A Trace-Based Generalized Multimodal SRL Framework for Reading-Writing Tasks

Debarshi Nath
IIT Bombay-Monash Research
Academy
debarshi.nath@iitb.ac.in

Dragan Gasevic
Department of
Human-Centred Computing
Monash University

Ramkumar Rajendran
Interdisciplinary Programme in
Educational Technology
Indian Institute of Technology
Bombay

ABSTRACT
Reading-Writing has a ubiquitous presence in almost all kinds of learning. While measurement frameworks for self-regulated learning exist, they are often very contextual and do not guarantee generalizability over more than a specific task. This doctoral project primarily aims to investigate the applicability of a common SRL measurement framework over a range of reading-writing tasks. The research also aims to investigate whether integrating log data, peripheral data like mouse clicks and keystrokes and eyetracking data reveal more information and improve the measurement of SRL.

Keywords
Self-Regulated Learning, Multimodal Analytics, Process Mining, Pattern Mining, Predictive Modelling, Temporal Analytics

1. INTRODUCTION
Writing is an essential part of thinking and learning, whether it be in a school context, in higher education or in a professional setting. Writing tasks is also a critical tool for intellectual and social development [8]. Reading, comprehending, and writing are extremely ubiquitous requirements for all kinds of learning setups. For this reason, developing self-regulation of learners in writing tasks has gained prominence in educational research for a long time [8]. Self-regulation in writing tasks has consequently been explored greatly over the years [8, 12, 1]. But the inception of digital learning environments has opened up new possibilities for understanding learners' mental processes and supporting proper learning strategies through the collection of trace data. Combining trace logs and other forms of multimodal data can reveal more information about learners' latent mental processes and can improve the current state of research [19].

There have been trace-based studies focusing on writing tasks [9, 7, 16, 11, 2]. However, a large number of these stud-

ies are very contextual; they are conducted in their ad-hoc learning environments and for their own specific reading-writing task. This statement can actually be made for most SRL-based studies, and rightly so because self-regulated learning is extremely contextual [18]. Most learning environments are so specific that they do not allow generalizations across multiple environments [15]. Researchers do adopt measuring protocols from other studies, but that again raises questions about the validity and reliability of such measurements as such measurements were designed for a very specific learning context.

A learner's adoption of strategies can also depend on the type of reading-writing task. There are three major kinds of reading comprehension- *literal*, *inferential* and *evaluative* [17]. There are four types of writing styles- *persuasive*, *narrative*, *expository* and *descriptive* [10]. The goal of the assignment can determine the style or combination of styles that a reader and writer may adopt. Despite these differences in reading and writing styles, writing tasks do have their commonalities across tasks- most involve reading, comprehending, and writing. With this view, we put up our case that creating a trace-based measurement protocol that can be used across multiple writing tasks can ease the pain of researchers who often have to conduct tedious controlled studies and manual coding to ascertain the validity of their trace data-based studies in their own context. Developing such a protocol can also help learning systems designers create universal learning environments which can support learners' self-regulation. Hence, we explore the possibility of generic trace-based measurement protocol that can measure SRL across multiple reading-writing tasks, and at the same time is able to identify the differences in self-regulation in each of these tasks.

In this doctoral project, we aim to investigate whether a trace-based measuring protocol designed, developed, and tested for one writing task can be used across multiple writing tasks. We also explore whether integrating multimodal data like eye-tracking with the existing log channel can improve the modeling of the learners.

2. RESEARCH QUESTIONS
The following are the research questions that we aim to answer in this doctoral project:

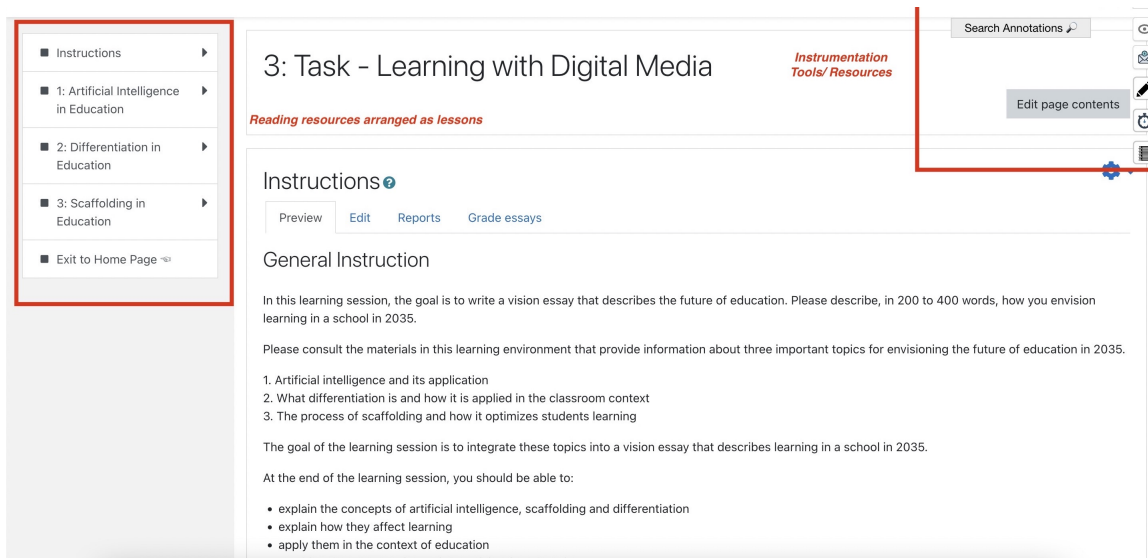1. **RQ1:** How do students' SRL strategies change when

Figure 1: The Learning Environment

they engage in various reading-writing tasks with different goals?

2. **RQ2:** Do data from multiple sensors (like logs + eye-gaze) improve the detection of SRL strategies in learners, as compared to a single channel (i.e., logs)?

3. **RQ3:** Do prediction models trained on task-independent reading-writing multimodal data (data combined from multiple tasks) perform equivalently as for that in a specific reading-writing task?

3. THEORETICAL FRAMEWORK

Self-Regulated Learning (SRL) is a theoretical umbrella that encompasses cognitive, metacognitive, behavioural, and affective aspects of learning [14]. While different theoretical models have large commonalities between them as they try to capture alternate views of the same process, there exist subtle differences based on the aspects on which their central focus lies [14]. A large majority of these models view SRL as a cyclic process comprising three phases- Preparatory, Performance and Reflection. In the theoretical framework that we use, theoretically-grounded patterns of atomic user actions are mapped to higher-level SRL processes. Thus, different SRL processes have been operationalised using patterns of meaningful learner actions. A detailed description of the theoretical framework along with the exact list of patterns used to identify SRL processes is present in [5].

4. METHODS

Over the duration of the doctoral project, we aim to collect data from two (or more) reading-writing tasks, both with different content and different overall goals, and we aim to investigate them with a single trace-based SRL measurement framework. We will investigate whether the same framework is sufficient to capture the differences between the tasks, and what are the similarities as well. In the second year of PhD, we have focussed on collecting the data for one reading-writing task (specifically the one explained in section 4.2).

A schematic diagram of our study design is represented in Fig 2.

4.1 Study Setup

The lab study is being conducted at the Department of Educational Technology, IIT Bombay. English is not the first language of the participants in the research study, but they have studied or are currently enrolled in institutions where English is the primary language of instruction. All the participants are college-going students from diverse streams or disciplines. The participants are a mix of undergraduates, post-graduates or PhD students.

As part of the data, we are collecting their software logs, the eye-tracking data of the students, their facial recordings and screen recordings. The eye-tracking data is being collected using Tobii Pro Nano screen-based eyetracker sampled at 60Hz. The data is exported using Tobii Pro SDK on Python which offers an open-source solution to export the raw data collected using Tobii eyetrackers.

4.2 Procedure

The study uses a pre-post test design which comprises of a 90 min reading-writing task the learners are required to go through a set of reading materials pertaining to three topics and compose a piece of writing. The three topics are: (1) artificial intelligence, (2) differentiation in the classroom and (3) scaffolding of learning. The goal of the task is to compose an essay that gives an overview of the state of education in the year 2035 within 400 words. The task has been designed in a way that prompts the learner to use SRL skills and tools like highlighter, notetaker in the learning environment.

4.3 Learning Environment

The task had been created in a Moodle-based learning environment, as shown in fig 1. The learning environment consists of a catalogue and navigation area which contains the list of reading materials and a way to navigate between
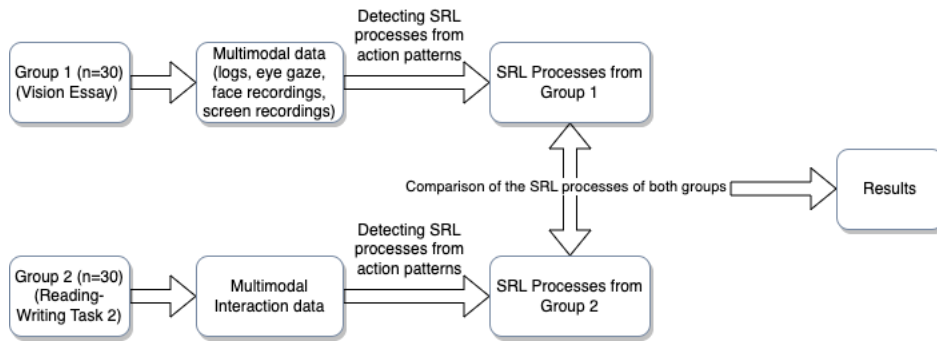
Figure 2: The Study Design

them, and also to the general instructions and rubric of the task. There is a reading area in the centre which displays the contents of the selected reading material. The environment is integrated with tools like annotation, planner, timer tools. There is a writing window that can be opened and closed at any time for writing the essay. The planner tool helped the learners to plan how much time they are going to spend on each part of the task, and the timer tool displayed the time left for the task. The annotation tool, based on the open web annotation tool hypothes.is, allowed the learners to highlight, annotate and take notes, and they can search for highlights, tags and notes they created earlier as well. Within this learning environment, we collected learners' trace data that includes: 1) navigational log, which stored the time-stamps for all page visits; 2) mouse trace data, which stored mouse clicks on pages and mouse scrolls; 3) keyboard strokes.

4.4 Trace-based Measurement Protocol for Log Data

For converting the raw logs into theoretical SRL processes, we use a trace parser. The parser first converts the raw logs into meaningful learning actions like RELEVANT_READING, PLANNER, GENERAL_INSTRUCTION. These set of learning actions give rise to our *action library*. Specific theoretical patterns of these actions are then mapped to higher-level SRL processes. The entire list of these SRL processes then gives our *process library*. The entire process of parsing has been detailed in [5]. The theoretical SRL processes that we obtain in our learning task are also listed in Table 1. Each of these processes is coded by experts from 2-action or 3-action sequences. Our learning system also provides us with the duration spent on each of these patterns of actions (and hence the SRL processes). We can add the duration spent on each of these SRL processes separately during the 90 min learning period and also count their occurrences which gives us the metrics such as those represented in Table 1.

4.5 Data Processing and Feature Extraction from eye-gaze data

For cleaning and processing the raw eye gaze data, we will be following the steps outlined for the Tobii I-VT Fixation Filter [13]. The steps involve gap fill-in interpolation, eye selection, and noise reduction among other steps.

We will extract two main features- fixations and saccades

and their derivatives from the eye-gaze data. For this purpose, we aim to use PyTrack [6], which is an open-source Python-based solution for analyzing eye-gaze data.

5. RESULTS

Table 1 represents the distribution of SRL processes within each category of SRL processes/subprocesses for 9 learners. The distribution is comparable to that presented in [5], where *Elaboration/Organisation*, *First Reading* and *Monitoring* emerged as the most prevalent SRL processes in the learners for the essay-writing task. A point to note is that the sample presented in [5] is from a population of learners whose first language is Dutch over 45 min of reading-writing, while the sample presented in this paper is from a population whose native language is not English over a period of 90 min.

Table 1: Distribution of SRL processes in the participants

Main Categories	Subcategories	Count	Duration (%)
Metacognition	Orientation	79	21.625
	Planning	10	0.375
	Monitoring	186	3.468
	Evaluation	17	0.574
Low Cognition	First Reading	267	36.974
	Re-reading	156	6.244
High Cognition	Elaboration/Organisation	349	30.739

6. FUTURE WORK

As introduced earlier, the objective of our task can determine the style or combination of styles that a reader and writer may adopt. To compare two examples, the vision essay in our learning task requires a learner to read and reflect on three readings- Artificial Intelligence in Education, Differentiation in Education and Scaffolding in Education and write a vision of education in 2035. The learner is expected to stay connected to the readings, but is also expected to combine them, go beyond what is there in the readings and imagine innovative scenarios in future where the information from these topics could be relevant. To contrast with this task, an argumentative task is a common form of academic writing where a learner is supposed to take a stance and make a for/against argument for a situation and back it up with evidences from the readings [4, 10]. Compared to the earlier vision essay, this task is rather restricted and the learner has to interpret the information, identify the relevant pieces of information from the readings, strictly adhere to facts and avoid misdirections in the text (if any) and put

up a case for the argument. We hypothesize that such contrasting tasks can impact the self-regulatory behaviour of the learners, even while going through the same content.

For the research questions that we aim to address, we will continue our data collection. Once the data is collected for the current reading-writing task, we will change the task in terms of its goal and content, and collect data (most likely from a classroom course). This will allow us to have a substantial amount of data to answer our research questions in the ways described in brief below.

6.1 RQ1

To address RQ1, we aim to investigate the differences in the SRL strategies of learners depending on different reading-writing tasks using the following methods-

(a) Comparing the distribution of counts and duration spent by the learners in the SRL process categories for each of the tasks.

(b) Comparing aggregate process models of the learners for each of the tasks.

(c) Sequential Pattern Mining to reveal dominant action patterns in each of the task.

6.2 RQ2

To answer RQ2, we aim to combine log data (logs + mouse and keyboard interactions) and eye-gaze information. We plan to investigate whether sufficient attention was given each page of the content during each of the SRL process, and filter out the pages based on whether adequate eye-gaze were pointed to them.

6.3 RQ3

RQ3 involves a problem of prediction, which involves the prediction of the SRL process of the learner based on the logs and eye-gaze data of the students. The problem can be taken up either as a classification problem of predicting the SRL process from the data or predicting the next SRL process of the learner based on their current SRL state. We will train and test independently for each task, and compare the performance of our model when trained and tested for all tasks combined.

Prior to combining them, we will abstract features from the channels (features like count of mouse clicks, scrolls and count of fixations and saccades in AOIs from eye-gaze data).

7. CONCLUSION

The doctoral project focuses on investigating whether a single SRL measurement framework can be generalized for multiple reading-writing tasks. The outcome will provide evidence for the applicability of SRL measurement frameworks for multiple tasks and hopefully, it will prompt more research toward building generic SRL models at least for a certain set of tasks that have commonalities between them. The SRL measurement frameworks are at this point very contextual and restricted in nature.

The multimodal aspect of the project also aims to investigate whether additional data channels can reveal more information about the nature of self-regulation in learners. We will explore whether the eye-gaze channel can inform the log data channel better, or vice versa.

We have so far collected data for 16 participants, and have presented a summary of the results of 9 participants after consideration of the quality of the data. Going ahead we aim to collect more data from participants engaged in this task, and also collect data from learners in newer reading-writing tasks with different content. Then we will be ready to answer our RQs in ways described in the last section.

The approach is not without its limitations. The multimodal aspect of the project (especially RQ2) is very investigative in nature, and the methods will depend on the researcher. How to fuse the data channels, which exact features to select, and how to ensure its explainability is yet to be decided and are challenges on their own. The data that we have collected so far has been collected in a controlled lab environment, and real-world data might not be as clean as ours. We aim to collect data for our further reading-writing tasks from a real-world classroom, and ensuring the quality of the data and choosing appropriate technological solutions for multimodal data collection are other challenges. We also need to ensure that the content for our further new reading-writing tasks is comparable in terms of their complexity to the current reading-writing task.

The applicability of the research can be diverse and can go beyond just the measurement of SRL in reading-writing tasks. Although our major focus is on correct and valid measurement, appropriate measurement can be used for scaffolding learners' self-regulation which is an area that has gained momentum in recent years. Prediction models that we aim to investigate can help in scaffolding further by telling researchers which SRL processes the learner is going to enter next at any instant of time, in realtime. This information can be used to personalize the scaffolding process. Although at this point we only aim to work with logs and eyetracking data channels, more data channels like physiological sensors (skin conductance, heart rate) and facial expressions could be integrated to reveal more information about SRL [3].

8. ADVICE SOUGHT

The answer to the following questions will greatly help in ensuring that my research progresses on the correct path:

1. What are the best methods for comparing event-based processes? (other than sequential pattern mining, process models and statistical differences of event occurrences)?

2. The events in an activity such as the learner actions in our task occur at uneven intervals. Is there a possibility of using classic temporal prediction models in such cases?

3. How to combine data from multimodal channels while still keeping the temporal nature of the process intact, especially when the sampling rates of the data channels are uneven and one data channel (log data) is not even periodic in nature?

4. Are webcam eyegaze detection comparable to screen-based eye trackers when detecting fixations within an AOI?

9. REFERENCES

[1] S. Abadikhah, Z. Aliyan, and S. H. Talebi. Efl students' attitudes towards self-regulated learning strategies in academic writing. *Issues in Educational Research*, 28:1–17, 01 2018.

[2] M. Bernacki, J. Byrnes, and J. Cromley. The effects of achievement goals and self-regulated learning behaviors on reading comprehension in technology-enhanced learning environments. *Contemporary Educational Psychology*, 37:148–161, 04 2012.

[3] E. Cloude, R. Azevedo, P. Winne, G. Biswas, and E. Jang. System design for using multimodal trace data in modeling self-regulated learning. *Frontiers in Education*, 7:928632, 08 2022.

[4] ETS. Gre® general test: Analytical writing question types.

[5] Y. Fan, J. van der Graaf, L. Lim, M. Raković, S. Singh, J. Kilgour, J. Moore, I. Molenaar, M. Bannert, and D. Gašević. Towards investigating the validity of measurement of self-regulated learning based on trace data. *Metacognition and Learning*, May 2022.

[6] U. Ghose, A. S., W. Boyce, H. Xu, and E. Chng. Pytrack: An end-to-end analysis toolkit for eye tracking. *Behavior Research Methods*, 52, 03 2020.

[7] A. Hadwin, J. Nesbit, D. Jamieson-Noel, J. Code, and P. Winne. Examining trace data to explore self-regulated learning. metacognition & learning, 2, 107-124. *Metacognition and Learning*, 2:107–124, 12 2007.

[8] L. A. Hammann. Self-regulation in academic writing tasks. 2005.

[9] D. Jamieson-Noel and P. Winne. Comparing self-reports to traces of studying behavior as representations of students' studying and achievement. *Zeitschrift Fur Padagogische Psychologie - Z PADAGOG PSYCHOL*, 17:159–171, 11 2003.

[10] R. JEFFREY. OPEN OREGON EDUCATIONAL, 2018.

[11] P. K., S. T., and K. R. Development of computer-based learning system for learning behavior analytics. *Indonesian Journal of Electrical Engineering and Computer Science*, 25(1):460 – 473, 2022. Cited by: 0; All Open Access, Gold Open Access.

[12] R. Nitta and K. Baba. *Self-regulation in the evolution of the ideal L2 self: A complex dynamic systems approach to the L2 motivational self system*, pages 367–396. 01 2015.

[13] A. Olsen. The tobii I-VT fixation filter- algorithm description, Mar 2012.

[14] E. Panadero. A review of self-regulated learning: Six models and four directions for research. *Frontiers in Psychology*, 8, 2017.

[15] J. Saint, A. Whitelock-Wainwright, D. Gašević, and A. Pardo. Trace-srl: A framework for analysis of microlevel processes of self-regulated learning from trace data. *IEEE Transactions on Learning Technologies*, 13(4):861–877, 2020.

[16] N. Srivastava, Y. Fan, M. Rakovic, S. Singh, J. Jovanovic, J. van der Graaf, L. Lim, S. Surendrannair, J. Kilgour, I. Molenaar, M. Bannert, J. Moore, and D. Gasevic. Effects of internal and external conditions on strategies of self-regulated learning: A learning analytics study. In *LAK22: 12th International Learning Analytics and Knowledge Conference*, LAK22, page 392–403, New York, NY, USA, 2022. Association for Computing Machinery.

[17] D. o. E. Victoria State Government. Comprehension.

[18] P. Winne. Improving measurements of self-regulated learning. *Educational Psychologist*, 45:267–276, 10 2010.

[19] P. Winne. Construct and consequential validity for learning analytics based on trace data. *Computers in Human Behavior*, 112:106457, 06 2020.