# Investigating Validity and Generalisability in Trace-Based Measurement of Self-Regulated Learning: A Multidisciplinary Study

Yige Song
The University of Melbourne
Parkville, Australia
yige.song@student.unimelb.edu.au

Eduardo Oliveira
The University of Melbourne
Parkville, Australia
eduardo.oliveira@unimelb.edu.au

Paula de Barba
Monash University
Clayton, Australia
Paula.deBarba@monash.edu

Michael Kirley
The University of Melbourne
Parkville, Australia
mkirley@unimelb.edu.au

Pauline Thompson
The University of Melbourne
Parkville, Australia
pauline.thompson@unimelb.edu.au

## Abstract

Self-regulated learning (SRL) skills are critical for effective learning and academic success. With the growing availability of trace data from students' online learning activities, researchers are increasingly leveraging this data to infer SRL processes. However, challenges remain regarding the validity of these inferences and their generalisability across diverse learning contexts. This study presents a structured approach to investigate these challenges by examining SRL behaviours in a multidisciplinary university cohort. The dataset includes 76 baseline survey responses, over 300 daily SRL survey submissions, and more than 6,000 sequences of recorded learning actions as trace data. Using mixed linear models and sequence mining, the analysis is grounded in SRL theory and evaluated through machine learning performance metrics. Our findings indicate consistent within-person patterns of SRL and online learning behaviours, supporting the concept of transferable, holistic skill development. Additionally, the results validate the trace-based detection of SRL engagement but highlight limitations in accurately detecting planning and reflection phases. These findings underscore the potential of automating SRL engagement detection while emphasising the need for multi-modal approaches to capture the full spectrum of SRL processes comprehensively.

## CCS Concepts

• **Applied computing** → **Education**.

## Keywords

Learning Analytics, Blended Learning, Trace data, Self-Regulated Learning, Trace-SRL, Mixed Linear Model, Sequence Mining

**ACM Reference Format:**
Yige Song, Eduardo Oliveira, Paula de Barba, Michael Kirley, and Pauline Thompson. 2025. Investigating Validity and Generalisability in Trace-Based Measurement of Self-Regulated Learning: A Multidisciplinary Study. In *LAK25: The 15th International Learning Analytics and Knowledge Conference (LAK 2025), March 03–07, 2025, Dublin, Ireland*. ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3706468.3706511



*LAK 2025, Dublin, Ireland*

ACM ISBN 979-8-4007-0701-8/25/03
https://doi.org/10.1145/3706468.3706511

## 1 Introduction

Self-regulated learning (SRL) involves students' ability to plan, monitor, and reflect on their learning activities [22, 46]. SRL skills are crucial for effective learning [22, 46]. Therefore, the measurement and support of students' SRL skills have become key areas of research in learning analytics. Traditionally, SRL skills can be measured by surveys or ongoing observations of students learning behaviours, which can be expensive and time-consuming [7, 43]. The increasing availability of trace data — records of student interactions with online learning environments, provides significant potential to automatically assess SRL at scale. To measure SRL from trace data, researchers need to translate trace data into meaningful SRL indicators (i.e., **trace-SRL**), and this area of research has received significant attention in the past few years [11, 26, 27]. The success of trace-SRL research can improve our understanding of students' online learning behaviours. Also, the result of trace-SRL research can guide us to develop interventions to support students' SRL skills at scale, ultimately improving students' learning outcomes.

Existing trace-SRL research (e.g. [11, 26, 27]) has focused on extracting learning patterns from trace data as indicators of SRL processes such as planning, engagement, and reflection. While promising, trace-SRL research faces two major challenges: **validity** (i.e., the accuracy of inferences about students' SRL processes) [11, 42] and **generalisability** (i.e., the applicability of methods across different learning contexts) [10]. Many existing trace-SRL frameworks are developed based on SRL theories (i.e., **theory-driven**; for example, see [26]), yet they have not been validated by other measures of SRL such as real-time survey data, raising concerns about their accuracy in accessing SRL from trace data. Additionally, regarding the trace-SRL frameworks that have been derived using trace data and other measures of SRL (i.e., **data-driven**; for example, see [11]), their generalisability is often limited. This is because these frameworks are usually developed within context-specific settings and are not designed to be applied in different disciplines [10, 11].

This study investigates the validity and generalisability of trace-SRL by conducting research in a multidisciplinary context, focusing on the same group of students enrolled in two first-year subjects. The research aims to advance the existing learning analytics literature and trace-SRL research in the following ways:

(1) For validity: It integrates daily SRL survey data with trace data for trace-SRL investigation. Additionally, it incorporates both data-driven and theory-driven approaches [11], allowing for a comparison of the two methods for the validity evaluation.
(2) For generalisability: The study takes place in multidisciplinary subjects, investigating the generalisability of trace-SRL research while offering insights into the consistency and variation of SRL processes across contexts.

This research contributes to the field by highlighting key considerations in the interpretation of trace data, towards more effectively supporting and improving students' SRL skills across diverse learning contexts.

# 2 Background

## 2.1 Trace Data and Self-Regulated Learning

In the past decade, learning management systems (LMS) such as Canvas have been widely adopted by education institutions to manage resources and deliver materials. At the same time, students' learning activities online have been recorded and saved in the form of trace data (aka clickstream data or log data), which provides an objective measure of students' online learning behaviours [42, 45]. Each record of trace data corresponds to a **learning action** that the student performed in their learning. By investigating the timestamp and the URL of a trace data record, researchers can derive the learning action represented by this trace record, such as submitting an assessment or viewing a specific lecture [34]. When learning a subject, each student can perform hundreds, if not thousands, of online learning actions, making trace data a valuable resource for understanding the students' process of learning.

A good understanding of how students manage their learning and engage with learning materials is crucial for interpreting the trace data. The widely-adopted self-regulated learning (SRL) theory characterises learning as the process involving the regulation of behaviour, cognition, emotions and motivation [3, 22, 46]. Most developed models broadly agree that SRL is a cyclical process including **three phases**: *planning* (aka preparation, forethought), *engagement* (aka enactment, monitoring, performing) and *reflection* (aka adapting, evaluation) [22, 46]. In the planning phase, learners analyse the task, set goals and make personal learning plans. In the engagement phase, learners execute the task by carrying out learning actions. At the same time, they control their learning using various strategies and monitor their performance. In the reflection phase, learners reflect on their learning, evaluate the learning outcome and adjust their learning plans accordingly, which completes the SRL cycle [3].

## 2.2 Trace-SRL

Numerous research shows that students' SRL skills can be reflected by their online learning activities recorded in trace data. Considering students' online learning activities as **sequences of learning actions** (e.g. a video view followed by assignment submission, followed by ...), students with high SRL skills have longer and more frequent sequences of learning actions recorded in trace data [9]. Moreover, in these action sequences, students reported with high SRL skills can exhibit certain patterns of learning, which can not be observed for students reported with low SRL skills [13, 20]. Due to the connection between trace data and students' self-reported SRL, learning analytics researchers have been attempting to translate trace data into meaningful SRL indicators. This area of research, known as **trace-SRL**, aims to directly access and measure students' SRL processes from trace data [32].

A typical trace-SRL study involves three key steps: (1) developing the **trace-SRL framework**, (2) extracting SRL processes from trace data using the trace-SRL framework, and (3) analysing and interpreting the patterns from these SRL processes.

The trace-SRL framework developed in *the first step* is commonly in the form of a **sequence library** - a map between sequences of learning actions from trace data and its underlying SRL processes. To build this library, two different approaches can be used: theory-driven and data-driven [11]. As the name suggests, the **theory-driven** approach starts with SRL theory, and then hypothesises how SRL processes can be manifested through sequences of learning actions. For example, Salehian Kia et al. [27] incorporated Winne and Hadwin's COPES SRL theory [41] into their specific study context to define the mapping between SRL processes and sequences of clicks. Similarly, Saint et al. [26] constructed their sequence library by adapting microanalytic SRL theories [12, 31]. On the other hand, the **data-driven** approach does not rely on SRL theories. It works by aligning trace data with other SRL measures such as self-reports. An action sequence is added to the trace-SRL framework if it consistently co-occurs with an SRL process recorded in other SRL measures, as this co-occurrence indicates a strong association between this action sequence and the SRL process [11].

Once the sequence library is established, the *second step* is to implement this trace-SRL framework to extract SRL processes from trace data. For instance, Saint et al. [26] used regular expressions (REGEX) to detect the presence of SRL processes in all action sequences. Finally, the *third step* involves analysing and interpreting the patterns from the detected SRL processes, and this can be done using various techniques depending on the research question. Researchers have been using clustering and process mining to evaluate patterns such as the phase transitions [19] and temporal characteristics [25] of SRL. Alternatively, sequence mining techniques can be applied to identify the frequent action sequences associated with each SRL phase [20].

## 2.3 Validity and Generalisability Trade-Off in Trace-SRL: Can We Have Both?

One advantage of measuring SRL from trace data is that the trace data is automatically collected by the LMS or other digital learning platforms. Therefore, this approach is inexpensive and can be easily scaled without interfering with students' learning. However, the approach of measuring SRL directly from trace data has yet to be widely adopted in research and educational practice due to two major challenges — **validity** and **generalisability** [10, 42], as introduced in the introduction.

For the trace-SRL framework to be valid, the trace data indicators must accurately measure the SRL phase they are intended to represent [42]. One study that aims for improved validity is Fan et al.'s research [11], which derives a trace-SRL framework and improves it using the "think-aloud" protocol as the reference point. This approach enhances the validity of trace-SRL research, because the think-aloud protocol is often considered as the "gold standard" of accessing learner's SRL processes due to its real-time, detailed report on the SRL processes during learning. However, applying this approach in a more authentic learning environment is a challenge. As the researchers mentioned, the limitation of the "think-aloud" protocol is that it is labour-intensive and intrusive to students' learning [11].

For a trace-SRL framework to be generalisable, it should be developed using a scalable approach that is applicable to broader learning contexts [10]. The research by Saint et al. [26], mentioned earlier, can be considered generalisable, as their framework is applicable to an authentic university subject's LMS. However, their framework is theory-driven and has not been validated using alternative SRL measures, which compromises its validity.

The above paragraphs highlight the trade-off between validity and generalisability in trace-SRL research: Ensuring validity often requires continuous surveying or observation to accurately capture students' SRL processes during learning; Conversely, achieving generalisability involves creating a model that works under different real-world learning contexts, which makes the systematic validation of the framework difficult. To investigate this trade-off, we need (1) a trace-SRL framework that can be applied to diverse learning contexts, (2) a more scalable way of accessing students' SRL processes in real-time or near real-time to validate the trace-SRL framework, and researchers should (3) conduct this validation in an authentic, multi-context learning environment.

## 3 Current Research

This research investigates the generalisability and validity of the trace-SRL framework in an authentic multidisciplinary university setting. It triangulates trace data (which represents online learning activities), the just-in-time survey data (which indicates the underlying SRL processes during learning activities), and a theory-driven trace-SRL framework [26] (which is the theoretical mapping between online learning activities and the underlying SRL processes).

Previous research has shown that students' learning activities and SRL can vary across different learning contexts [18, 21]. Given our multidisciplinary setting, we first evaluate how SRL survey data and trace data vary or stay consistent across these subjects for the same group of students.

- **RQ1**: What is the *variation* and *consistency* of SRL survey data and trace data for the same group of students across different subjects? In other words, when a group of students study different subjects, do their SRL and online learning behaviours vary?

The development of the trace-SRL sequence library can be either data-driven or theory-driven, as described in Section 2.2. For a comprehensive investigation of trace-SRL in multiple subjects, this research uses both data-driven and theory-driven approaches and then compares and contrasts their results.

We first use a data-driven approach to independently develop trace-SRL sequence libraries for each subject. This leads to our second research question:

- **RQ2**: In multidisciplinary subjects, how can we *develop* a data-driven trace-SRL sequence library and *evaluate* it using a theory-driven trace-SRL framework? In other words, can we identify indicators from trace data to capture the SRL process (informed by survey responses), that are generalisable across multidisciplinary subjects?

Next, we use a theory-driven approach to implement Saint et al.'s yet-to-be-validated trace-SRL framework [26], which was developed based on microanalytic SRL theories [12, 31]. This leads to our third research question:

- **RQ3**: How can we *implement* an existing theory-driven trace-SRL framework and *validate* it in a multidisciplinary subject setting? In other words, how accurately can an existing trace-SRL theory identify the different phases of SRL processes from trace data in our multidisciplinary setting?

## 4 Methodology

### 4.1 Overview

The process flow of this study is illustrated in Figure 1. This methodology section follows the flow of Figure 1, starting with the study context (Section 4.2), followed by the data collection and preprocessing steps (Section 4.3), and concluding with the detailed description of data analysis for addressing each research question (Section 4.4).

### 4.2 Study Context

The study is conducted in an Australian university for two first-year subjects, a computing subject (COMP) and a science subject (SCIE). COMP is a core subject for students pursuing a computing major, and the final result of the subject is counted in their final weighted average mark of the degree. SCIE is a compulsory subject for the Bachelor of Science degree, covering sustainability and scientific methods. The final result of SCIE is marked on a pass (i.e., score > 50%) or fail (i.e., score < 50%) basis, and the students taking this subject need to secure a pass result to complete their degree. Both subjects are delivered in a blended learning mode, combining in-person lectures and workshops, with corresponding materials and learning exercises provided online. The assessments in COMP include weekly programming tasks, individual programming projects, and mid-term and final exams. In contrast, the assessments in SCIE consist of weekly quizzes and a combination of individual and group reports.

The participants of the research are 39 first-year bachelor of science students. They have signed the ethics form (#22276), voluntarily agreeing to participate in the research. 38 of them were actively involved in the research process, whose response was included in this research. All of the participants took both COMP

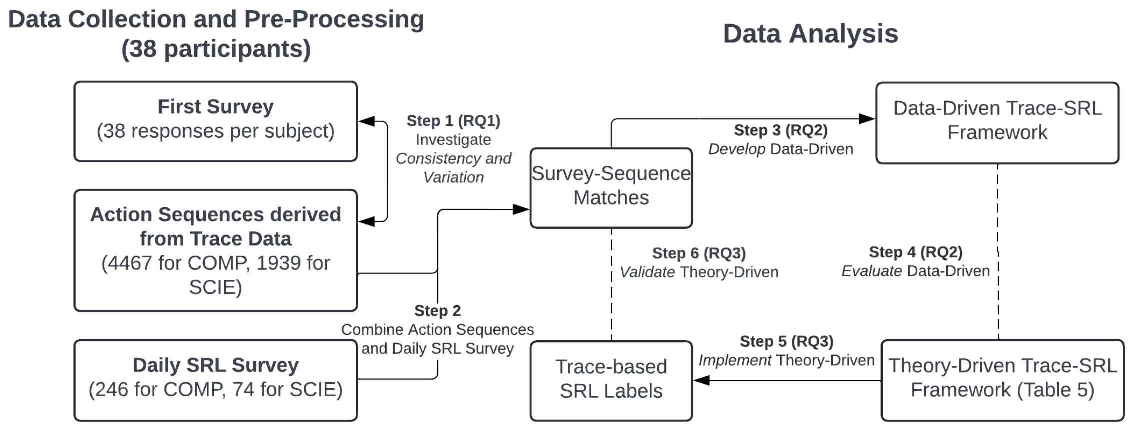


**Figure 1: The process flow of this study, with each arrow outlining a research step. The same research process was applied for both COMP and SCIE subjects.**

and SCIE in the semester 1 of 2022. There were 17 weeks in that semester: twelve teaching weeks (i.e., 1, 2, ..., 12); a week for the mid-semester break between week 7 and 8; an exam revision week; and three exam weeks.

## 4.3 Data Collection and Pre-Processing

There are three types of data collected for this study: A survey collected at the beginning of the semester (i.e., first survey) about SRL skills; a survey collected daily in weeks 4, 6, 8 and 9 (i.e., daily SRL survey) about SRL related learning activities completed in that day; and the trace data automatically recorded in the Canvas LMS system. The first survey serves as the baseline indicator for students' SRL skills, and it is used for answering RQ1; the daily SRL survey is collected "just-in-time" at the end of each day, allowing for accurate data collection at scale in the natural, large-scale and longitudinal setting over the semester, while avoiding the issues of memory distortions or forgetfulness seen in retrospective reports. It is used for answering RQ2 and RQ3; the trace data is used for answering all the research questions.

*4.3.1 First Survey.* For the first survey, participants were asked to answer an SRL questionnaire for both subjects in Likert scales. The SRL questionnaire is an eight-item rubric, examining students' SRL skills and strategies specific to their learning on the LMS. The eight-item SRL rubric is based on an upcoming work by de Barba et al. (under review), incorporating four validated SRL scales, each abbreviated as MSLQ [24], OSLQ [1], MAI [29], and A-SRL [4]. These scales have been verified and are widely used in higher education, encompassing 211 items aimed at measuring different aspects of SRL skills.

Additionally, students were also asked about their motivation and the priority for each subject in the first survey. This is because these two factors directly influence students' goal-setting in the SRL planning phase, contributing significantly to the overall SRL process [3, 22]. The motivation questions are based on de Barba et al. [8], covering aspects including value beliefs, self-efficacy, performance or mastery approach goals, individual interests, and control beliefs [23]. The priority questions asked students to rank the importance of each subject in relation to other enrolled subjects and their non-academic commitments [17].

*4.3.2 Daily SRL Survey.* The daily SRL survey used in this study is a concise SRL questionnaire adapted from Harrison and Vallin's short version of the metacognitive awareness inventory (MAI) scale [14]. For each subject, students were able to select multiple answers in a multiple-choice format, indicating the learning activities they performed on the LMS during the day. Each learning activity was then mapped to a corresponding SRL phase. Specifically, the activity "investigating before starting" indicates SRL phase 1 – planning; "reviewing materials", "clarifying understanding", and "learning new materials" correspond to SRL phase 2 – engagement; and "evaluating what they have learned so far" represents SRL phase 3 – reflection. The daily SRL survey link was sent to students at the end of each day in weeks 4, 6, 8, and 9, and the 38 participants generated 246 daily SRL survey responses for COMP and 74 for SCIE.

*4.3.3 Trace Data.* The trace data of this research was automatically recorded in the Canvas LMS and accessed via API requests. Each record of trace data corresponds to a **learning action** performed by a student, capturing details such as the student's ID, the URL of the visited site, and the timestamp of the action. To ensure a consistent analysis across subjects with different designs, previous research has proposed a general method for categorising learning actions based on their URLs (e.g., [28, 39]). Building on these works, our research classifies learning actions into four categories: *materials* for accessing subject content; *assessments* for accessing assessments and past submissions; *gateway* for visits to the subject's main page; and *other* for miscellaneous and administrative tasks. The details of this categorisation are outlined in Table 1.

According to the education psychology literature, students do not perform each learning action in isolation. Instead, they effectively regulate their effort to complete a sequence of (relevant) learning actions without interruption [36]. Therefore, to better represent this process, existing research typically groups learning actions

| Label | Corresponding LMS page | Symbol |
|---|---|---|
| materials | modules, files, videos, wiki pages | m (M) |
| assessments | assignment, quiz, grades, submissions | a (A) |
| gateway | subject main view | g (G) |
| other | announcements, subject forum... | o (O) |
| start/end | N.A. | _ |

**Table 1: Learning actions and their corresponding labels, corresponding LMS pages and symbols. A lowercase symbol represents a single visit to the corresponding LMS page, while an uppercase symbol indicates continuous visits. "start" and "end" action represents the start and end of an action sequence.**

that are temporally adjacent to each other, and uses these groups as the unit for further analysis [5, 9, 26, 27, 34]. Each group of adjacent learning actions are termed **sequence of learning actions**, or simply, sequence or action sequence in our context. To derive the sequences, a critical assumption on determining whether two consecutive learning actions are part of the same sequence should be made [16]. Existing research typically makes this assumption depending on the design of the subject being investigated [34]. We grouped learning actions occurring within one hour of each other together because one hour is the typical duration of the lectures and workshops in COMP and SCIE. In the end, 4467 sequences were generated for COMP, and 1939 sequences were generated for SCIE.

## 4.4 Data Analysis

*4.4.1 RQ1 - Consistency and Variation of Survey and Trace Across Contexts.* For RQ1, we investigated the first survey data and trace data between two subjects within the same group of students (step 1 in Figure 1). The data being used is shown in Table 2. Specifically, the subject-specific SRL, motivation and priority collected in the first survey were included. Regarding the trace data, we calculated some commonly used summary statistics of action sequences, including their duration, number of actions and total counts [9, 28]. We also included the number of active days as part of the trace data measure because the regularity of students' learning signifies their level of engagement [15].

A multi-method approach combining statistical tests, correlation analysis and mixed models was used to answer RQ1. The analysis was completed using Python scipy.stats and statsmodels library [30].

Specifically, the analysis involves three steps: cross-subject comparison, within-student correlation and variance decomposition. For cross-subject comparison, we used Wilcoxon signed-rank test [40] to compare survey and trace data measures across subjects at a group level. For within-student correlation, we investigated whether individual students show consistency in these measures across the two subjects. Given that the trace data and survey data were collected and processed using the same way across subjects, we used Pearson correlation analysis to assess their numerical linear relationships. For variance decomposition, we used mixed linear models to calculate how much of the variance in trace and survey measures is attributed to subject-specific versus student-specific factors [15]. Since this is not a prediction task, we created two null models for each of the survey and trace measure, with each null model containing only random effects. In the first model, the random effect is the subject (i.e., COMP or SCIE), and in the second, it is the student (i.e., their ID). We calculated the Intraclass Correlation Coefficient (ICC) for each model to estimate the proportion of variance explained by each random effect: the student ($ICC_{\text{Stud}}$) and the subject ($ICC_{\text{Subj}}$). We also checked the relevant assumptions for linear models (such as normality of residuals, homoscedasticity of residuals, and absence of influential points) and reported only the results that satisfy these assumptions.

*4.4.2 RQ2 - Development and Evaluation of Data-Driven Trace-SRL.* To answer RQ2, we used a data-driven approach to develop the trace-SRL sequence library. The data used for this analysis includes students' daily SRL survey responses and trace data from both subjects.

The first step in this process is to map daily SRL surveys with their corresponding trace data records. To achieve this, we generated **survey-sequence matches** (step 2 in Figure 1) by aligning each daily SRL survey response with the same students' sequences of learning actions that were recorded on the same day. This resulted in 211 unique daily SRL surveys in COMP being aligned with a total of 466 action sequences, while 47 unique daily SRL surveys were aligned with 69 action sequences in SCIE. To ensure a one-to-one alignment, we concatenated all action sequences occurring on the same day in further analysis. In the future analysis, we will refer to the SRL phase reported in the daily SRL survey as the "**survey-based SRL label**".

Recall that students might report engaging in multiple SRL phases within a day. We focused on analysing the subset of survey-sequence matches which has only one survey-based SRL label. This ensures the extraction of phase-exclusive trace indicators. In COMP, there are 9 survey-sequence matches for SRL phase one (planning). This means that, among all the survey-sequence matches, 9 instances have survey data where students reported only engaging in SRL phase one on that day. Similarly, 108 survey-sequence matches were observed for phase two (engagement), and 10 for phase three (reflection) in COMP. In SCIE, the survey-sequence matches for SRL phases one, two, and three are 5, 21, and 3, respectively. To illustrate the results, we present several examples of survey-sequence matches that have one and only one survey-based SRL label, as shown in Table 3.

The examples demonstrate that for the same student in a subject, their learning behaviours exhibited from learning actions can have consistent characteristics (e.g. sequence length and duration) [37]. Also, deriving the SRL phase unambiguously from trace data is difficult and sometimes impossible [11] (see the student with ID 85 in Table 3).

Extending beyond the above examples, we then systematically extracted the frequent action sequences associated with each survey-based SRL labels (step 3 in Figure 1). This can be achieved using sequence mining algorithms, which have been previously applied in learning analytics research (e.g., [20, 47]). In our study, we identified common N-grams (i.e., subsequences containing N actions in an action sequence) within the survey-sequence matches corresponding to each survey-based SRL label. For example, for two sequences "*mga*" and "*aga*", the 1-gram subsequence *g* and *a* have

| Data Source | Measure | Description |
|---|---|---|
| First Survey | SRL | The average score in the SRL questionnaire |
| | Motivation | The average score in the motivation questionnaire |
| | Priority | The subject priority in ranks |
| Trace | Median Sequence Duration | The median duration of action sequences in seconds |
| | Median Sequence Length | The median number of learning actions in action sequences |
| | Total Sequences | The total number of action sequences |
| | Active Days | The number of days that students have action sequences on the LMS |

**Table 2: The survey data and trace data measures used in the investigation of RQ1.**

| Student.ID | Day | Subject | Survey-Based SRL Labels | Action Sequence | Duration |
|---|---|---|---|---|---|
| 85 | W6 Fri | COMP | 1 - Planning | _gA_ | 00:00:09 |
| 85 | W8 Fri | COMP | 2 - Engagement | _gA_ | 00:01:12 |
| 85 | W6 Mon | COMP | 3 - Reflection | _gA_ | 00:00:21 |
| 22 | W6 Fri | SCIE | 1 - Planning | _mgAgamAm_gAgmamGmAga_ | 01:13:42 |
| 22 | W8 Fri | SCIE | 2 - Engagement | _omogAoaoaoaoaoaOamg_ | 01:54:54 |
| 22 | W4 Fri | SCIE | 3 - Reflection | _gMgamAgMgMaM_ | 01:02:52 |

**Table 3: Example survey-sequence matches that have only one survey-based SRL label for two participants. As shown, each survey-sequence match includes a survey-based SRL phase label obtained from the daily SRL survey. It also contains the learning activities completed online on that day, represented as a sequence of learning actions (see Table 1). For instance, *_gA_* indicates that the student begins by visiting the subject's main page, followed by continuous visits to subject assessments, and concludes their online learning activities for the day.**

the frequency of 2 because they appear in both sequences, whereas $m$ has the frequency of 1. Similarly, the 2-gram subsequences and their frequencies are $\{ga : 2, mg : 1, ag : 1\}$, and 3-gram subsequences are $\{mga : 1, aga : 1\}$.

To identify common N-grams for survey-sequence matches, we implemented the Generalised Sequential Pattern (GSP) sequence mining algorithm [35]. The algorithm starts by identifying subsequences consisting of a single action and then iteratively builds to include 2-action subsequences, discarding those that do not appear frequently. After conducting repeated experiments, we set the frequency threshold at 0.15, meaning only subsequences occurring in more than 15% of the time are included in the SRL subsequence library. In the end, we summarised the "typical subsequences" for each SRL phase to construct the trace-SRL sequence library (the result will be shown in Figure 6).

After obtaining this trace-SRL sequence library driven by the survey-based SRL labels, we compared it with the existing theory-driven framework from Saint et al. [26] (step 4 in Figure 1), introduced in the following section.

*4.4.3 RQ3 - Implementation and Validation of Theory-Driven Trace-SRL.* The investigation of RQ3 involves the implementation and validation of Saint et al.'s theory-driven trace-SRL framework [26]. Comparable to our data-driven sequence library developed for answering RQ2, the theory-driven framework also maps action sequences to SRL phases. The only difference is that the theory-driven framework further divides each SRL phase into various micro-processes. In our research, we adapted this framework and removed the action sequences that can not be obtained from our more generalised classification of learning action as shown in Table 1. The adapted theory-driven framework is shown in Table 4. Then, we used trace data and survey-based SRL labels together to perform the validation of this adapted framework.

To address RQ3, we first implemented the theory-driven framework (Table 4) to label each action sequence with the corresponding SRL phases (step 5 in Figure 1). Consistent with the original approach [26], we developed a regular expression (REGEX) parser to achieve this. Note that a sequence can be assigned multiple SRL phase labels. For example, according to the theory-driven framework, an action sequence of "*_gMg_*" should be labelled as both phases 1 and 2.

Because this theory-driven framework labels trace data with SRL phases directly, we can label all action sequences (i.e., 4,467 for COMP and 1,939 for SCIE), regardless of whether daily SRL survey responses are available. This also means that all survey-sequence matches created for RQ2 received **trace-based** SRL labels (by the theory-driven framework) in addition to their existing **survey-based** SRL labels. Since the trace-based SRL labels were calculated independently from the survey data, we performed exploratory data analysis to ensure that the trace-based SRL labels are comparable with the survey-based SRL labels to start with (details described in Section 5.3).

The next step is to validate the theory-driven framework (step 6 in Figure 1). The framework is considered validated if it can correctly label trace data (i.e., trace-based SRL labels) with the SRL phases reflected in the daily SRL survey (i.e., survey-based SRL labels). Therefore, the validation process was approached as a classification task - for each survey-sequence match created earlier, its trace-based SRL label is the predicted label and its survey-based SRL label serves as the true label. We calculated the true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN)

| SRL Phases | SRL Micro-Processes | Description | Action (Sub)sequence |
|---|---|---|---|
| Planning | Plan Materials | Setting a materials learning plan by accessing materials pages then switching back to the subject main view | [g/G][m/M][g/G] (i.e., g or G, followed by m or M, then g or G) |
| | Plan Assessments | Setting an assessment working plan by accessing materials pages then switch back to the subject main view | [g/G][a/A][g/G] |
| Engagement | Engage Materials | Consistently work on learning materials | M |
| | Engage on Assessments | Consistently work on subject assessments | A |
| | Strategy Change | Transit from materials to assessments | [m/M][a/A] |
| Reflection | Reflecting on Learning | Switch to materials after accessing assessments | [a/A][m/M] |

**Table 4: The adapted theory-driven trace-SRL framework [26], which is used in this research to answer RQ3.**

[33] for each SRL phase. For example, if the framework generates a trace-based label of SRL phase 1, but SRL phase 1 is not present in the survey-based label, this prediction should be classified as a false positive for SRL phase 1. The same process was applied to check SRL phases 2 and 3. Finally, commonly used performance measures for classification tasks, including accuracy, precision, recall, and F1 score [33], were employed to evaluate the performance of this task.

- Accuracy: $\frac{TP+TN}{TP+TN+FP+FN}$. This metric represents the proportion of correct predictions.
- Precision: $\frac{TP}{TP+FP}$. High precision indicates a low rate of false positives.
- Recall: $\frac{TP}{TP+FN}$. High recall indicates a low rate of false negatives.
- F1 score: $\frac{2}{Precision^{-1}+Recall^{-1}}$. A high F1 score reflects a good balance between precision and recall.

# 5 Results

## 5.1 RQ1 - Consistency and Variation of Survey and Trace Across Contexts

The result of our investigation on the cross-subject comparison, within-student correlation and variance decomposition of the first survey and trace data across subjects is shown in Table 5.

The cross-subject comparison result in Table 5 indicates that for the same group of students, all measures show significant differences (p-value $< 0.05$) between subjects. Specifically, students reported having significantly higher SRL, motivation and priority in COMP compared to SCIE, accompanied by more active days and more sequences of learning actions. This is likely due to SCIE being an overall less demanding, pass-or-fail graded but compulsory subject.

Despite the observed cross-subject differences, we found consistencies within the same student across subjects. For the first survey data, although students' motivation and priority for COMP versus SCIE showed no clear relationship, their SRL across the two subjects was significantly correlated (p-value $< 0.001$). Besides the first survey data, some trace measures of online learning behaviours also showed significant within-student correlations (Table 5).

To quantify the influence of subject design and student-specific factors on the variation in the first survey and trace data, we used mixed linear models to calculate the Intraclass Correlation Coefficient (ICC) for students and subjects. Due to violations of model assumptions, the results for some measure are not reported in Table 5. Among the reported results, the subject design explained over 50% of the variation in motivation and active days across the two subjects. However, the variation in the SRL measure was primarily attributed to individual student factors.

The overall results demonstrate the differences in survey and trace measures at the group level between subjects, while also highlighting the consistency of SRL and online learning behaviours at the individual student level.

## 5.2 RQ2 - Development and Evaluation of Data-Driven Trace-SRL

The data-driven trace-SRL sequence library developed using the GSP sequence mining algorithm is shown in Table 6. The results indicate that the sequence libraries tend to be subject-specific. That is, COMP and SCIE tend to have their own action subsequences for each SRL phase with minimal overlap. For example, a higher proportion of "_ga_" and a lower proportion of "m" are observed in COMP's phase 3 (reflection), whereas this pattern is seen in SCIE's phase 1 (planning).

Still, some overlaps exist in the trace-SRL sequence library across subjects. We compared our data-driven trace-SRL sequence library with the theory-driven framework adapted from Saint et al. [26] and identified a consistent pattern. Specifically, the common pattern is the "*ma*" subsequence, representing transitions from accessing learning materials to performing assessment-related actions. This pattern frequently appears in both COMP and SCIE when students reported engaging in SRL phase 2 (Table 6). This subsequence also appears in the SRL phase 2 in the theory-driven framework, as shown in Table 4. According to Saint et al. [26], moving from studying materials (m) to completing assignments (a) follows the metacognitive process of acquiring knowledge and checking understanding. Therefore, the presence of this action sequence indicates students' active engagement with the learning task on the LMS.

The overall results demonstrate the difficulty of developing a "one-size-fit-all" data-driven trace-SRL framework. However, one action sequence in the data-driven framework (i.e., "*ma*") stands

| Measure: $M$ | Cross-Subject Comparison | | Within-Student Correlation | | Variance Decomposition | |
|---|---|---|---|---|---|---|
| | $W$ | $P(M_{\text{COMP}} = M_{\text{SCIE}})$ | $r$ | $P(r = 0)$ | $ICC_{\text{Stud}}$ | $ICC_{\text{Subj}}$ |
| SRL | 31.5 | **<0.001** | 0.75 | **<0.001** | 0.58 | 0.16 |
| Motivation | 3.5* | **<0.001** | -0.01 | 0.92 | | 0.65 |
| Priority | 5.5 | **<0.001** | 0.12 | 0.53 | | |
| Median Sequence Duration | 235.5 | **0.049** | 0.13 | 0.45 | | |
| Median Sequence Length | 0* | **<0.001** | 0.78 | **<0.001** | | |
| Total Sequences | 0 | **<0.001** | 0.46 | **0.003** | | |
| Active Days | 0 | **<0.001** | 0.47 | **0.003** | | 0.72 |

**Table 5: The investigation results on the consistency and variation of first survey data and trace data across two subjects (i.e. COMP and SCIE). Wilconxon paired test was used to examine the cross-subject comparison, with the test statistics $W$ and p-value $P(M_{\text{COMP}} = M_{\text{SCIE}})$ reported. The * symbol indicates the use of normal approximation due to zero values. Pearson correlation analysis was used to evaluate the within-student correlation of COMP versus SCIE, with the correlation coefficient $r$ and p-value $P(r = 0)$ of the test statistics reported. Linear mixed models and the corresponding inter-class correlation were used to calculate the variance decomposition by student-specific factors ($ICC_{\text{Stud}}$) and subject-specific factors ($ICC_{\text{Subj}}$). Note that the result is shown only when the corresponding model converged and satisfied the linear mixed model assumptions.**

| | Phase 1 | Phase 2 | Phase 3 |
|---|---|---|---|
| COMP | a_ ↑<br>_gmg_ ↑<br>m_ ↓<br>ma ↓ | ***ma*** ↑ | _ga_ ↑<br>m ↓<br>am ↓ |
| SCIE | _ga_ ↑<br>m ↓ | am ↑<br>m_ ↑<br>***ma*** ↑<br>a_ ↓ | N.A. |

**Table 6: The data-driven trace-SRL sequence library developed. This framework was derived from survey-sequence matches, recording the action subsequences associated with each survey-based SRL (phase) label. After extensive testing, we decided to associate a sequence (subsequence) of actions with a given SRL phase if its proportion in this SRL phase is over 10% higher (i.e., ↑) than both (a) its proportion in all other phases, and (b) its overall proportion across all data. The same criteria applies if its proportion is 10% lower (i.e., ↓). Note that only the longest subsequences are presented (e.g., if *_gmg_* is eligible, include *_gmg_* rather than *gm*). Also note that for clearer representation, the action symbols in Table 1 are all converted to lowercase letters in this table (e.g., 'a' can indicate either a single assessment action or continuous assessment actions). Not all typical actions are included due to duplication and consideration of their meaning, and SRL phase 3 in SCIE is omitted due to insufficient sample size.**

out. It shows consistency across contexts and is also supported by the theory-driven framework.

## 5.3 RQ3 - Implementation and Validation of Theory-Driven Trace-SRL

To address RQ3, we implemented the theory-driven trace-SRL framework to generate trace-based SRL labels and validated it using survey-based SRL labels. Because of our multi-disciplinary setting, we first conducted explorative data analysis to ensure comparable results across subjects. We began the explorative data analysis by investigating into all of the action sequences (i.e., 4467 for COMP, 1939 for SCIE) and calculating the proportion of their corresponding trace-based SRL labels. The result is shown in Table 7.

Table 7 shows that the proportion of trace-based SRL labels and micro-processes in COMP and SCIE are overall consistent, with two observable differences: COMP has a higher proportion of "plan materials" micro-process, and SCIE has more "engage materials" micro-process. This can be attributed to the greater amount of reading materials and consecutive videos embedded within the LMS page for SCIE.

Next, we continued the explorative data analysis by comparing the proportion of trace-based and survey-based SRL labels. Both COMP and SCIE followed a weekly subject schedule, allowing us to conduct this analysis by weeks. For trace-based SRL labels, the proportion of a given phase was calculated by dividing the number of action sequences associated with that phase in a week by the total number of action sequences in that week. For survey-based SRL labels, the proportion was calculated by dividing the number of daily SRL surveys including that SRL phase by the total number of surveys collected in that week. The comparison results are presented in Figure 2.

The plot shown in Figure 2 indicates that overall, the proportions of both trace-based and survey-based SRL labels fluctuate over the weeks. However, the trace-based and survey-based labels do not align for all SRL phases. For phase 2 (engagement), the proportions of trace-based and survey-based labels are comparable, consistently showing similar trends over the weeks. In contrast, phase 1 (planning) and phase 3 (reflection) exhibit less alignment between the two methods.

As outlined in Section 4.4.3, the theory-driven framework can be validated by framing it as a classification task. The evaluation results of the classification are shown in Figure 3. Consistent with the proportion analysis presented in the previous Figure 2, this result also highlights an imbalance of SRL phase 2 versus SRL phases 1 and 3.

According to Figure 3, the classification accuracy of phase 2 exceeds 0.7, and phases 1 and 3 achieve slightly lower but close

| Trace-Based SRL Labels | Phase 1 | | Phase 2 | | | Phase 3 |
|---|---|---|---|---|---|---|
| SRL Micro-Processes | Plan Materials | Plan Assessments | Engage Materials | Engage Assessments | Change Strategy | Reflect |
| COMP | 0.198 | 0.146 | 0.353 | 0.655 | 0.271 | 0.200 |
| SCIE | 0.123 | 0.143 | 0.485 | 0.545 | 0.249 | 0.261 |

**Table 7: The proportion of trace-based SRL labels and SRL micro-processes among all action sequences, derived using the theory-driven framework in Table 4 and the regular expression (REGEX) parser.**

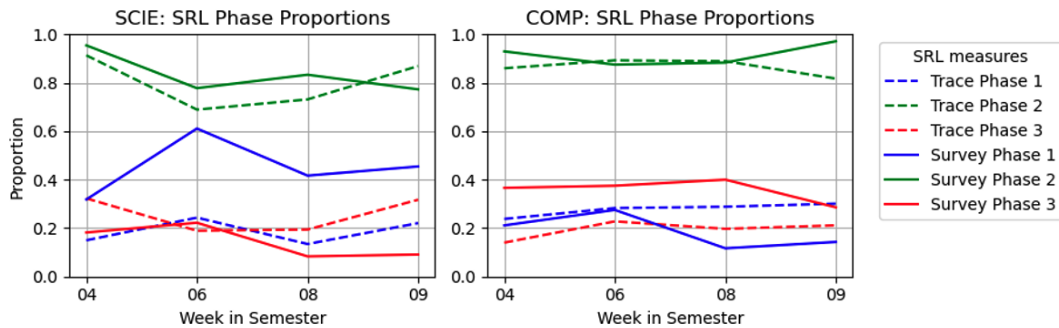


**Figure 2: Weekly SRL phase proportions in SCIE and COMP, comparing trace-based with survey-based SRL labels. Solid lines represent survey-based SRL phase labels, while dashed lines represent trace-based SRL phase labels.**

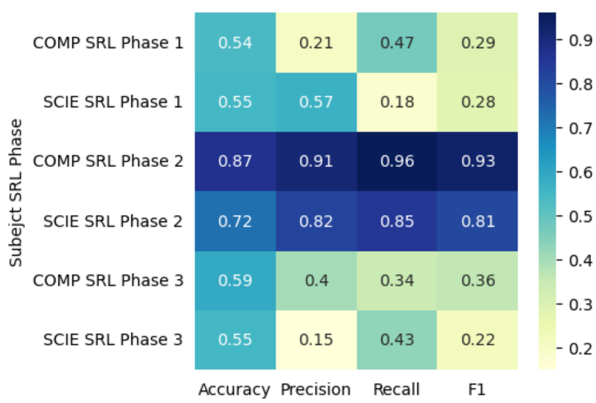


**Figure 3: Performance evaluation of trace-based SRL label in predicting survey-based SRL label across COMP and SCIE subjects. The heatmap displays the Accuracy, Precision, Recall, and F1 score for each SRL phase, with higher values indicating better predictive performance.**

accuracy. However, when examining precision and recall (i.e., evaluating false positive and false negative rates), the performance on phases 1 and 3 are much worse than phase 2. Specifically, for phases 1 and 3, the theory-driven framework either fails to capture enough true occurrences of the phase (i.e., low recall) or incorrectly predicts the presence of the phase (i.e., low precision), resulting in imbalanced performance, as indicated by their low F1 scores. In contrast, the framework demonstrates a better balance between precision and recall for SRL phase 2, reflected in its high F1 scores. Notably, these results are consistent across both COMP and SCIE.

## 6 Discussion and Implication

### 6.1 Individual Consistency Highlights the Importance of Holistic Skill Development

The result of RQ1 indicates that on the one hand, subject-specific factors construct cognitive and motivational demands [44], thus shaping students' subject-specific SRL and online learning behaviours at the group level. On the other hand, individual students demonstrate consistent learning behaviours and SRL across different subjects. Notably, 58% of the variation in SRL is attributable to personal factors, compared to 16% attributed to subject design. This finding aligns with the aptitude aspect of SRL, suggesting it can be viewed as an internal trait that reflects students' learning tendencies across various contexts [7]. Earlier research by Vermetten et al. [37] supports this perspective, showing that while students adapt their learning strategies to meet context-specific requirements, their strategies remain strongly correlated across contexts. This consistency is likely due to individual learning habits or general skills in regulating learning [2].

The result implies that teaching students general self-regulation skills (like planning, reflection, and monitoring) and guiding them to adopt more effective learning strategies can benefit them across subjects. This supports holistic skill development rather than focusing on isolated subject-based interventions. During holistic skill development, students should be encouraged to apply their holistic learning strategies and skills to meet subject-specific needs.

## 6.2 The Context-Specific Nature of Trace-SRL and the Role of LMS Structure

The investigation of RQ2 leads to a major finding: coming up with an "one-size-fits-all" trace-SRL framework is a challenge. Specifically, the data-driven results vary across subjects, and there are discrepancies between the data-driven results and the theory-driven framework. These findings can be attributed to the context-specific nature of SRL and the structure of the LMS, as explained below.

First, SRL and learning behaviours are inherently context-specific. Since self-regulation is embedded in the complex interplay of learners, learning behaviours, and the environment [46], students continuously adapt their online learning behaviours based on subject requirements and disciplinary contexts [18, 21]. Given the variability in subject requirements and disciplines, an "one-size-fits-all" approach is unlikely to effectively capture self-regulation from students' learning behaviours.

Second, the structure of the LMS further accentuates the context-specific nature of trace-SRL. Recall that trace-SRL research does not directly measure SRL or observe students' learning behaviours; instead, it infers SRL phases from students' online actions recorded in trace data. As prior research highlights, traditional LMS platforms often lack a defined structure that supports consistent SRL measurement [10]. It means that instructors may organise learning materials and use the LMS in arbitrary ways, with no standardised structure across subjects or disciplines [21]. This inconsistency can result in trace data that is misaligned with SRL theory and varies between learning contexts (e.g., subjects), which further complicates the development of a generalised framework for measuring SRL from trace data.

Overall, the finding highlights that contextual factors, such as subject discipline, assignments, and LMS structure, are crucial considerations when interpreting trace data for SRL measurement.

## 6.3 Advancing with Automating Trace-SRL Detection or Reassessing with Multi-Modal Data? It Depends

The investigation of RQ3 shows that the existing trace-SRL method can capture the engagement SRL phase from trace data reasonably well, but it struggles to perform well on the planning and reflection phases. To explain this finding, we return to the definition of SRL.

At its core, SRL describes learners' cognitive and metacognitive processes during learning [46]. While some of these processes are recorded in trace data, others are not. For instance, students often interact with LMS platforms to complete cognitive processes, such as accessing resources during the engagement phase to acquire knowledge. In these cases, the learning actions captured in trace data provide observable footprints of students' cognitive activities. Conversely, many cognitive and metacognitive processes, such as setting personal plans, monitoring progress, or reflecting on learning, occur internally and can not be captured by trace data. This interpretation explains the results observed in RQ3. Additionally, most existing learning platforms are primarily designed to manage learning resources [10]. These platforms do not explicitly guide students through the "planning-engagement-reflection" cycle or automatically record this process. This limitation makes it even more challenging to capture the more internal planning and reflection processes compared to the externally manifested engagement process [38].

The result has important implications for guiding future research, particularly on whether to advance trace-SRL research by automating the detection process, or reassessing SRL measurement with multi-modal data. First, supported by just-in-time daily survey data, the research presents promising evidence for the validity and generalisability of using trace data to measure the engagement phase of SRL. This finding suggests that future research could focus on automating the accurate measurement of the SRL engagement phase without requiring direct input from students. Such advancements are a crucial step towards the ultimate goal of trace-SRL research: enabling large-scale, automated detection and support for students' SRL. Second, the observed imbalance in the ability to capture the planning and reflection phases is not necessarily a negative outcome. Instead, it illustrates how survey data complements trace data by capturing cognitive and metacognitive aspects that trace data alone can not. This emphasises the multi-faceted nature of SRL. Researchers interested in the entire SRL process should therefore adopt multi-modal data collection and analysis methods to provide a more holistic and comprehensive understanding.

## 7 Conclusion, Limitation and Future Works

This research advances our understanding of using trace data to measure self-regulated learning by developing and validating a novel methodology in authentic learning environments. By integrating trace data and two types of surveys, and leveraging both data-driven and theory-driven frameworks across multidisciplinary university subjects, our study addresses critical gaps in understanding the validity and generalisability of trace-based SRL measurements. Our investigation has three key findings: (RQ1) Individual students demonstrate consistent online learning behaviours and self-regulated learning patterns across subjects, even when subject design and context vary substantially; (RQ2) Despite extensive analysis, we were unable to identify universal trace data indicators for SRL phases that could be reliably applied across multidisciplinary subjects; (RQ3) Trace data shows particular effectiveness for measuring the engagement phase of SRL, providing reliable indicators of active learning behaviours. However, it proves less reliable for capturing the reflection and planning phases, suggesting these metacognitive processes may require alternative measurement approaches.

While this study provides important insights for interpreting trace data in SRL measurement and support, two main limitations warrant consideration. First, our intentional selection of subjects with significant differences in design and workload led to an imbalanced volume of data across subjects, particularly in trace data and daily SRL survey responses. While we chose to preserve the authentic nature of the data rather than employ resampling methods, this decision may introduce biases and affect the comparability of results across subjects. This limitation reflects the inherent tension between maintaining validity and achieving balanced datasets in educational research. Second, our use of just-in-time daily surveys, while enabling large-scale longitudinal data collection throughout the semester, relies on self-reported data. As noted in recent literature [6] and highlighted through peer review, self-report measures face inherent challenges regarding accuracy and potential biases. However, this limitation also underscores the importance of our multi-method approach in triangulating SRL measurements.

With the insights and implications outlined in this research, we identify promising directions for future research in SRL measurement and development. From a measurement perspective, future research can advance in two possible key areas. First, developing and validating automated systems for large-scale detection of the engagement phase using trace data, building on our finding that these behaviours are most reliably captured through digital traces. Second, designing integrated measurement approaches that strategically combine multiple data sources to achieve more comprehensive and accurate SRL assessment, particularly for capturing reflection and planning phases.

From the perspective of supporting SRL development, our findings point to two primary directions. Research could focus on developing interventions that promote holistic SRL skill development, leveraging our finding that these skills show consistency across contexts. Additionally, there is a need to create adaptive support systems that account for the varying reliability of trace-based measures across different SRL phases, ensuring that support strategies align with the strengths and limitations of available measurement approaches.

These directions offer promising pathways for advancing both the measurement and support of self-regulated learning in higher education. By focusing on automated detection of engagement patterns while developing integrated approaches for holistic assessment, future research can address the current limitations in SRL measurement while supporting scalable implementations across diverse educational contexts.

## References


[1] Lucy Barnard, William Y. Lan, Yen M. To, Valerie Osland Paton, and Shu-Ling Lai. 2009. Measuring self-regulation in online and blended learning environments. *The Internet and Higher Education* 12, 1 (Jan. 2009), 1–6. https://doi.org/10.1016/j.iheduc.2008.10.005

[2] John B. Biggs. 1987. *Student approaches to learning and studying* (1. publ ed.). Australian Council for Educational Research, Melbourne.

[3] Jaclyn Broadbent and Paula De Barba. 2023. Self-regulated learning and the student experience in online higher education. In *Research Handbook on the Student Experience in Higher Education*, Chi Baik and Ella R. Kahu (Eds.). Edward Elgar Publishing, 179–190. https://doi.org/10.4337/9781802204193.00021

[4] Magno Carlo. 2011. Validating the Academic Self-regulated Learning Scale with the Motivated Strategies for Learning Questionnaire (MSLQ) and Learning and Study Strategies Inventory (LASSI). *The International Journal of Educational and Psychological Assessment* 7(2) (May 2011). https://blog.ssrn.com/2023/01/31/ssrn-uses-doi-for-preprint-articles/

[5] Zijin Chen, Yige Song, Andrew Valentine, Paula De Barba, Shannon Rios, and Eduardo Oliveira. 2024. Do Beliefs About Intelligence Drive Engagement in Online Learning?: Exploring the Impact of Growth and Fixed Mindsets. *ASCILITE Publications* (Nov. 2024), 223–233. https://doi.org/10.14742/apubs.2024.1421

[6] Heeryung Choi, Philip H. Winne, Christopher Brooks, Warren Li, and Kerby Shedden. 2023. Logs or Self-Reports? Misalignment Between Behavioral Trace Data and Surveys When Modeling Learner Achievement Goal Orientation. In *LAK23: 13th International Learning Analytics and Knowledge Conference*. ACM, Arlington TX USA, 11–21. https://doi.org/10.1145/3576050.3576052

[7] Timothy J. Cleary and Gregory L. Callan. 2017. Assessing Self-Regulated Learning Using Microanalytic Methods. In *Handbook of Self-Regulation of Learning and Performance* (2 ed.), Dale H. Schunk and Jeffrey A. Greene (Eds.). Routledge, 338–351. https://doi.org/10.4324/9781315697048-22

[8] Paula G. de Barba, Gregor E. Kennedy, and Mary D. Ainley. 2016. The role of students' motivation and participation in predicting performance in a MOOC. *Journal of Computer Assisted Learning* 32, 3 (June 2016), 218–231. https://doi.org/10.1111/jcal.12130

[9] Paula G. de Barba, Donia Malekian, Eduardo A. Oliveira, James Bailey, Tracii Ryan, and Gregor Kennedy. 2020. The importance and meaning of session behaviour in a MOOC. *Computers & Education* 146 (March 2020), 103772. https://doi.org/10.1016/j.compedu.2019.103772

[10] Jiahui Du, Khe Foon Hew, and Lejia Liu. 2023. What can online traces tell us about students' self-regulated learning? A systematic review of online trace data analysis. *Computers & Education* 201 (Aug. 2023), 104828. https://doi.org/10.1016/j.compedu.2023.104828

[11] Yizhou Fan, Joep Van Der Graaf, Lyn Lim, Mladen Raković, Shaveen Singh, Jonathan Kilgour, Johanna Moore, Inge Molenaar, Maria Bannert, and Dragan Gašević. 2022. Towards investigating the validity of measurement of self-regulated learning based on trace data. *Metacognition and Learning* 17, 3 (Dec. 2022), 949–987. https://doi.org/10.1007/s11409-022-09291-1

[12] Jeffrey Alan Greene and Roger Azevedo. 2009. A macro-level analysis of SRL processes and their relations to the acquisition of a sophisticated mental model of a complex system. *Contemporary Educational Psychology* 34, 1 (Jan. 2009), 18–29. https://doi.org/10.1016/j.cedpsych.2008.05.006

[13] Feifei Han and Robert A. Ellis. 2023. Self-reported and digital-trace measures of computer science students' self-regulated learning in blended course designs. *Education and Information Technologies* 28, 10 (Oct. 2023), 13253–13268. https://doi.org/10.1007/s10639-023-11698-5

[14] George M. Harrison and Lisa M. Vallin. 2018. Evaluating the metacognitive awareness inventory using empirical factor-structure evidence. *Metacognition and Learning* 13, 1 (April 2018), 15–38. https://doi.org/10.1007/s11409-017-9176-z

[15] Jelena Jovanović, Mohammed Saqr, Srećko Joksimović, and Dragan Gašević. 2021. Students matter the most in learning analytics: The effects of internal and instructional conditions in predicting academic success. *Computers & Education* 172 (Oct. 2021), 104251. https://doi.org/10.1016/j.compedu.2021.104251

[16] Vitomir Kovanović, Dragan Gašević, Srećko Joksimović, Marek Hatala, and Olusola Adesope. 2015. Analytics of communities of inquiry: Effects of learning technology use on cognitive presence in asynchronous online discussions. *The Internet and Higher Education* 27 (Oct. 2015), 74–89. https://doi.org/10.1016/j.iheduc.2015.06.002

[17] Brian R. Little. 2015. The integrative challenge in personality science: Personal projects as units of analysis. *Journal of Research in Personality* 56 (June 2015), 93–101. https://doi.org/10.1016/j.jrp.2014.10.008

[18] Wannisa Matcha, Dragan Gašević, Nora'ayu Ahmad Uzir, Jelena Jovanović, Abelardo Pardo, Lisa Lim, Jorge Maldonado-Mahauad, Sheridan Gentili, Mar Pérez-Sanagustín, and Yi-Shan Tsai. 2020. Analytics of Learning Strategies: Role of Course Design and Delivery Modality. *Journal of Learning Analytics* 7, 2 (Sept. 2020), 45–71. https://doi.org/10.18608/jla.2020.72.3

[19] Sina Nazeri, Marek Hatala, and Fatemeh Salehian Kia. 2023. When to Intervene? Utilizing Two Facets of Temporality in Students' SRL Processes in a Programming Course. In *LAK23: 13th International Learning Analytics and Knowledge Conference*. ACM, Arlington TX USA, 293–302. https://doi.org/10.1145/3576050.3576095

[20] Jeremy T. D. Ng, Yiming Liu, Didier S. Y. Chui, Jack C. H. Man, and Xiao Hu. 2023. Leveraging LMS Logs to Analyze Self-Regulated Learning Behaviors in a Maker-based Course. In *LAK23: 13th International Learning Analytics and Knowledge Conference*. ACM, Arlington TX USA, 670–676. https://doi.org/10.1145/3576050.3576111

[21] Yeonjeong Park, Ji Hyun Yu, and Il-Hyun Jo. 2016. Clustering blended learning courses by online behavior data: A case study in a Korean higher education institute. *The Internet and Higher Education* 29 (April 2016), 1–11. https://doi.org/10.1016/j.iheduc.2015.11.001

[22] Paul R. Pintrich. 2000. The Role of Goal Orientation in Self-Regulated Learning. In *Handbook of Self-Regulation*. Elsevier, 451–502. https://doi.org/10.1016/B978-012109890-2/50043-3

[23] Paul R. Pintrich. 2003. A Motivational Science Perspective on the Role of Student Motivation in Learning and Teaching Contexts. *Journal of Educational Psychology* 95, 4 (Dec. 2003), 667–686. https://doi.org/10.1037/0022-0663.95.4.667

[24] Paul R. Pintrich, David A. F. Smith, Teresa Garcia, and Wilbert J. Mckeachie. 1993. Reliability and Predictive Validity of the Motivated Strategies for Learning Questionnaire (Mslq). *Educational and Psychological Measurement* 53, 3 (Sept. 1993), 801–813. https://doi.org/10.1177/0013164493053003024

[25] John Saint, Dragan Gašević, Wannisa Matcha, Nora'Ayu Ahmad Uzir, and Abelardo Pardo. 2020. Combining analytic methods to unlock sequential and temporal patterns of self-regulated learning. In *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge*. ACM, Frankfurt Germany, 402–411. https://doi.org/10.1145/3375462.3375487

[26] John Saint, Alexander Whitelock-Wainwright, Dragan Gasevic, and Abelardo Pardo. 2020. Trace-SRL: A Framework for Analysis of Microlevel Processes of Self-Regulated Learning From Trace Data. *IEEE Transactions on Learning Technologies* 13, 4 (Oct. 2020), 861–877. https://doi.org/10.1109/TLT.2020.3027496

[27] Fatemeh Salehian Kia, Marek Hatala, Ryan S. Baker, and Stephanie D. Teasley. 2021. Measuring Students' Self-Regulatory Phases in LMS with Behavior and Real-Time Self Report. In *LAK21: 11th International Learning Analytics and Knowledge Conference*. ACM, Irvine CA USA, 259–268. https://doi.org/10.1145/3448139.

3448164

[28] Mohammed Saqr, Sonsoles López-Pernas, Jelena Jovanović, and Dragan Gašević. 2023. Intense, turbulent, or wallowing in the mire: A longitudinal study of cross-course online tactics, strategies, and trajectories. *The Internet and Higher Education* 57 (April 2023), 100902. https://doi.org/10.1016/j.iheduc.2022.100902

[29] Gregory Schraw and Rayne Sperling Dennison. 1994. Assessing Metacognitive Awareness. *Contemporary Educational Psychology* 19, 4 (Oct. 1994), 460–475. https://doi.org/10.1006/ceps.1994.1033

[30] Skipper Seabold and Josef Perktold. 2010. Statsmodels: Econometric and Statistical Modeling with Python. Scipy, 92–96.

[31] Melody Siadaty, Dragan Gašević, and Marek Hatala. 2016. Measuring the impact of technological scaffolding interventions on micro-level processes of self-regulated workplace learning. *Computers in Human Behavior* 59 (June 2016), 469–482. https://doi.org/10.1016/j.chb.2016.02.025

[32] Melody Siadaty, Dragan Gašević, and Marek Hatala. 2016. Trace-based Micro-analytic Measurement of Self-Regulated Learning Processes. *Journal of Learning Analytics* 3, 1 (April 2016). https://doi.org/10.18608/jla.2016.31.11

[33] Marina Sokolova and Guy Lapalme. 2009. A systematic analysis of performance measures for classification tasks. *Information Processing & Management* 45, 4 (July 2009), 427–437. https://doi.org/10.1016/j.ipm.2009.03.002

[34] Yige Song, Eduardo Oliveira, Michael Kirley, and Pauline Thompson. 2024. A Case Study on University Student Online Learning Patterns Across Multidisciplinary Subjects. In *Proceedings of the 14th Learning Analytics and Knowledge Conference*. ACM, Kyoto Japan, 936–942. https://doi.org/10.1145/3636555.3636939

[35] Ramakrishnan Srikant and Rakesh Agrawal. 1996. Mining sequential patterns: Generalizations and performance improvements. In *Advances in Database Technology — EDBT '96*, Gerhard Goos, Juris Hartmanis, Jan Van Leeuwen, Peter Apers, Mokrane Bouzeghoub, and Georges Gardarin (Eds.). Vol. 1057. Springer Berlin Heidelberg, Berlin, Heidelberg, 1–17. https://doi.org/10.1007/BFb0014140 Series Title: Lecture Notes in Computer Science.

[36] Allen M. Tough. 1971. *The adult's learning projects: a fresh approach to theory and practice in adult learning*. Number no. 1 in Research in education series. Ontario Institute for Studies in Education, Toronto.

[37] Yvonne J. Vermetten, Hans G. Lodewijks, and Jan D. Vermunt. 1999. Consistency and variability of learning strategies in different university courses. *Higher Education* 37, 1 (1999), 1–21. https://doi.org/10.1023/A:1003573727713

[38] Olga Viberg, Mohammad Khalil, and Martine Baars. 2020. Self-regulated learning and learning analytics in online learning environments: a review of empirical research. In *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge*. ACM, Frankfurt Germany, 524–533. https://doi.org/10.1145/3375462.3375483

[39] Alexander Whitelock-Wainwright, Yi-Shan Tsai, Kayley Lyons, Svetlana Kaliff, Mike Bryant, Kris Ryan, and Dragan Gašević. 2020. Disciplinary differences in blended learning design: a network analytic study. In *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge*. ACM, Frankfurt Germany, 579–588. https://doi.org/10.1145/3375462.3375488

[40] Frank Wilcoxon. 1992. Individual Comparisons by Ranking Methods. In *Breakthroughs in Statistics*, Samuel Kotz and Norman L. Johnson (Eds.). Springer New York, New York, NY, 196–202. https://doi.org/10.1007/978-1-4612-4380-9_16 Series Title: Springer Series in Statistics.

[41] Philip Winne and Allyson Hadwin. 1998. Studying as self-regulated learning. In *Metacognition in Educational Theory and Practice.*, Douglas Hacker, John Dunlosky, and Arthur Graesser (Eds.). Taylor And Francis, 277–304.

[42] Philip H. Winne. 2020. Construct and consequential validity for learning analytics based on trace data. *Computers in Human Behavior* 112 (Nov. 2020), 106457. https://doi.org/10.1016/j.chb.2020.106457

[43] Philip H. Winne and Nancy E. Perry. 1999. Measuring Self-Regulated Learning. In *Handbook of Self-Regulation*, Monique Boekaerts, Moshe Zeidner, and Pintrich Paul (Eds.). Elsevier Science & Technology.

[44] Christopher Wolters and Paul Pintrich. 1998. Contextual differences in student motivation and self-regulated learning in mathematics, English, and social studies classrooms. *Instructional Science* 26, 1/2 (1998), 27–47. https://www.jstor.org/stable/23371263

[45] Dan Ye and Svoboda Pennisi. 2022. Using trace data to enhance Students' self-regulation: A learning analytics perspective. *The Internet and Higher Education* 54 (June 2022), 100855. https://doi.org/10.1016/j.iheduc.2022.100855

[46] Barry J. Zimmerman. 2000. Attaining Self-Regulation. In *Handbook of Self-Regulation*. Elsevier, 13–39. https://doi.org/10.1016/B978-012109890-2/50031-7

[47] Ayça Çebi, Rafael D. Araújo, and Peter Brusilovsky. 2023. Do individual characteristics affect online learning behaviors? An analysis of learners sequential patterns. *Journal of Research on Technology in Education* 55, 4 (July 2023), 663–683. https://doi.org/10.1080/15391523.2022.2027301