

# Applying Data Science and Learning Analytics Throughout a Learner's Lifespan

Goran Trajkovski  
*Western Governors University, USA*

Marylee Demeter  
*Western Governors University, USA*

Heather Hayes  
*Western Governors University, USA*

A volume in the Advances in Educational  
Technologies and Instructional Design (AETID)  
Book Series



Published in the United States of America by

IGI Global  
Information Science Reference (an imprint of IGI Global)  
701 E. Chocolate Avenue  
Hershey PA, USA 17033  
Tel: 717-533-8845  
Fax: 717-533-8661  
E-mail: [cust@igi-global.com](mailto:cust@igi-global.com)  
Web site: <http://www.igi-global.com>

Copyright © 2022 by IGI Global. All rights reserved. No part of this publication may be reproduced, stored or distributed in any form or by any means, electronic or mechanical, including photocopying, without written permission from the publisher. Product or company names used in this set are for identification purposes only. Inclusion of the names of the products or companies does not indicate a claim of ownership by IGI Global of the trademark or registered trademark.

Library of Congress Cataloging-in-Publication Data

Names: Trajkovski, Goran, 1972- editor. | Demeter, Marylee, 1972- editor. | Hayes, Heather, 1978- editor.

Title: Applying data science and learning analytics throughout a learner's lifespan / Goran Trajkovski, Marylee Demeter, Heather Hayes editors.

Description: Hershey, PA : Information Science Reference, 2022. | Includes bibliographical references and index. | Summary: "This publication examines novel and emerging applications of data science and sister disciplines in gaining insights from data to inform interventions into the learners' journey and interactions with an academic or training institution, covering topics such as building models of learners for success, using data to inform courseware and assessmentware development"-- Provided by publisher.

Identifiers: LCCN 2021060930 (print) | LCCN 2021060931 (ebook) | ISBN 9781799896449 (hardcover) | ISBN 9781799896456 (paperback) | ISBN 9781799896463 (ebook)

Subjects: LCSH: Education--Data processing. | Data mining.

Classification: LCC LB1028.43 .A64 2022 (print) | LCC LB1028.43 (ebook) | DDC 371.33/4--dc23/eng/20211223

LC record available at <https://lcn.loc.gov/2021060930>

LC ebook record available at <https://lcn.loc.gov/2021060931>

This book is published in the IGI Global book series Advances in Educational Technologies and Instructional Design (AE-TID) (ISSN: 2326-8905; eISSN: 2326-8913)

British Cataloguing in Publication Data

A Cataloguing in Publication record for this book is available from the British Library.

All work contributed to this book is new, previously-unpublished material. The views expressed in this book are those of the authors, but not necessarily of the publisher.

For electronic access to this publication, please contact: [eresources@igi-global.com](mailto:eresources@igi-global.com).

# Chapter 1

## Mind the Gap: From Typical LMS Traces to Learning to Learn Journeys

**Carmel Kent**

*The Open University, UK*

**Abayomi Akanji**

*EDUCATE Ventures Research, UK*

**Benedict du Boulay**

*EDUCATE Ventures Research, UK*

**Ibrahim Bashir**

*EDUCATE Ventures Research, UK*

**Thomas G. Fikes**

*Arizona State University, USA*

**Sue A. Rodríguez De Jesús**

 <https://orcid.org/0000-0002-1710-3167>

*Arizona State University, USA*

**Alysha Ramirez Hall**

*Arizona State University, USA*

**Paul Alvarado**

*Arizona State University, USA*

**Jennifer E. Jones**

*Arizona State University, USA*

**Mutlu Cukurova**

 <https://orcid.org/0000-0001-5843-4854>

*University College London, UK*

**Varshita Sher**

*EDUCATE Ventures Research, UK*

**Canan Blake**

*University College London, UK*

**Arthur Fisher**

*EDUCATE Ventures Research, UK*

**Juliet Greenwood**

*Arizona State University, USA*

**Rosemary Luckin**

*University College London, UK*

### ABSTRACT

*Many universities aim to improve students' 'learning to learn' (LTL) skills to prepare them for post-academic life. This requires evaluating LTL and integrating it into the university's curriculum and assessment regimes. Data is essential to provide evidence for the evaluation of LTL, meaning that available*

DOI: 10.4018/978-1-7998-9644-9.ch001

*data sources must be connected to the types of evidence required for evaluation. This chapter describes a case study using an LTL ontology to connect the theoretical aspects of LTL with a university's existing data sources and to inform the design and application of learning analytics. The results produced by the analytics indicate that LTL can be treated as a dimension in its own right. The LTL dimension has a moderate relationship to academic performance. There is also evidence to suggest that LTL develops at an uneven pace across academic terms and that it exhibits different patterns in online as compared to face-to-face delivery methods.*

## **INTRODUCTION**

This study was initiated to explore how, and to what degree, students at Arizona State University (ASU) acquire skills in Learning to Learn (LTL).

Helping students learn to learn is a worthwhile aim, but it needs to be actioned through explicit teaching and reflective assessment. If LTL is one of the aims of higher education, it also needs to be assessed at various stages to ensure its growth. Just as the students routinely get transcripts that reflect their performance grades, there needs to be some validation about their increasing effectiveness as learners (Molenaar et al., 2019). However, most learning management systems (LMSs) record students' actions largely as 'users' of the university rather than as learners. For example, LMSs often record students watching a video or submitting an assignment, rather than activities that offer more evidence about LTL, such as reacting to feedback (Suraworachet et al., 2021) and self-reflecting on their learning activity (Lau et al., 2017).

### **A Working Definition for LTL**

The concept of LTL is derived from the definition of learning. Scholars continually debate the definition of learning, but most of them would agree that learning is a process, that it involves change that follows experience (Schunk, 2012), and that, for the most part, it is internal and so invisible (Lefrançois, 2019). These characteristics - all have direct consequences on how learning can be evaluated.

In this chapter, LTL was conceptualized as a process of improvement in self-regulated learning (SRL). Self-regulated learners are “*meta-cognitively, motivationally, and behaviourally active participants in their own learning process*” (Zimmerman, 1989, p. 4).

LTL depends on each learner's ability and willingness to reflect on and thus improve their self-regulated learning capability (Education Council, 2006; Hautamäki et al., 2002). To measure this process, it is necessary to adopt a temporal viewpoint of the progress learners make in their ability to understand (and adapt) their own learning strategies, strengths, and motivation as learners (The Campaign for Learning, 2007). In the current study, the researchers operationalized the LTL journeys of university students as a temporal sequence of “snapshots”, each of which might contain some evidence about students' SRL (or SRL-related) capabilities. Examples of the types of snapshots, along with the process of suggesting which moments should be captured as snapshots are given in the ‘LTL Ontology’ section. The temporal order of these snapshots and their individual strengths of evidence about students' SRL capability were used to create an overall view of their LTL journeys.

## Learning to Learn in the Literature

Being aware of one's learning processes, one's inclinations and then improving them is a crucial educational objective for many learners in the 21st century. Learners live in a constantly changing environment and must adapt to its changing demands. These uncertainties call for the need to make it easier for learners to transition across disciplines and become independent learners. In fact, LTL is mentioned as one of the eight critical competencies in the recommendations for lifelong learning, which were adopted by the Education Council and the European Parliament in December 2006 and revised in 2018 (Education Council, 2006, 2018).

As LTL is defined in this study as an improvement in students' SRL capabilities, other related constructs such as expectations, motivation, goals, and self-regulatory learning strategies immediately come into play (Eccles & Wigfield, 2002; Robbins et al., 2004). These constructs related to LTL are both malleable and highly context-sensitive (e.g., Carver & Scheier, 1981; Richardson, 2012; Wolters et al., 2003). This raises the question as to how an academic institution can know how well its students are able to self-regulate their learning processes? The indicators are far from being straightforward and directly observable.

SRL is based on students' self-awareness, regulatory skills, and their ability to then transfer those *learning skills* to domains other than those in which they were originally developed (Epstein, 2019; Halpern, 1998). Note that the authors are not referring to the transfer of problem-solving skills or conceptual understanding from one domain to another, for example, from arithmetic to algebra, but to the *transfer of the learning skills themselves* from one context to another. For instance, learning to self-explain effectively in one context to choosing to self-explain in an unrelated context because it is an effective learning strategy is an example of such a transfer (see, for example, Chi et al., 1989). Awareness of one's own learning focuses on knowing what one knows, what one does not know, what are the skills one has gained, and what are one's attitudes and desires with regards to learning. Regulatory skills comprise individuals' ability to organize their actions to achieve their learning goals (Matrić, 2018). They include learning management skills such as planning what to do next, setting goals, monitoring progress, and then reflecting on it all (Zimmerman, 2002).

Research has shown that the observable outcomes related to LTL, such as improved academic performance and retention, are also closely related to constructs such as SRL, engagement, motivation, and self-efficacy (Richardson, 2012). LTL is also related to personality traits, which are widely studied in the learning literature (e.g., conscientiousness, openness, agreeableness, emotional stability, extraversion; Poropat, 2009). Nevertheless, it was hypothesized in this work that, despite these interrelations between these related pedagogic constructs, LTL represents a unique dimension of learning.

The research reported here is concerned with the challenge of how the gap between the desired and the existing data can be bridged when exploring LTL. Specifically, two research questions were explored:

**Research Question 1 (RQ1):** Is LTL, specifically the SRL sub-construct, a distinct dimension from other, well-researched dimensions, such as academic performance and engagement?

**Research Question 2 (RQ2):** How does LTL, specifically the SRL sub-construct, change over time, in particular across academic terms?

Table 1. Main datasets used

Data	Granularity	Description
Academic performance, enrolment data and socio-demographic data	One record per student, course, year, and term	See bullet points <i>o</i> , <i>p</i> and <i>q</i> in Figure 2 below. These are the contextual data of the students, containing information such as each student's admission, courses taken, term, academic performance and loans.
Assignment dataset from the LMS	Daily	See bullet point <i>s</i> in Figure 2. These are records of assignments that were either submitted or not. The same student could submit several assignments each week, which would result in several records. The data contain the student's anonymized ID, the contextual assignment information, dates associated with the submission, their grade etc.
Clickstream data from the LMS	Weekly	See bullet point <i>r</i> in Figure 2. These contain students' digital footprints of their engagement, such as sessions, clicks, and duration.

## METHODS

As described above, LTL is considered as an improvement in students' SRL, which is a mental attribute. Researchers have tried to study various observable approximations to it, such as heart-rate variability (Spann et al., 2017), question-asking and engagement in online discussions (Pardo et al., 2016), and the association of learning strategies with teachers' feedback (Matcha et al., 2019). Some researchers have used self-reporting mechanisms, such as interviews or questionnaires (Pintrich et al., 1991). While there are questionnaires which measure LTL-related constructs, they do not capture the external signifiers of LTL ecologically as they take place and are harder to scale to a large number of students. To overcome these shortcomings, some researchers have used clickstream data from LMSs (Cicchinelli et al., 2018; Motz et al., 2019). LMS clickstreams are themselves limited since they are rarely developed to reveal the contextual complexity and multi-dimensionality of constructs such as SRL (Boulton et al., 2018). Other researchers have used other approaches to assess LTL-related constructs. For example, Hender-son (2018) employed a phenomenological approach to study the engagement of history students with a simulation. He collected memos and diagrams throughout the students' learning process, following Corbin and Strauss's (2014) memos and diagrams protocol, and complemented it with semi-structured interviews with some of the students. Another example is the use of the Baker Rodrigo Cumpaugh Monitoring Protocol (BROMP), a well-validated protocol for quick and time-synchronized quantitative field observations of students' affect and behaviour (Baker et al., 2018). BROMP has been used to evaluate different teacher practices (Hymavathy et al., 2014), and students' regulatory skills and engagement (Downer et al., 2010; Hymavathy et al., 2014).

## Data Sets and Cohort

This section describes the data sources the researchers used in the analysis, as well as the details of the observed cohort of students. This study was conducted using data from two courses at ASU. The raw data included students' demographic information, their enrolment and learning context information such as courses, scholarships and grades, details about the submission of assignments and the digital footprints of their work in the LMS. The observed cohort consisted of students who took one or other of two courses: BIO181 (the most basic course in the Biology programme) and PSY101 (the most basic course in the Psychology programme). The students were enrolled in two different study delivery modalities (online

**Mind the Gap**

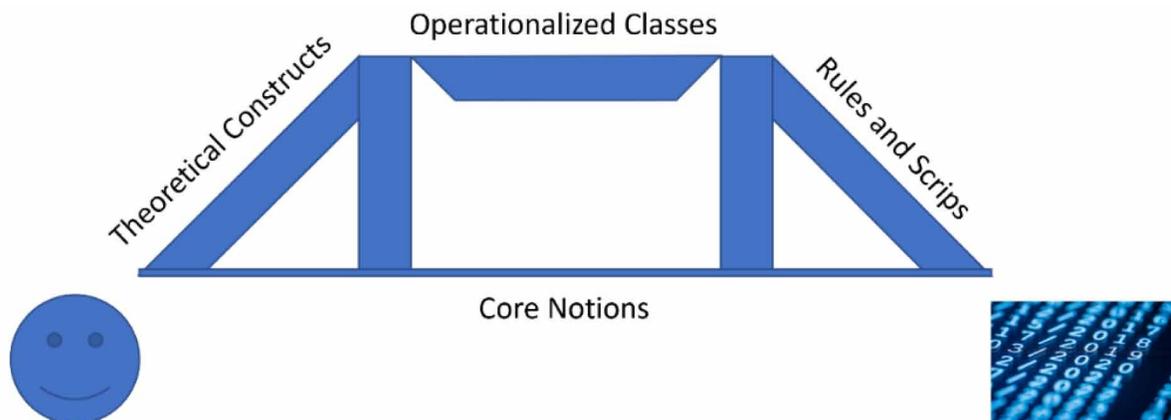
*Table 2. Cohort data*

Courses	Description	Number of students
BIO181	Biology's typical first course	2,056
PSY101	Psychology's typical first course	1,895
Terms	Description	Number of students
Spring of 2019		520
Summer of 2019		298
Fall of 2019		979
Spring of 2020		924
Summer of 2020		454
Fall of 2020		950
Sessions	Description	Number of students
A	Seven and a half weeks of online learning.	1,872
B	Seven and a half weeks of online learning.	1,653
C	Fifteen weeks of F2F learning (runs in parallel with Sessions A and B).	467

or face-to-face – F2F) during 2019 and 2020. Each of these two academic years consisted of three terms (Fall, Spring, Summer). In each term, students could study in one of the three possible sessions: Session A, which typically spanned the first seven and a half weeks of the term; Session B, typically the last seven and a half weeks of the term; or Session C, which was fifteen weeks long and ran in parallel with A and B. Students learning in the online provision (called ASU Online – ASUO) were enrolled in Sessions A and/or B, while students enrolled on the F2F provision took the longer Session C. In some instances, especially in the F2F sessions, students were still active in the LMS even after the end of their session.

Tables 1 and 2 below show detailed information about the data sources used and the cohort of students observed, respectively. The numbers in Table 2 do not always match up since there were intersections

*Figure 1. An ontology as a bridge*



among the elements in the categories. For example, a specific student might be participating in both courses in the same term and session. In that case, the student's instance would be counted twice in the courses (once in each course) but would be counted just once for the term and session.

## The LTL Ontology

Most LMSs are designed to provide 'institutional analytics' and are missing a layer of semantics needed to provide any sense of the students' behavioural and emotional points of view. To build this semantic bridge, the authors developed an LTL ontology, which expresses the complexities of the LTL concepts, along with their relationships and properties. As mentioned earlier, there is a gap between the data sources which are typically found in higher educational institutions, and data that would be helpful for analyzing LTL. Here the authors address this gap and how they designed the required bridge.

## Connecting Theoretical and Operationalized Constructs

Most LMSs use concepts such as courses, assignments and grades, rather than theoretical concepts such as engagement and SRL. The LTL ontology was designed to distinguish between theoretical constructs and observable bits of behaviour (see Figure 1 below).

At the left-hand end of this bridge are the **theoretical constructs** of LTL, such as SRL, attention, reflection, motivation and so on. Along the ground are the **core notions**, such as a course, a session and programme. The main span of the bridge is made up of what the researchers have called **operationalized classes**. These describe observable "bits of behaviour" that students engage in, such as taking part in a group assignment or undertaking some peer review tasks. At the right-hand end of the bridge is the data.

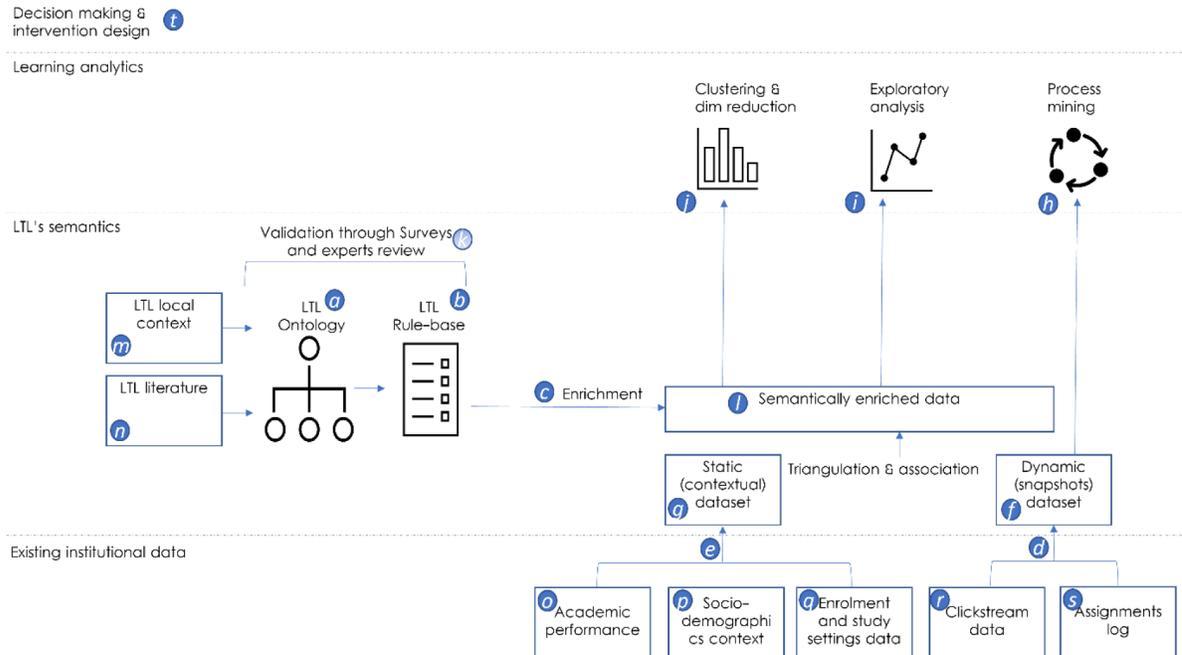
The ontology expresses the associations between theoretical constructs and their possible operationalizations and represents the fact that a theoretical construct might be operationalized in more than one way. For example, to operationalize student engagement, one might want to collect data about their interactions in class discussions, as well as their engagement with the course materials. The other way round is also possible. Namely, a single observable behaviour might contribute to the operationalization of more than one theoretical construct. For example, a student attending their instructor's office hours might operationalize a behaviour of help-seeking but might also indicate their engagement and motivation in the course.

## Connecting Temporal and Static Information

To capture the notion of LTL as a process, the authors built the ontology around the concepts of a *snapshot* and a *sequence of snapshots*. A snapshot is a time-stamped event capturing data about a student's behaviour. The snapshot includes properties such as the snapshot's type (e.g., submitting an assignment) and its granularity (e.g., occurring at a specific instant in time or across a whole week). A sequence of snapshots captures the order and frequency of individual snapshots and thus encapsulates the essence of an LTL journey.

## Mind the Gap

Figure 2. Analytic architecture to bridge between the desired and existing data. The circled bullet point annotations are used in the body of the text to further explain each component. Bullet point *k* is planned for future research



## Connecting Existing and Desired Data Points

To bridge between the semantics of the existing data and LTL, the researchers derived a set of features, represented via a coding system that classified those snapshots providing some evidence of LTL together with a score indicating the degree of relevance of its evidence to LTL. The LTL coding system was validated, both by experienced tertiary educators (who are among the authors of this chapter) and by identifying significant correlations with well-established metrics, such as the Grade Point Average (GPA). To do that, an LTL rule-base was designed (seen at the righthand end of the bridge in Figure 1, and some of which is shown in Appendix 1). This set of rules was designed to enrich the existing evidence as collected from the raw dataset (see bullet points *o*, *p*, *q*, *r* and *s* in Figure 2).

Two of the engineered features that were created using the LTL rule-base were an LTL Classification scheme and a measure of the strength of the LTL Evidence. Each LTL Classification feature assigned an LTL-related annotation to the activities described in the snapshots. The classification types were Analysis, Assessment, Engagement, Feedback, Motivation, Practice, Reflection and Review. The LTL evidence strength used a numerical scale from  $-3$  to  $3$ , which indicated the strength and valence of the evidence related to LTL that was associated with different snapshots. Negative values (from  $-3$  to  $-1$ ) were assigned if the evidence was considered to be associated negatively with LTL,  $0$  if the evidence was considered to be independent of LTL, or positive (from  $1$  to  $3$ ) if the evidence was considered to be associated positively with LTL.

The above engineered features provided an insight into how different snapshots (such as interactions with the LMS) and outcomes (such as assignment grades) contributed to each student's LTL evidence base. The researchers could now quantify how the LTL evidence strength for different LTL types changed across each term/session. Since these features were computed for each snapshot, the researchers could compare the LTL evidence of an individual student's particular snapshot with the average value of all snapshots for that student. Similarly, they could also juxtapose each student's LTL evidence against the average of all students in the same course during a specific term or session.

## The Analytic Architecture

The semantic layer described above (see bullet point *c* in Figure 2 below) was created by the LTL rule-base (see bullet point *b* in Figure 2 below). It was built based on experts' pedagogical knowledge and based on a literature review about LTL, both of which were described by the ontology (see bullet point *a* in figure 2).

The way the LTL ontology was designed informed both a temporal view of the data (to capture behavioural trends, see bullet point *f* in Figure 2) and a static view of the data (to capture the rich contextual characteristics of the students' LTL journeys, see bullet point *g* in Figure 2). To analyse the temporal nature of LTL, while not neglecting the rich context ASU maintained for their students, the researchers opted for a hybrid analytical approach, as described below, involving process mining (bullet point *h* in Figure 2), exploratory analysis (bullet point *i*) and dimension reduction & clustering (bullet point *j*). They also tried validating the whole approach using survey data (bullet point *k*). However, they could not complete this due to a low response rate (see more in the Limitations and future work section).

## Process mining

Process mining (Van Der Aalst, 2012) (bullet point *h* in Figure 2) is a commonly used technique for analyzing and monitoring the processes of students' behaviour (Saint, 2021). The researchers used it to model the LTL journeys, which were operationalized by the students' sequences of snapshots. Those sequences exhibit the order and frequency of various LTL related snapshots throughout an academic term. The researchers employed Heuristic Miner (HM), a commonly used process mining algorithm to model the snapshots data (Weijters et al., 2006), which creates heuristic nets, implemented by using the pm4py Python library (Berti et al., 2019). HM's aim is to generalize the process models. It seeks generic rules or patterns that the sequences of snapshots most frequently follow, which makes it more robust with respect to noise and outliers compared to other algorithms.

The HM nets were modelled with nodes representing snapshots, annotated by the combination of their corresponding LTL classification and the strength of LTL evidence. To make the models more interpretable the researchers decided to reduce the number of possible nodes. For that, they merged some of the combinations of LTL classification and evidence strength and reduced the number of possible nodes to 13. The merging was based on the semantics of those classifications, separately merging all the positive and negative evidence levels together. Then, the researchers chose a threshold of 0.05% and filtered out all nodes appearing with a lower frequency than that (See the filtered list in Appendix 2).

The directed edges between the nodes corresponded to temporal relationships between the snapshot nodes. For instance, an edge connecting snapshots *s1* and *s2* corresponds to the observation that snapshot *s1* was followed by snapshot *s2*. The edge connecting the two snapshots could be embellished with extra

## Mind the Gap

Table 3. Principal Component Analysis, rotation converged in 5 iterations. Only loadings above 0.3 are shown. The bolded numbers suggest a simple solution

Behavioural feature	Component 1 - Engagement	Component 2 – Academic performance	Component 3 - LTL
The average number of cursor clicks (transformed by square root)	<b>0.825</b>		
The average time (number of minutes) spent on the LMS's section of each course (transformed by square root).	<b>0.793</b>		0.316
An average LTL evidence level of the Practice category (as seen in assignments like lab works, simulations, case studies, problem sets and quizzes).	<b>0.735</b>		
The average number of sessions a student spent on the LMS (transformed by square root).	<b>0.707</b>		0.356
The average submission earliness (how early a student submits a course' assessment, before the due date) (transformed by log10).	<b>0.661</b>		
The average percentage of grade points a student earned for their course' assessment.	0.661	<b>0.409</b>	0.371
The number of grades a student has gained that are higher than their personal average grade.		<b>0.897</b>	
The number of grades a student has gained that are lower than their personal average grade.			
The number of grades a student has gained that are higher than the average of all students' grades (transformed by square root).	0.613	<b>0.657</b>	
The number of grades a student has gained that are lower than the average of all students' grades (square root).		<b>0.721</b>	
Average LTL evidence level of assessment (as seen in events such as final examinations and assignment).	0.598		<b>0.463</b>
Average LTL evidence level of engagement (seen through the number of clicks, sessions, and minutes spent online).			<b>0.782</b>
The percentage of assignments not submitted out of all assignments that should have been submitted (transformed by log10).			<b>-0.684</b>

information, such as how many instances were found to follow this edge, or the average time taken per instance to go from  $s1$  to  $s2$ .

## Dimension reduction and clustering

Dimension reduction and clustering were carried out to complement the process mining approach with the context of the data. The researchers carried out a factor analysis to identify whether LTL features formed a unified dimension and whether this was a distinct dimension from other, well-researched dimensions, such as academic performance and engagement.

Using the dimensions identified, they clustered (Trivedi et al., 2015) the semantically enriched aggregated dataset to identify various learner profiles (del Valle & Duffy, 2007). In future work, the authors suggest that these clusters should be validated on a larger cohort of students and that learning interventions are designed to be tailored for each of them.

In the next section, the results of the hybrid approach to learning analytics are detailed to answer both the research questions.

## **RESULTS**

In this section, the authors present the models that resulted from their ontology-informed analysis.

### **RQ1: Is LTL Indeed a Distinct Dimension?**

A principal components analysis (PCA) was run on the 13 variables aggregated from the dynamic dataset (after removing five variables as they were not adequate for a PCA) and 4,208 cases (after removing eight outliers). Each case represented a unique combination of a student, a term, a session within the term, and a course. The 13 variables, along with the transformations that were applied to them to make them adequate for a PCA, are shown in Table 3 below.

To inspect the suitability of the PCA, the correlation matrix showed that all variables had at least one correlation coefficient greater than 0.3. Variables with a correlation above 0.85 were considered as too highly correlated and some were therefore removed. The overall Kaiser-Meyer-Olkin (KMO) (Kaiser, 1974) measure was 0.809, with individual KMO measures all greater than 0.7 (except for the number of low course-standardized grades, transformed by square root, which the researchers decided to leave in since it was semantically important). Bartlett's test of sphericity was statistically significant ( $p < .001$ ), indicating that the data was likely to be factorizable.

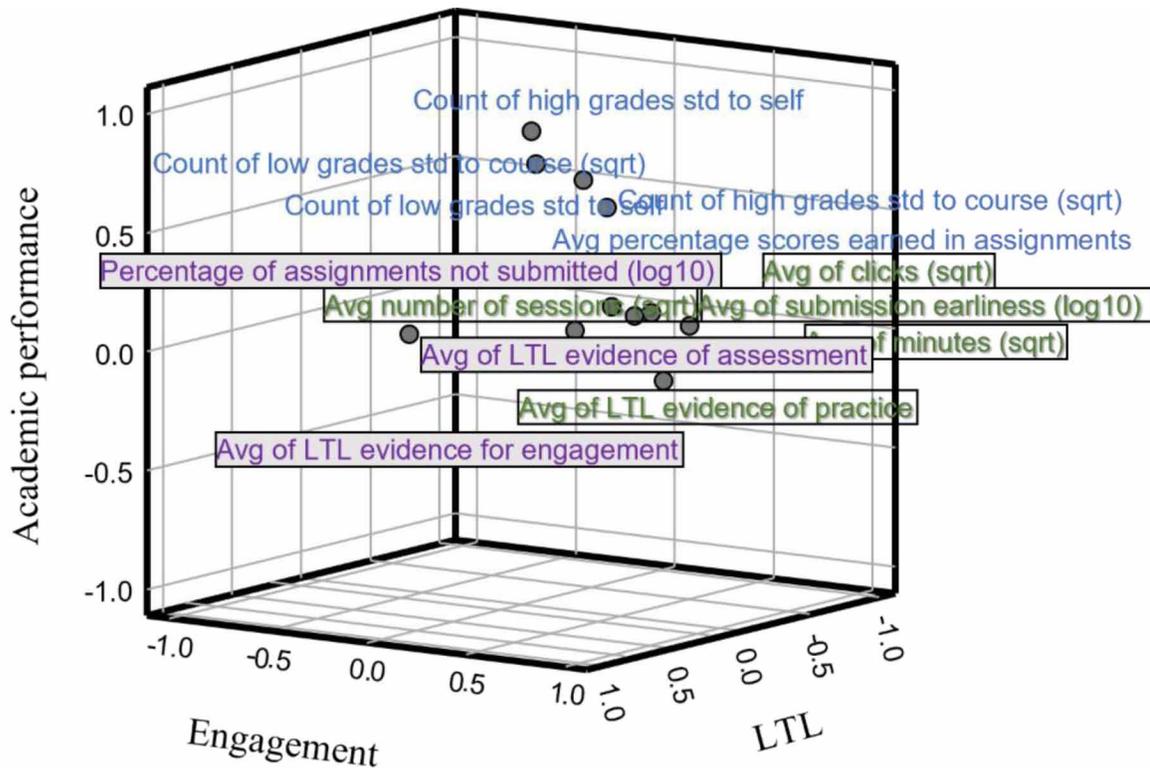
The PCA revealed three components that had eigenvalues greater than one and explained 69.15% of the total variance (out of which component 1 explained 46.69%, component 2 explained 13.66% and component 3 explained 8.79% of the total variance). Visual inspection of the scree plot indicated that three components should be retained (Cattell, 1966). In addition, a three-component solution met the interpretability criterion. A Varimax orthogonal rotation was employed to aid interpretability, which exhibited a 'simple structure' (Thurstone, 1947). The interpretation of the data was consistent with the theoretical assumptions about LTL with strong loadings of engagement items on component 1, grades on component 2, and LTL evidence items on component 3. Component loadings and communalities of the rotated solution are presented in Table 3 above.

It is interesting to note that the engagement component included the LTL evidence for practice, which might suggest that practice and engagement were strongly related. It is also interesting to note that the percentage of points earned (a standardized calculated feature) had a loading larger than 0.3 on all three components. For interpretability reasons, the researchers decided to base their analysis on associating it with the academic performance component, although statistically, it bears on all three components. It is shown already that academic performance and engagement are moderately related and yet still signify different dimensions (Kent et al., 2016), and as it will be shown later in this section, academic performance was also moderately related, yet signified a different dimension than LTL. The number of assignments that were not submitted was negatively loading onto the LTL component, which also suggests that LTL goes hand in hand with staying on track with the academic workload.

Once the researchers had gained some understanding of how students' behaviour could be explored, they turned to clustering the students. For that, they added to the three components already revealed (see Figure 3) the five LTL evidence variables, which were excluded from the PCA beforehand to create a

## Mind the Gap

Figure 3. A component plot of the PCA solution. It appears that grades (text without an outline), engagement (text with an outline but no fill) and LTL (text with outline and grey fill) were three distinct dimensions, which strengthens the hypothesis that LTL was a distinct dimension and should be evaluated separately



distribution that was appropriate for a PCA. These five variables were namely – the LTL classifications of reflection, motivation, review, unclassified and analysis.

The researchers then carried a K-means cluster analysis (Lloyd, 1982) to explore various clustering solutions, used after various sorting options. They considered standardizing the eight variables but decided against it because they used completely different scales, and the solution on the standardized version was weaker. Finally, the chosen solution converged after 11 iterations.

Three clusters of students were identified: cluster 1 with 267 cases, cluster 2 with 2,818 cases, and cluster 3 with 1123 cases (see Figure 4 below). The proportions were not ideal, with cluster 1 being very small. However, further analysis showed that cluster 1 was more distant from clusters 2 & 3 (3.35 and 3.61 respectively in the 8-dimensional space of the 8 variables) than they were from each other (1.82). This was deemed potentially interesting (taking into account the other variables) and was therefore further explored.

It seems from the analysis below that the three clusters could be characterized roughly as:

- Cluster 1 (in dots in Figure 4), which the researchers termed as ‘**The Versatile (Adaptable) Achievers**’: is the smallest cluster, which contains students exhibiting relatively high grades and high LTL levels;

Figure 4. Percentage of cases in each cluster. Eight cases were not clustered due to missing values

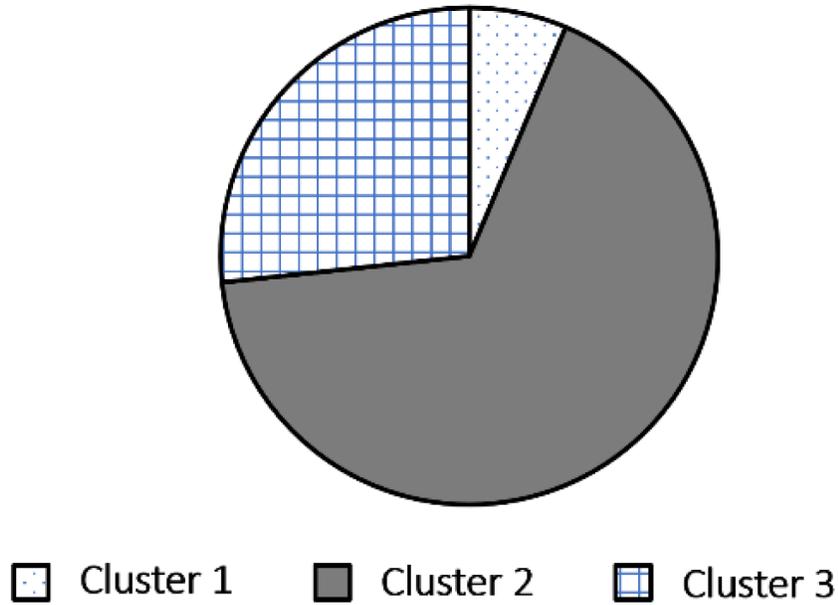
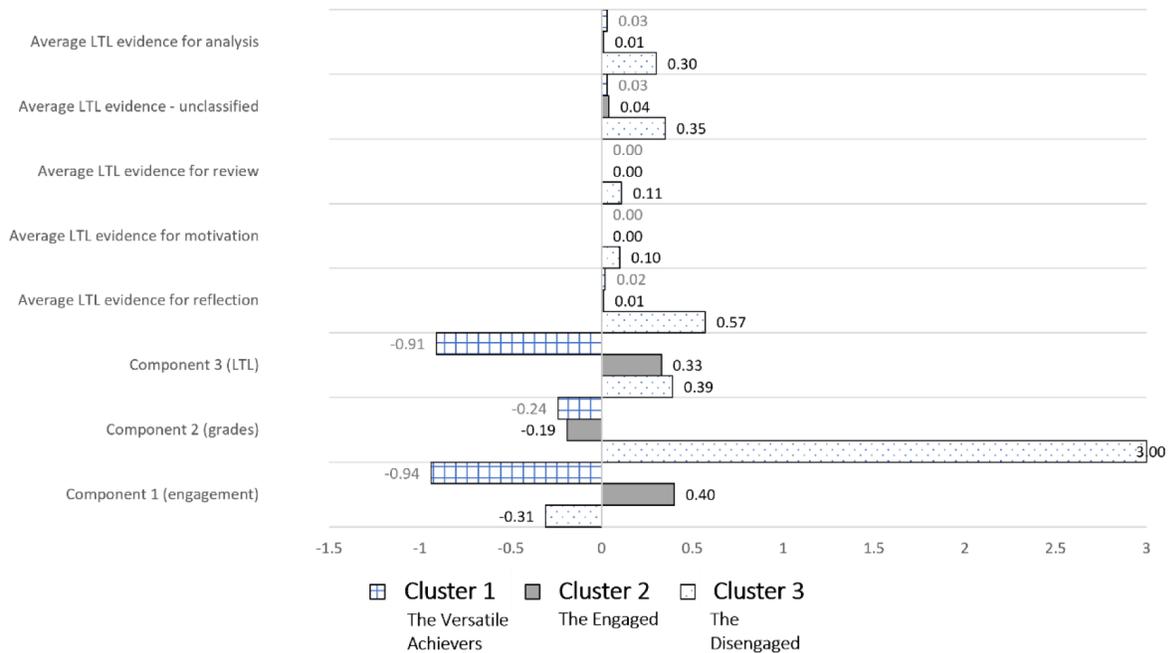
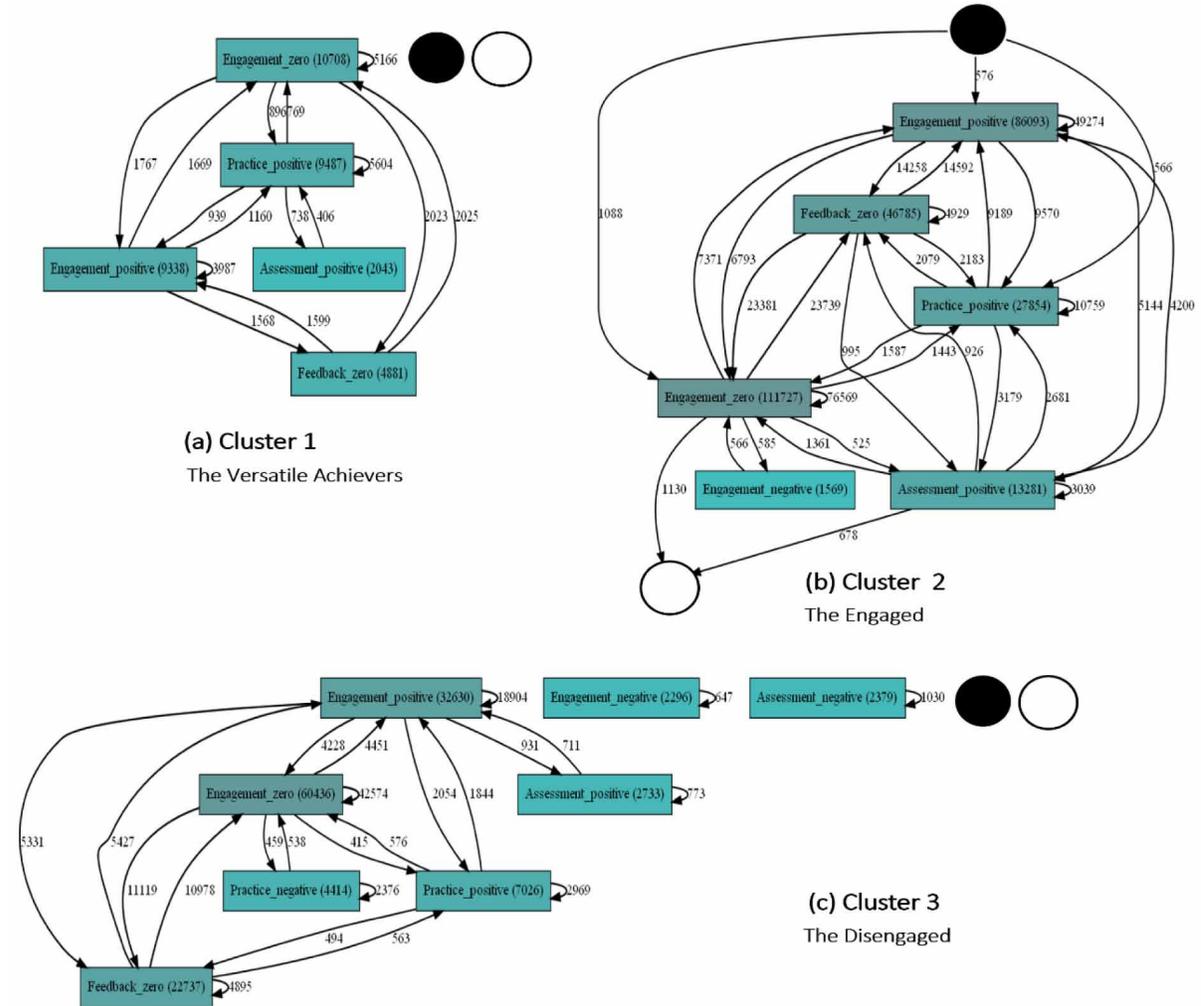


Figure 5. The three clusters using a final centres profiling (i.e., the mean of each variable within the various clusters): cluster 1 in dots, cluster 2 with solid fill and cluster 3 with a grid fill



**Mind the Gap**

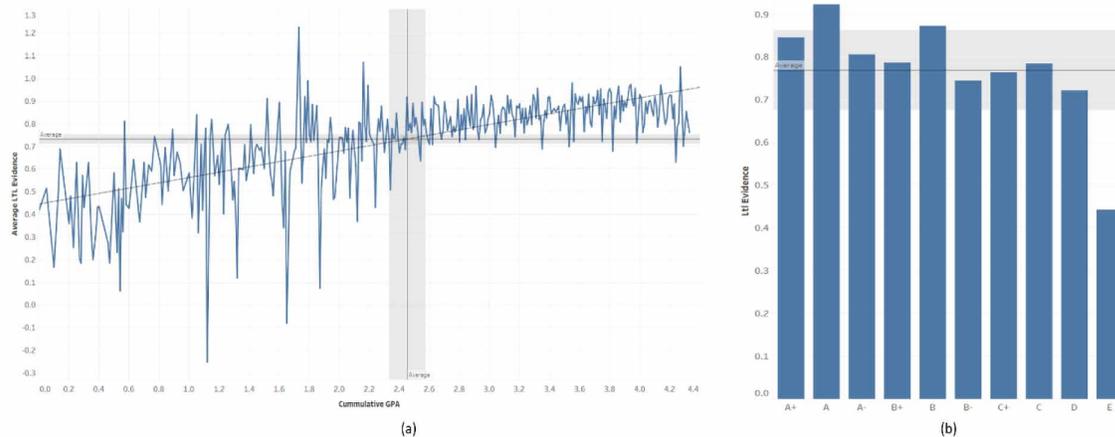
Figure 6. The clusters' Heuristic nets: (a: left top): cluster 1 (the Versatile Achievers); (b: right top): cluster 2 (the Engaged); and (c: bottom): cluster 3 (the Disengaged)



- Cluster 2 (with solid fill in Figure 4), termed ‘**The Engaged**’: was by far the largest cluster, with the most engaged students in it;
- Cluster 3 (with grid fill in Figure 4), termed as ‘**The Disengaged**’: is the cluster that might require the most attention in designing interventions, since the students in it are characterized with relatively low levels of engagement, low grades, and low LTL levels.

As shown in Figure 5 below and in the post hoc analysis shown in Appendix 3: (i) the differences between cluster 1 (the *Versatile Achievers*) and the other two clusters in terms of LTL evidence and grades were statistically significant, (ii) both clusters 1 (the *Versatile Achievers*) & 2 (the *Engaged*) had a statistically significantly higher level of the LTL component as compared to cluster 3 (the *Disengaged*), and (iii) cluster 2 (the *Engaged*) had a statistically significantly higher level of the engagement compo-

Figure 7. (a: left) Relationship of the strength of average LTL evidence (Y axis) against the cumulative GPA (X axis); (b: right) The relationship between the strength of evidence of LTL (Y axis) and the grade awarded (X axis)



ment over cluster 1 (the *Versatile Achievers*), which was statistically significantly higher than cluster 3 (the *Disengaged*).

The clusters were chosen to maximize the differences among cases in different clusters. All variables had a significant impact on determining the clusters, as shown in the ANOVA table in Appendix 3.

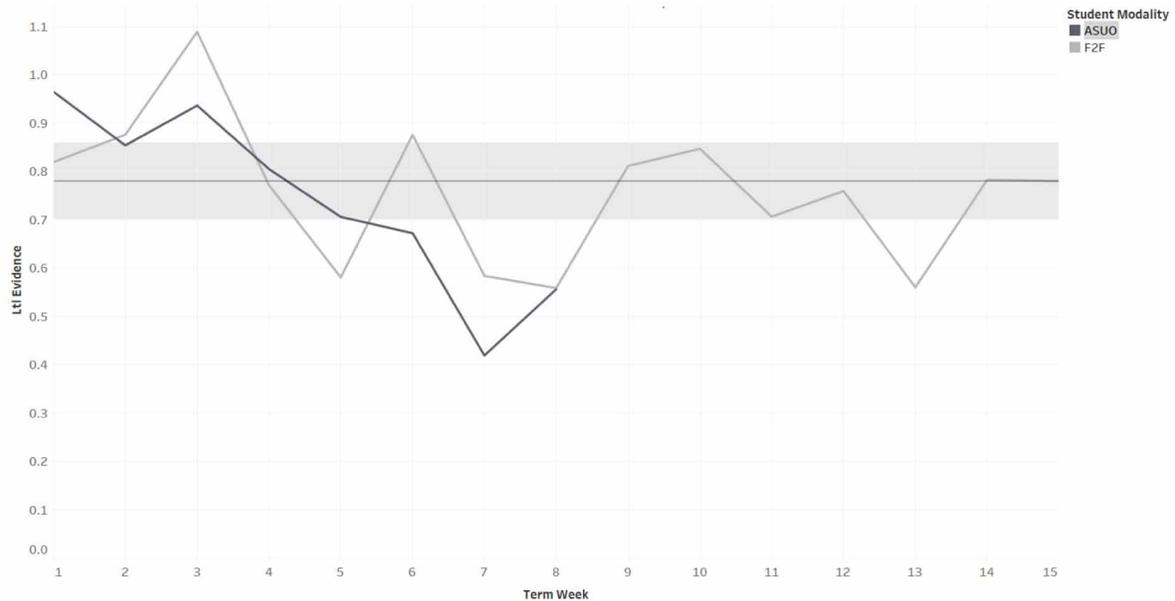
To further strengthen the observations from the ANOVA analyses, the researchers decided to investigate the process models from the three students' clusters using HM (see Figure 6 below). The black and the white discs indicate the start and the end state nodes of those process models respectively. In two of the HMs they seem disconnected as the edges connecting them to the rest of the model were not manifested frequently enough. As can be seen in Figure 6, the HM of the *Disengaged* cluster showed the highest number of instances of negative engagement snapshots and subsequent negative assessment snapshots when compared to the remaining two clusters. These observations support the hypothesis that poor academic performance is associated with this cluster. Surprisingly, even though the *Versatile Achievers* cluster was characterized as the group with the highest grades, their engagement levels were lower when compared to the *Engaged* cluster and they registered fewer instances of positive practice snapshots following positive engagement snapshots. In other words, even though the *Engaged* cluster students' engagement level with the course curriculum was higher and they were more likely to transform their positive engagement levels into positive practice, their grades were not reflective of their efforts. Recently, Suraworachet et al. (2021) have shown that high SRL students react to their feedback faster than low SRL students. Similarly, these findings are supported to some extent by the three process models, shown below, where the ratio of positive to zero LTL snapshots which directly follow being given feedback was 79%, 61% and 49% for clusters 1, 2 and 3 respectively.

To further understand the relationship between academic performance and LTL, the researchers examined how the average value of students' LTL (across all LTL types) changed in relation to their cumulative GPA (Figure 7(a)) and their cumulative grades (Figure 7(b)).

As can be seen, higher academic performance corresponds to higher average values of LTL evidence. Figure 7(a) shows that when the cumulative GPAs are below average, there is a much higher variance

## Mind the Gap

Figure 8. The average LTL evidence (Y axis) throughout the term weeks (X axis), for online students (ASUO in black, attending shorter sessions A and B) and F2F students (F2F in grey, attending 15 weeks sessions C)



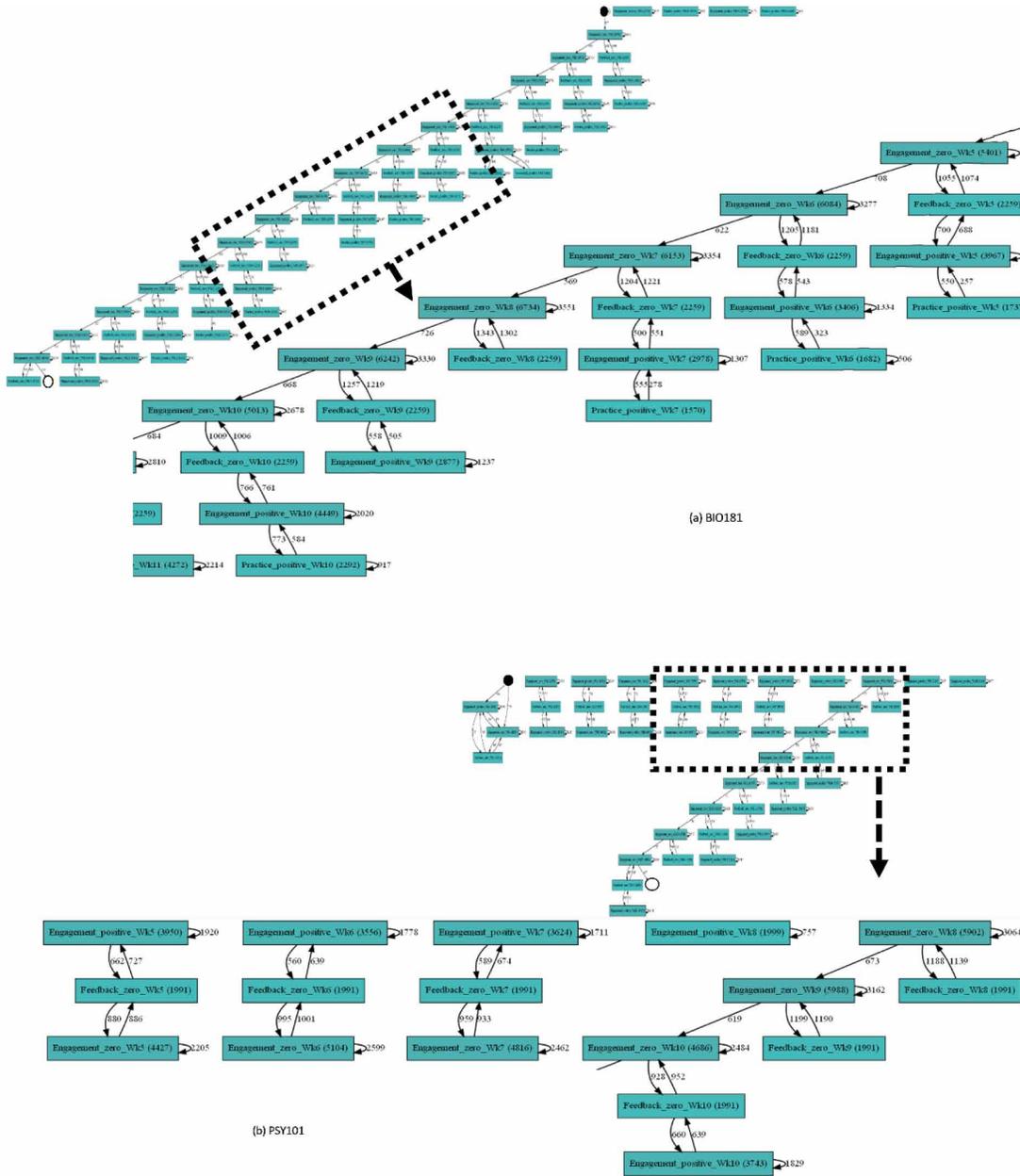
and inconsistency in the average LTL evidence, with most values of LTL evidence levels also lying below the average. However, for above-average cumulative GPAs, a tighter spread of LTL evidence can be seen, with almost all the values consistently lying above the average. A Spearman's rho correlation coefficient was used to assess the relationship between a student's cumulative GPA and the strength of their LTL evidence (as the examined data were not normally distributed). There was a moderate positive relationship between the two, i.e., a student's cumulative GPA and their LTL evidence,  $r=0.46$ ,  $p<0.001$ . Since the LTL evidence was not normally distributed, Mann-Whitney U test was used to show that the average LTL evidence was significantly higher for students with A grades (including A+, A & A-) than those with C grades (including C & C+),  $U=251755.50$ ,  $p<0.001$ .

## RQ2: How Does LTL Change throughout the Term?

Next, the researchers explored how LTL levels changed throughout the term (see Figure 8 below).

Both teaching modalities (online and F2F) show a clear pattern of decrease in the LTL evidence across the term. Finer grained interpretations, which are more descriptive and should be taken more cautiously might suggest other similar trends. For example, both the online and F2F samples show higher than median LTL evidence before the middle of the term (week 4 for online, week 7 for F2F) followed by a steep drop at the middle of the term (which might be around the half-term evaluations), followed by a decelerating drop thereafter. Interestingly, the F2F students are showing a partial "bounce back" after the mid-term slump, which does not seem to happen for the online students. However, even when the F2F students do recover their LTL levels, the LTL evidence generally stays below the median. For both teaching modalities, there is another local dip near the end of the session (possibly around the time of

Figure 9. (a: above) the Heuristic Net for BIO181; (b: below): the Heuristic net for PSY101. Within each of the figures, the authors have added a magnified image of weeks 5-10 to show the details of the patterns



the final evaluations). Note that the assumption underlying this analysis was that the behavioural patterns of sessions A and B were similar and were therefore combined. Future analysis, however, should look into potential differences between them.

To further examine whether the learning design or the content might have an impact on the pattern of change of the LTL evidence levels throughout the term, the researchers compared the heuristic nets from

## ***Mind the Gap***

the biology course BIO181 (in Figure 9(a)) with that of the psychology course PSY101 (Figure 9(b)). Given that the two courses differ widely based on factors such as instructional design, overall learning goals, content, criteria, and evaluation methods, they hypothesized there would be a visible difference in the process flows in the two courses. On the contrary, both the heuristic models were shown to be quite similar in their design. This might be due to LTL's factors being relatively resilient or not affected by the learning design. However, it might also be that the process models or the LTL classifications are not sensitive enough to those differences and validating it with other courses could shed some light on it.

Within each week, the researchers observed patterns of learners transitioning from zero engagement to positive engagement through feedback (whereas in the biology course it has also transferred to positive practice). This pattern is presumably signalling an effective LTL process flow. However, the patterns during weeks 8 and 9 were strikingly different from the rest of the weeks, where a lower translation rate to LTL evidence of practice and engagement were found. This is consistent with the findings shown in Figure 8, where there are dips in LTL levels in the middle of both the long and short sessions and another dip towards their end.

## **Summary of the Results**

**RQ1:** A dimension reduction analysis has shown that academic performance, engagement, and LTL are three distinct dimensions, although the standardized assignment scores (which are a strong proxy for academic performance) also loaded on all three components. Although the LTL component contributed the least to the variation in the data, to answer the first research question the researchers suggest that LTL can be referred to as a dimension on its own with a moderate relationship to academic performance. It was also interesting to note that, similarly to the moderate relationship between engagement and academic outcome in higher education (Kent et al., 2016), the data also provide evidence for a moderate relationship between academic outcome and LTL. This was more evident for students with above-average academic performance, which is consistent with evidence about the relationship between academic performance and engagement (Boulton et al., 2018). It might suggest that for students on the below-average side, academic performance might be also affected by factors other than LTL.

The cluster analysis identified three distinct clusters based on three components (engagement, academic performance, and LTL). The clusters were characterized as the Versatile Achievers (cluster 1), the Engaged (cluster 2) and the Disengaged (cluster 3). From a process mining perspective, the Disengaged cluster registered the highest number of instances of negative engagement and subsequent negative assessment. These observations align with the poor academic performance and engagement levels associated with this cluster, but also might suggest that low engagement precedes low academic performance. Clearly, the researchers do not have evidence of a causal relationship, but future research could look to validate that. Also, the Versatile Achievers cluster was characterized as the group with the highest grades. This is despite the fact that students in the Versatile Achievers cluster showed significantly lower engagement levels compared to the Engaged cluster students, and they registered fewer instances of positive practice levels following positive engagement. Even though the Engaged cluster students' engagement levels were higher and they were more likely to transform their positive engagement into positive practice, the evidence shows that they were not able to transform it into higher academic performance. Their grades were not reflective of their efforts. Interestingly, their LTL evidence was lower than the Versatile Achievers, which might explain this inability to translate engagement into performance.

**RQ2:** Both the online and F2F modalities exhibited similar LTL levels patterns across the weeks of the term, regardless of the content. They both exhibited a general decline and a further local decline during the middle of the term. However, the F2F students showed a partial recovery pattern after the half-term, which was not exhibited by online students. Given the small number of observations for F2F, the noise in the time series, and the limited number of courses in the sample, it is hard to know whether the recovery is real or whether the 15 week F2F time series is simply decreasing at a lower rate. Even if the recovery pattern were real, it is hard to hypothesize why this difference would be observed. It might be due to the modality or due to the different length of the sessions, as the online courses are half as long as the F2F courses; or it could be due to subject matter or instructional design differences (the longer, F2F course was just in biology). Regardless, it is apparent that LTL levels were generally in a decline throughout the term, both for online and F2F students.

## **CONCLUSION**

In this chapter, the authors have discussed the process of implementing a practical operationalization of a pedagogical construct (such as LTL), which usually remains largely theoretical or is explored in a qualitative manner. For universities to base students' assessment and feedback on those meaningful constructs, a semantic bridge must be used to connect the selected theoretical concepts with the actual data that is typically collected to assess students' progress. The case study reported here was carried out on an initial cohort of more than 4,000 students who took one of two undergraduate courses during 2019-2020. It undertook an end-to-end process that began from the university's decision to focus on the concept of LTL, through the development of an ontology, and on to a modeling and analysis framework dedicated to LTL.

## **LIMITATIONS AND FUTURE RESEARCH**

Data from further courses and departments should be added, for ecological and internal validity and to enrich the framework with the contexts of various instructional designs, assignment types, content domains, reflection mechanisms, and an enlarged cohort of students. To better understand if there are significant differences in students' LTL patterns between online and F2F teaching provisions, data from students participating in equal-length sessions would need to be obtained, which the researchers did not have. Furthermore, enriching the raw data layer with additional data sources to be transformed into both the dynamic and static datasets is crucial to further refine the LTL models. Specifically, enriching the dynamic dataset with further levels of granularity (such as daily and even per millisecond) could give us a more accurate understanding of students' behavioural patterns. Enriching the static (contextualized) dataset with data sources bearing more LTL-related semantics, such as students' reflection journals, could reduce the theory-practice gap even further.

An additional expert review process should be undertaken to review the rule-base, extending the initial review carried out internally by the researchers and via the literature review.

To validate the relevance of the findings about students' SRL progress, the researchers suggest that a set of standardized questionnaires be administered to validate students' SRL levels. Ideally, these could be administered to a sample of the students, at specific points in time during their studies to establish a

## **Mind the Gap**

longitudinal baseline. For example, the researchers suggest basing those questionnaires on the findings of a well-validated meta-analysis of adults' SRL and educational attainment (Sitzman & Ely, 2011) to emphasize the following dimensions: goal setting (Bernard et al., 2009), self-efficacy (Pintrich, 1991), effort and persistence (Elliot et al., 1999). In this work, the researchers did try to conduct this external validation, however not enough responses were obtained to enable the validation to take place, so this remains as a limitation and a future research opportunity (see bullet point *k* in Figure 2).

Reflecting on possible future research on LTL, the researchers suggest looking at questions such as - do students in different programs and/or different delivery methods (such as online vs. F2F) make different amounts of progress in LTL throughout their learning journey? Do they show different kinds of SRL capability by the time of their graduation? Do the identified clusters lead to differences in either the overall progress or kind of LTL? Further, future research should explore the association between LTL and SRL through the diversity, equity, and inclusion lenses, given that there might be significant variations across different groups (e.g., gender, ethnicity). The diversities that exist amongst the student population across universities may have important implications in how LTL and SRL may be conceptualized and studied.

Practically, this study's findings could suggest interventions, all of which should be closely validated. For example, the finding that LTL manifests a separate outcome dimension calls for a new outlook on the implementation of evaluation-feedback cycles. The finding that LTL levels were generally in a decline throughout the term might call for a half-term intervention to refresh and boost students' reflective capabilities.

To close the theoretical-practical gap, educational institutions must address bridging three main gaps between: (1) evaluating LTL vs. explicitly fostering it in the curriculum; (2) assessing LTL vs. universities' more traditional assessment regimes, and (3) the data that is typically collected about students vs. the data that is required to evaluate LTL. This chapter has been very much focused on the third gap (i.e., the data collection gap). However, addressing the data collection gap reveals an opportunity to also address the instructional design gap and the assessment gap. Although many studies have investigated the self-regulation processes of academic learning and the personal and contextual factors involved, it is still a challenge *"to build and test instructional models that support and promote SRL within contexts that are full of new information technologies"* (Nuñez, et al., 2011, p.275). Some classroom-based methods that can be used by instructors, such as paying attention to students' prior knowledge, providing clear feedback, helping students organize concepts and terminology for the particular discipline that they are studying, as well as explicitly getting students to reflect on their learning as distinct from the subject that they are learning, can help learners make sense of their learning. Instructors may need to be reminded that students might have difficulty deploying their metacognitive skills and that this difficulty might manifest itself in different ways and at different stages of learning and level. This is sometimes even more the case when using technology-enhanced learning environments, as these environments sometimes add extra difficulties due to their open nature and students not knowing how to deploy self-regulating processes while learning (Azevedo, 2005).

Finally, the use of ontologies as the basis of any analysis framework also opens the possibility of using semantic technologies to integrate publicly available ontologies and data sources, semantic search and analysis, and semantically enriched decision support systems (Sabou et al., 2005).

## ACKNOWLEDGMENT

The authors wish to thank the School of Life Sciences and the School of Social and Behavioral Sciences' Department of Psychology at ASU for their participation in workshops that lead to the ontologies and questions addressed in this paper. The authors also wish to thank the Office of the Provost at ASU and the Office of the Dean at ASU's EdPlus for their support and creative input.

## REFERENCES

- Baker, R., Ocumpaugh, J., & Andres, A. (2018). BROMP Quantitative Field Observations: A Review. In R. Feldman (Ed.), *Learning Science: Theory, Research, and Practice*. Academic Press.
- Berti, A., van Zelst, S. J., & van der Aalst, W. (2019). *Process mining for python (PM4Py): bridging the gap between process and data science*. arXiv preprint arXiv:1905.06169.
- Boulton, C. A., Kent, C., & Williams, H. T. (2018). Virtual learning environment engagement and learning outcomes at a 'bricks-and-mortar' university. *Computers & Education, 126*, 129–142. doi:10.1016/j.compedu.2018.06.031
- Carver, C. S., & Scheier, M. F. (1981). A control-systems approach to behavioral self regulation. In L. Wheeler (Ed.), Vol. 2, pp. 107–140). *Review of personality and social psychology*. Sage.
- Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research, 1*(2), 245–276. doi:10.1207/15327906mbr0102\_10 PMID:26828106
- The Campaign for Learning. (2007). *What is learning to learn?* <https://www.campaign-for-learning.org.uk/Handlers/Download.ashx?IDMF=07f4cb7b-6f5f-42d2-8831-842c5f189797>
- Chi, M. T. H., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science, 13*(2), 145–182. doi:10.1207/15516709cog1302\_1
- 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. In *Proceedings of the 8th International conference on learning analytics and knowledge (LAK)* (pp. 191-200). 10.1145/3170358.3170381
- Corbin, J., & Strauss, A. (2014). *Basics of Qualitative Research*. Sage.
- Downer, J. T., Booren, L. M., Lima, O. K., Luckner, A. E., & Pianta, R. C. (2010). The Individualized Classroom Assessment Scoring System (inCLASS): Preliminary reliability and validity of a system for observing preschoolers' competence in classroom interactions. *Early Childhood Research Quarterly, 25*(1), 1–16. doi:10.1016/j.ecresq.2009.08.004 PMID:23175598
- del Valle, R., & Duffy, T. M. (2007). Online learning: Learner characteristics and their approaches to managing learning. *Instructional Science, 37*(2), 129–149. doi:10.1007/11251-007-9039-0
- Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology, 53*(1), 109–132. doi:10.1146/annurev.psych.53.100901.135153 PMID:11752481

## **Mind the Gap**

Education Council. (2006) Recommendation of the European Parliament and the Council of 18 December 2006 on key competencies for lifelong learning. Official Journal of the European Union, 30.

Education Council. (2018). *Council Recommendation on key competences for lifelong learning*. ST/9009/2018/INIT, 1–13

Elliot, A. J., & McGregor, H. A. (2001). A 2 × 2 achievement goal framework. *Journal of Personality and Social Psychology*, 80(3), 501–519. doi:10.1037/0022-3514.80.3.501 PMID:11300582

Halpern, D. F. (1998). Teaching Critical Thinking for Transfer Across Domains. *The American Psychologist*, 53(4), 449–455. doi:10.1037/0003-066X.53.4.449 PMID:9572008

Hautamäki, J., Arinen, P., Eronen, S., Hautamäki, A., Kupiainen, S., Lindblom, B., Niemivirta, M., Pakaslahti, L., Rantanen, P., & Scheinin, P. (2002). *Assessing learning-to-learn: a framework*. National Board of Education.

Henderson, S. A. (2018). Pedagogical contraband: A phenomenological approach to understanding student engagement during simulations. *Dissertations and Theses @ UNI*. 936. <https://scholarworks.uni.edu/etd/936>

Hymavathy, C., Krishnamani, V., & Sumathi, C. P. (2014). Analyzing learner engagement to enhance the teaching-learning experience. In *Proceedings of the International Conference on MOOC, Innovation and Technology in Education (MITE)*. IEEE. 10.1109/MITE.2014.7020243

Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, 39(1), 31–36. doi:10.1007/BF02291575

Kent, C., Laslo, E., & Rafaeli, S. (2016). Interactivity in online discussions and learning outcomes. *Computers & Education*, 97, 116–128. doi:10.1016/j.compedu.2016.03.002

Lau, C., Sinclair, J., Taub, M., Azevedo, R., & Jang, E. E. (2017). Transitioning self-regulated learning profiles in hypermedia-learning environments. In *Proceedings of the 7th International Learning Analytics & Knowledge Conference (LAK)* (pp. 198-202). 10.1145/3027385.3027443

Lloyd, S. (1982). Least squares quantization in PCM. *IEEE Transactions on Information Theory*, 28(2), 129–137. doi:10.1109/TIT.1982.1056489

Matcha, W., Gašević, D., Uzir, N. A. A., Jovanović, J., & Pardo, A. (2019, March). Analytics of learning strategies: Associations with academic performance and feedback. In *Proceedings of the 9th International Conference on Learning Analytics & Knowledge* (pp. 461-470). 10.1145/3303772.3303787

Motz, B., Quick, J., Schroeder, N., Zook, J., & Gunkel, M. 2019. The validity and utility of activity logs as a measure of student engagement. In *Proceedings of the 9th International Conference on Learning Analytics & Knowledge (LAK)*. Association for Computing Machinery. 10.1145/3303772.3303789

Matrić, M. (2018). Self-regulatory systems: Self-regulation and learning. *Journal of Process Management. Nuclear Technology*, 6(4), 79–84. doi:10.5937/jouproman6-19338

- Molenaar, I., Horvers, A., Dijkstra, R., & Baker, R. (2019). Designing Dashboards to support learners' Self-Regulated Learning. *Proceedings 9th International Conference on Learning Analytics & Knowledge*, 1-12.
- Motz, B., Quick, J., Schroeder, N., Zook, J., & Gunkel, M. 2019. The validity and utility of activity logs as a measure of student engagement. In *Proceedings of the 9th International Conference on Learning Analytics & Knowledge (LAK)*. Association for Computing Machinery. 10.1145/3303772.3303789
- Pardo, A., Han, F., & Ellis, R. A. (2016, April). Exploring the relation between self-regulation, online activities, and academic performance: A case study. In *Proceedings of the 6th International Conference on Learning Analytics & Knowledge (LAK)* (pp. 422-429). 10.1145/2883851.2883883
- Pintrich, P., Smith, D., Garcia, T., & McKeachie, W. (1991). *A manual for the use of the motivated strategies for learning questionnaire (MSLQ)*. National Center for Research to Improve Postsecondary Teaching and Learning.
- Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological Bulletin*, 135(2), 322–338. doi:10.1037/a0014996 PMID:19254083
- Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance: A systematic review and meta-analysis. *Psychological Bulletin*, 138(2), 353–387. doi:10.1037/a0026838 PMID:22352812
- Sabou, M., Wroe, C., Goble, C., & Stuckenschmidt, H. (2005). Learning domain ontologies for semantic web service descriptions. *Journal of Web Semantics*, 3(4), 340–365. doi:10.1016/j.websem.2005.09.008
- Saint, J., Fan, Y., Singh, S., Gasevic, D., & Pardo, A. (2021, April). Using process mining to analyze self-regulated learning: a systematic analysis of four algorithms. In *Proceedings of the 11th International Learning Analytics and Knowledge Conference (LAK)* (pp. 333-343). 10.1145/3448139.3448171
- Schunk, D. H. (2012). *Learning Theories—An Educational Perspective* (6th ed.). Pearson Education.
- Spann, C. A., Schaeffer, J., & Siemens, G. (2017, March). Expanding the scope of learning analytics data: Preliminary findings on attention and self-regulation using wearable technology. In *Proceedings of the 7th international learning analytics & knowledge conference (LAK)* (pp. 203-207). 10.1145/3027385.3027427
- Suraworachet, W., Villa-Torrano, C., Zhou, Q., Asensio-Pérez, J. I., Dimitriadis, Y., & Cukurova, M. (2021). Examining the relationship between reflective writing behaviour and self-regulated learning competence: A time-series analysis. In *European Conference on Technology Enhanced Learning* (pp. 163-177). Springer. 10.1007/978-3-030-86436-1\_13
- Thurstone, L. L. (1947). *Multiple-factor analysis: A development and expansion of The Vectors of Mind*. Academic Press.
- Trivedi, S., Pardis, Z. A., & Heffernan, N. T. (2015). The utility of clustering in prediction tasks. *arXiv*, (September), 1–11.
- Van Der Aalst, W. (2012). Process mining. *Communications of the ACM*, 55(8), 76–83. doi:10.1145/2240236.2240257

## **Mind the Gap**

Wolters, C. A., Pintrich, P. R., & Karabenick, S. A. (2005). Assessing academic self-regulated learning. In *What do children need to flourish?* (pp. 251–270). Springer. doi:10.1007/0-387-23823-9\_16

Weijters, A. J. M. M., van Der Aalst, W. M., & De Medeiros, A. A. (2006). Process mining with the heuristics miner-algorithm. *Technische Universiteit Eindhoven. Tech. Rep. WP, 166*, 1–34.

Zimmerman, B. J. (1989). A social cognitive view of self-regulated academic learning. *Journal of Educational Psychology, 81*(3), 329–339. doi:10.1037/0022-0663.81.3.329

Zimmerman, B. (2002). Becoming a Self-Regulated Learner: An Overview. *Theory into Practice, 41*(2), 64–70. doi:10.1207/15430421tip4102\_2

## **KEY TERMS AND DEFINITIONS**

**Clickstream:** Data are traces of users' operations within digital systems, such as which objects users clicked on and when and the duration through which they engaged with them.

**Cluster Analysis:** Is a common statistical analysis technique, focused on grouping a set of entities (e.g., students' data) in such a way that entities clustered together are concerned as more similar to each other than to those in other clusters.

**Learning Management System (LMS):** Is a software application used by educational institutions to plan, implement and assess students' learning process.

**Learning to Learn (LTL):** Is conceptualized as a process of improvement in the self-regulated learning (SRL) abilities of students.

**Ontology:** Is a formal, machine-readable representation of a specific body of knowledge. The ontology would typically include the definition of the relevant concepts, categories, properties and the relationships between them.

**Process Mining:** Is a data science technique used to analyze and identify common process-based patterns, based on event logs.

**Self-Regulated Learning (SRL):** Refers to students' abilities to use their metacognition, planning, monitoring, and evaluating their own learning process.

## APPENDIX 1: EXAMPLES FOR THE LTL RULE BASE

Table 4. a list of the LTL engineered features and examples of the rules applied to compute them

Engineered Features	Descriptions	Examples rules
LTL classification	This feature classifies the snapshots with the type of LTL activity. The possible classification types are: Analysis, Assessment, Engagement, Feedback, Motivation, Practice, Reflection and Review.	A snapshot will be classified as Engagement if it's about students' engagement with their study (using indications such as the number of sessions, clicks, duration etc); Motivation if a student submits a bonus assignment; Reflection if a student engages with a reflective journal; Analysis if a student is involved in research; Assessment if a student submits a final exam; Feedback if the instructor has inputted on a student's assignment; Practice if the student has submitted a lab work and; Review if their assignment is related to the lecture.
LTL evidence	The LTL evidence is on a numerical scale, with values from -3 to 3, which indicates the strength of the evidence related to LTL that is associated with different snapshots. Values from -3 to -1 would be assigned if the evidence is considered to be associated negatively with LTL, 0 if they are not considered indicative of LTL at all, or 1 to 3 if the evidence is considered to be associated positively with LTL.	For example, LTL evidence would be assigned as -2 when the snapshot is about an assignment that was not submitted. In contrast, when a snapshot indicates a submission of an assignment that is classified as a Reflection snapshot, the LTL evidence would be assigned +3, as this is a strong positive indication for LTL.
Course-standardized grade	For an assignment submission snapshot, it compares a student's <i>grade</i> to all <i>students' average grades</i> (in a course/term/session)	HIGH: if the student's grade is equal to or higher than the average course grade. LOW: if the student's grade is lower than the average course grade.
Self-standardized grade	For an assignment submission snapshot, it compares a student's grade to that student's average grade in all assignments (in a course/term/session)	HIGH: if a student's grade is equal to or higher than the student's average grade. LOW: if a student's grade is lower than the student's average grade.
Course-standardized LTL evidence	For an engagement snapshot, it compares the student's engagement level to all students' average engagement levels (within a course/term/session).	+2: if the student's engagement count is equal to or higher than the course average engagement count. -2: if the student's engagement count is lower than the course average engagement count.
Self-standardized LTL evidence	For an engagement snapshot, it compares the <i>student's engagement</i> level to their <i>average engagement</i> level (in a course/term/session).	+2: if the student's engagement count is equal to or higher than their average engagement count. -2: if the student's engagement count is lower than their average engagement count.

## APPENDIX 2: THE MERGED AND FILTERED LIST OF NODES USED IN THE PROCESS MODELS

All the merged combinations which do not appear below, did not cross the 0.05% frequency threshold.

1. LTL classification Feedback, LTL evidence level zero
2. LTL classification Assessment, LTL evidence level negative

## Mind the Gap

3. LTL classification Assessment, LTL evidence level positive
4. LTL classification Analysis, LTL evidence level positive
5. LTL classification Motivation, LTL evidence level positive
6. LTL classification Practice, LTL evidence level negative
7. LTL classification Practice, LTL evidence level positive
8. LTL classification Reflection, LTL evidence level positive
9. LTL classification Review, LTL evidence level positive
10. LTL classification Unclassified, LTL evidence level positive
11. LTL classification Engagement, LTL evidence level zero
12. LTL classification Engagement, LTL evidence level negative
13. LTL classification Engagement, LTL evidence level positive

## APPENDIX 3: FURTHER DETAILS ABOUT THE CLUSTERING ANALYSIS

Table 5. ANOVA table for the clustering solution, all the variables contributed significantly to the solution

	Mean Square	F	Significance
Component 1 (engagement)	735.36	1130.06	0.00
Component 2 (grades)	1288.45	3323.67	0.00
Component 3 (LTL)	634.81	908.75	0.00
reflection Average LTL evidence for	38.23	267.10	0.00
Average LTL evidence for motivation	1.20	175.60	0.00
Average LTL evidence for review	1.36	110.73	0.00
Average LTL evidence - unclassified	12.69	147.41	0.00
Average LTL evidence for analysis	10.10	156.88	0.00

### Post-Hoc Comparisons Between the Clusters

A Games-Howell post hoc test was used because the homogeneity of variance was violated. The main significant results are detailed below:

Significant differences in Component 1 (engagement)

- The Engaged cluster had a statistically significant higher level of Comp1 (engagement) than the Versatile Achievers cluster (mean increase of 0.72), 95% CI [0.62, 0.81] ( $p < 0.001$ ) and than the Disengaged cluster (mean increase of 1.34), 95% CI [1.27, 1.41], ( $p < 0.001$ ).
- The Versatile Achievers cluster had a statistically significant higher level of Comp1 (engagement) than the Disengaged cluster (mean increase of 0.62), 95% CI [0.51, 0.73], ( $p < 0.001$ ).

Significant differences in Component 2 (grades)

- The Versatile Achievers cluster had a statistically significant higher level of Comp2 (grades) than the Engaged cluster (mean increase of 3.19), 95% CI [3.04, 3.35], ( $p < 0.001$ ), and than the Disengaged cluster (mean increase of 3.24), 95% CI [3.08, 3.40], ( $p < 0.001$ ).

A significant difference in Component 3 (LTL)

- The Versatile Achievers cluster had a statistically significant higher level of Comp3 (LTL) than the Disengaged cluster (mean increase of 1.29), 95% CI [1.16, 1.43], ( $p < 0.001$ ).
- The Engaged cluster had a statistically significant higher level of Comp3 (LTL) than the Disengaged cluster (mean increase of 1.24), 95% CI [1.15, 1.32], ( $p < 0.001$ ).

Significant differences in LTL evidence for reflection

- The Versatile Achievers cluster had a statistically significant higher level of LTL evidence for reflection than the Engaged cluster (mean increase of 0.56), 95% CI [0.39, 0.72], ( $p < 0.001$ ) and than the Disengaged cluster (mean increase of 0.54), 95% CI [0.37, 0.71], ( $p < 0.001$ ).

Significant differences in LTL evidence for motivation

- The Versatile Achievers cluster had a statistically significant higher level of LTL evidence for motivation than the Engaged cluster (mean increase of 0.098), 95% CI [0.05, 0.14], ( $p < 0.001$ ) and than the Disengaged cluster (mean increase of 0.097), 95% CI [0.05, 0.14], ( $p < 0.001$ ).

Significant differences in LTL evidence for review

- The Versatile Achievers cluster had a statistically significant higher level of LTL evidence for review than the Engaged cluster (mean increase of 0.10), 95% CI [0.05, 0.16], ( $p < 0.001$ ).

Significant differences in LTL evidence for unclassified LTL

- The Versatile Achievers cluster had a statistically significant higher level of evidence for unclassified LTL than the Engaged cluster (mean increase of 0.31), 95% CI [0.23, 0.40], ( $p < 0.001$ ) and than the Disengaged cluster (mean increase of 0.32), 95% CI [0.24, 0.41], ( $p < 0.001$ ).

Significant differences in LTL evidence for LTL classification of analysis

- The Versatile Achievers cluster had a statistically significant higher level of evidence for LTL analysis than the Engaged cluster (mean increase of 0.29), 95% CI [0.20, 0.37], ( $p < 0.001$ ), and than the Disengaged cluster (mean increase of 0.27), 95% CI [0.18, 0.35], ( $p < 0.001$ ).