Moos, D. C. (2014). Setting the stage for the metacognition during hypermedia learning: What motivation constructs matter? *Computers and Education*, *70*, 128–137. https://doi.org/10.1016/j.compedu.2013.08.014

Morehead, K., Dunlosky, J., Rawson, K. A., Blasiman, R., & Hollis, R. B. (2019). Note-taking habits of 21st Century college students: implications for student learning, memory, and achievement. *Memory*, *27*(6), 807–819. https://doi.org/10.1080/09658211.2019.1569694

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

Panadero, E. (2017). A Review of Self-regulated Learning : Six Models and Four Directions for Research. *Frontiers in Psychology*, *8*, 1–28. https://doi.org/10.3389/fpsyg.2017.00422

Pardo, A. (2018). A feedback model for data-rich learning experiences. *Assessment & Evaluation in Higher Education*, *43*(3), 428–438. https://doi.org/10.1080/02602938.2017.1356905

Pardo, A., Gasevic, D., Jovanovic, J. M., Dawson, S., & Mirriahi, N. (2018). Exploring Student Interactions with Preparation Activities in a Flipped Classroom Experience. *IEEE Transactions on Learning Technologies*. https://doi.org/10.1109/TLT.2018.2858790

Pardo, A., Jovanovic, J., Dawson, S., Gašević, D., & Mirriahi, N. (2019). Using learning analytics to scale the provision of personalised feedback. *British Journal of Educational Technology*, *50*(1), 128–138. https://doi.org/10.1111/bjet.12592

Pardo, A., Poquet, O., Martinez-Maldonado, R., & Dawson, S. (2017). Provision of Data-Driven Student Feedback in LA & EDM. *Handbook of Learning Analytics*, 163–174. https://doi.org/10.18608/hla17.014

Pérez-Sanagustín, M., Hilliger, I., Alario-Hoyos, C., Kloos, C. D., & Rayyan, S. (2017). H-MOOC framework: reusing MOOCs for hybrid education. *Journal of Computing in Higher Education*, *29*(1), 47–64. https://doi.org/10.1007/s12528-017-9133-5

Proctor, B. E., Prevatt, F. F., Adams, K. S., Reaser, A., & Petscher, Y. (2006). Study Skills Profiles of Normal-Achieving and Academically-Struggling College Students. *Journal of College Student Development*, *47*(1), 37–51.

Rachal, C. K., Daigle, S., & Rachal, W. S. (2007). Learning problems reported by college students: Are they using learning strategies? *Journal of Instructional Psychology*, *34*(4), 191–199. http://content.ebscohost.com.proxy-remote.galib.uga.edu/ContentServer.asp?T=P&P=AN&K=28349624&S=R&D=slh&EbscoContent=dGJyMNXb4kSep7Y4yNfsOLCmr0qep7BSs6e4Sq6WxWXS&ContentCustomer=dGJyMPPd30m549+B7LHfhOoA

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

Reich, J., & Ruipérez-Valiente, J. A. (2019). The MOOC pivot. *Science*, *363*(6423), 130–131. https://doi.org/10.1126/science.aav7958

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

Roberts, B. W., & Mroczek, D. (2008). Personality trait change in adulthood. *Current Directions in Psychological Science*, *17*(1), 31–35. https://doi.org/10.1111/j.1467-8721.2008.00543.x

Rodríguez, M. F., Correa, J. H., Pérez-Sanagustín, M., Pertuze, J. A., & Alario-Hoyos, C. (2017). A MOOC-based flipped class: Lessons learned from the orchestration perspective, In *Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics)*, Springer Verlag. https://doi.org/10.1007/978-3-319-59044-8{_}12

Schunk, D. H. (2008). Metacognition, self-regulation, and self-regulated learning: Research recommendations. *Educational Psychology Review*, *20*(4), arXiv 0507464v2, 463–467. https://doi.org/10.1007/s10648-008-9086-3

Schwendimann, B. A., Rodríguez-Triana, M. J., Vozniuk, A., Prieto, L. P., Boroujeni, M. S., Holzer, A., Gillet, D., & Dillenbourg, P. (2016). Perceiving learning at a glance : A systematic literature review of learning dashboard research, *10*(100), 1–14. https://doi.org/10.1109/TLT.2016.2599522

Sedrakyan, G., Järvelä, S., & Kirschner, P. (2016). Conceptual framework for feedback automation and personalization for designing learning analytics dashboards, 1–3. https://lirias.kuleuven.be/handle/123456789/579647

Sedrakyan, G. [Gayane], De Weerdt, J., & Snoeck, M. (2016). Process-mining enabled feedback: "tell me what i did wrong" vs. "tell me how to do it right". *Computers in Human Behavior*, *57*, 352–376. https://doi.org/10.1016/j.chb.2015.12.040

Shaffer, D. W., Collier, W., & Ruis, A. R. (2016). A tutorial on Epistemic Network Analysis: Analyzing the Structure of Connections in Cognitive, Social, and Interaction Data. *Journal of Learning Analytics*, *3*(3), 9–45.

Shah, D. (2019). By The Numbers: MOOCs in 2019 - Class Central. https://www.classcentral.com/report/mooc-stats-2019/

Shokri, O., Kadivar, P., Farzad, V. E., & Sangari, A. A. (2007). Role of personality traits and learning approaches on academic achievement of university students. *Psychological Research*, *9*(3-4), 65–84. https://www.sid.ir/En/Journal/ViewPaper.aspx?ID=102080

Siemens, G., & Baker, R. S. J. d. (2012). Learning Analytics and Educational Data Mining: Towards Communication and Collaboration. *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge - LAK '12*, 252–254. https://doi.org/10.1145/2330601.2330661

Sonnenberg, C., & Bannert, M. (2015). Discovering the Effects of Metacognitive Prompts on the Sequential Structure of SRL-Processes Using Process Mining Techniques. *Journal of Learning Analytics*, *2*(1), 72–100.

Sutherland, S. E. (2004). An introduction to systematic reviews. *Journal of Evidence-Based Dental Practice*, *4*(1), 47–51. https://doi.org/10.1016/j.jebdp.2004.02.021

Tlili, A., Essalmi, F., Jemni, M., Kinshuk, & Chen, N. S. (2016). Role of personality in computer based learning. *Computers in Human Behavior*, *64*, 805–813. https://doi.org/10.1016/j.chb.2016.07.043

Trigwell, K., & Prosser, M. (1991). *Improving the quality of student learning: the influence of learning context and student approaches to learning on learning outcomes* (tech. rep. No. 9). Kluwer Academic Publishers.

Van der Kleij, F. M., Feskens, R. C. W., & Eggen, T. J. H. M. (2015). Effects of Feedback in a Computer-Based Learning Environment on Students' Learning Outcomes: A Meta-Analysis. *Review of Educational Research*, *85*(4), 475–511. https://doi.org/10.3102/0034654314564881

van der Aalst, W. M. P. (2011). *Process Mining*. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-19345-3

Winne, P. H. (2006). How Software Technologies Can Improve Research on Learning and Bolster School Reform. *Educational Psychologist*, *41*(1), 5–17. https://doi.org/10.1207/s15326985ep4101

Winne, P. H. (2013). Learning Strategies, Study Skills, and Self-Regulated Learning in Postsecondary Education. *Higher Education: Handbook of Theory and Research*, (28), 337–403. https://doi.org/10.1007/978-90-481-8598-6

Winne, P. H. (2014). Issues in researching self-regulated learning as patterns of events. *Metacognition and Learning*, *9*(2), 229–237. https://doi.org/10.1007/s11409-014-9113-3

Winne, P. H., Gupta, L., & Nesbit, J. C. (1994). Exploring Individual Differences in Studying Strategies Using Graph Theoretic Statistics. *Alberta Journal of Educational Research*, *40*(2), 177–93.

Winne, P. H., & Hadwin, A. F. (1998). Studying as Self-Regulated Learning. *Metacognition in educational theory and practice*, *93*, 277–304.

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

Winne, P. H., Jamieson-Noel, D., & Muis, K. (2002). *Methodological issues and advances in researching tactics, strategies, and self-regulated learning* (Vol. 12).

Wise, A., & Shaffer, D. W. (2015). Why Theory Matters More than Ever in the Age of Big Data. *Journal of Learning Analytics*, *2*(2), 5–13. https://doi.org/10.18608/jla.2015.22.2

Wise, A., Speer, J., Marbouti, F., & Hsiao, Y.-T. (2013). Broadening the notion of participation in online discussions: examining patterns in learners' online listening behaviors. *Instructional Science*, *41*(2), 323–343. https://doi.org/10.1007/s11251-012-9230-9

Woodfield, R., Jessop, D., & McMillan, L. (2006). Gender differences in undergraduate attendance rates. *Studies in Higher Education*, *31*(1), 1–22. https://doi.org/10.1080/03075070500340127

Yang, M., & Carless, D. (2013). The feedback triangle and the enhancement of dialogic feedback processes. *Teaching in Higher Education*, *18*(3), 285–297. https://doi.org/10.1080/13562517.2012.719154

Zeegers, P. (2001). Approaches to learning in science : A longitudinal study. *British Journal of Educational Psychology*, (71), 115–132.

Zhang, L.-F. (2003). Does the big five predict learning approaches? *Personality and Individual Differences*, (34), 1431–1446. www.elsevier.com/locate/paid

Zhou, M., & Winne, P. H. (2012). Modeling academic achievement by self-reported versus traced goal orientation. *Learning and Instruction*, *22*(6), 413–419. https://doi.org/10.1016/j.learninstruc.2012.03.004

Zimmerman, B. J. (2011). Motivational Sources and Outcomes of Self-Regulated Learning and Performance. *Handbook of self-regulation of learning and performance*, 49–64.