Analytics for the Measurement of Process Dimensions of Self-Regulated Learning and Feedback Impact

John Saint[1,2][0000-0002-8338-9300]

[1] University of Edinburgh, United Kingdom
[2] Regent's University London, United Kingdom

Abstract. Blended learning environments have the potential to provide educators with valuable insights into learner behaviours and strategies. Capturing and analysing learner data, using traditional frequency-based statistical methods, is a challenge if the objective is to understand self-regulated learning (SRL) as a dynamic process. Current research on SRL has recognised the potential of data science methods for analysis of temporal processes. The current studies, however, rely on think-aloud data and other self-report measures, which are costly and intrusive. The use of trace data is a promising alternative, though raw trace data, as captured by contemporary learning environments, is insufficient to represent self-regulation processes (aka micro-level SRL processes) as theorised in well-known models. To address these challenges, this research aims to 1) improve the measurement of SRL by deriving micro-level processes from trace data; 2) analyse these micro-level processes for temporal associations; 3) explore how such temporal associates between micro-level processes are correlated with learning strategies; and 4) assess the impact of formative data-driven feedback on these SRL processes. We undertook a preliminary study and found that certain temporal activity traits relate to performance in the summative assessments attached to the course, mediated by strategy type. In addition, more strategically minded activity, embodying learner self-regulation, generally proves to be more successful than less disciplined reactive behaviours.

Keywords: Learning Analytics, Self-Regulated Learning, Micro-Level Processes, Student Feedback.

1 Motivation

Self-regulation is a key skill for strategically mature students as it informs how effectively they process feedback (both internal and external) and act upon it [1]. In addition, administering developmental feedback to students has a significant effect on their learning journey and academic performance [2]. From a metacognitive viewpoint, exponents of self-regulated learning (SRL) are able to succeed by assessing, planning, assimilating, organising, and self-evaluating in an ongoing cycle [3]. Therefore, significant benefits can be realised in identifying, articulating, and optimising patterns of SRL. Much of the research around measuring SRL, however, is based not on authentic process data, but on variants of self-report data capture e.g. [4] [5].

Data generated in blended learning environments, specifically those from learning management systems (LMS), provide opportunities for researchers to unlock insights into learner behaviours and strategies. The use of LMS trace data is a promising alternative to self-report data, as it eliminates potential issues of data objectivity and reliability. However, raw trace data cannot, in and of itself, represent self-regulation processes as theorised in well-known models. Therefore, the derivation of SRL macroprocesses and associated microprocess, as demonstrated by Siadaty et al. [6], provides a strong methodological platform on which the current study can build.

2 Research Area

2.1 Modelling SRL

Winne [7] identifies three key aspects of SRL: 1) Cognitive Tactic; 2) Cognitive Strategy; 3) Metacognition. This articulates a learner's management of their own cognitive tactics, and the development of an overarching knowledge management strategy, encompassing self-awareness. Zimmerman's model of SRL also provides a strong and conceptually interpretable model: Self-regulation is presented as a cycle of forethought (planning), performance (of learning event), and self-reflection [8].

In the context of Learning Analytics (LA), Winne advises underpinning SRL research with a proven SRL model. and provides a framework for mapping trace data events to 'inferences' and categorising them to phases of his SRL model [9]. There are studies that harvest pure trace data to unlock insights into cognitive tactics and learning strategies e.g. [10] [11] [12]. They stop short, however, of truly articulating SRL. This study aims to use pure LMS-generated trace data as its source, eliminating the empirical shortcomings of self-report data, yet retaining the vital characteristics of SRL.

2.2 SRL Microlevel Process Analysis

Microlevel process analysis is one of the responses to the challenges of capturing and articulating SRL. Cleary consolidates various characteristics of the field into a broad definition: "…a highly specific or fine-grained form of measurement that targets behaviours or processes as they occur in real time across authentic contexts…" [13, p. 330]. Greene and Azevedo [5] present a mapping of micro-level processes e.g., goal-setting or content evaluation, to overarching macro-level processes e.g., planning or monitoring. In this context, the macro-level processes represent elements of the chosen SRL model.

Siadaty et al. [6] build on this substantially by developing a hybrid self-report/trace-based protocol of SRL microanalysis. They posit an SRL model which positions 1) Planning, 2) Engagement, and 3) Evaluation & Reflection as its macro-level processes/SRL phases. Micro-level processes, such as Task Analysis or Working on Task, are categorised to their corresponding macro-level processes. Siadaty et al.'s method was empirically validated in two studies on the self-regulatory patterns of knowledge workers in technology-enhanced environments [14] [15]. These studies provide a critical empirical SRL bedrock. They do not, however, explore inter/intra-strategy temporal differences. Additionally, although the impact of various scaffolding interventions is assessed, it does not represent a true study of feedback as a mediator of SRL. This study seeks to address this empirical gap.

2.3 Event-based Process Analysis

Process Mining (PM) is an analytical discipline that straddles data mining, machine learning, and business process modelling. Being data-driven but process-centric, we can view it as a missing link between data science and process science [9]. In many PM studies, the processes are of a tactic-level granularity. This study aims to harness PM for micro and macro-level analysis. PM provides a vital temporal dimension that is not afforded by traditional statistical methods e.g., [10] [12]. Some studies recognise time as a dimension, but this is restricted to measurement of time on task, and not a reflection of true inter-process temporal dynamics e.g., [16]. The current study seeks to unlock insights into the temporal sequence of study activities as exhibited by exponents of SRL.

Bannert et al. [4] use process mining techniques to analyse think-aloud data logged from a student-group's navigation through an LMS. Their aim is to provide a comparison of process models of high and low performing students. Lust et al. [10] use clustering to identify user-profiles through learner behaviours, identifying profiles through frequency of activity. Kovanovic et al. [12], Jovanovic et al. [17], and Fincham et al. [18] all demonstrate sophisticated deployments of learner clustering around user-profiles and strategic learning sequences. The resultant group comparisons are insightful but leave a clear empirical gap for intra and inter-strategy articulation in the context of SRL micro-analysis. The current study aims to provide this specific comparative analysis.

Providing quality feedback in HE is inherently challenging. These challenges have intensified with the increasing massification of education. Two significant studies, both using the same high-volume LMS trace data, provide valuable insights into the impact of customised automated feedback. Pardo et al. [19] demonstrate a positive association between feedback messaging and both student satisfaction and assessment performance. Fincham et al. [18] detected tactical transition and strategic improvement (linked to assessment performance) as the result of feedback interventions. The current study aims to build on this research to assess the impact of feedback on SRL patterns in learners.

2.4 Research Questions

RQ1. To what extent can micro-level SRL processes be derived from trace data collected by conventional virtual learning environments?

RQ2. To what extent can temporal associations between micro-level SRL processes be derived through the analysis of trace data?

RQ3. What are the differences in SRL micro-level processes exhibited by students following different learning strategies?

RQ4. What is the impact of conditionally administered, analytics-based formative feedback on the micro-level SRL processes of students who follow different learning strategies?

3 Methodology

The trace data for this study come from two sources: The first were collected from an LMS attached to a computing course at an Australian university. The datasets provide LMS trace data from four cohorts of a course, spanning 2014 to 2017 [20]. The course was based on a flipped classroom pedagogy and the data relate to students' engagement with the online activities as preparation for the face-to-face learning sessions. Each time a student engaged with an element of the LMS, a learning event record was generated. These events, which are collectively called trace data, provide the source for our analyses. The second will be generated from a course to run at a London university in 2019.

The starting point of PM is a dataset in the form of an event log. The required elements to run a PM algorithm are: *Case*, a process instance; *Activity*, a well-defined step in a broader process; *Timestamp*, providing the temporality that is key to this study. Each LMS event record contains a *student ID* number (which serves as our PM case), a completed *study action* (which serves as our PM activity), and a *timestamp.*

To identify micro-level SRL processes (RQ1), we will extract trace data and utilise the mapping method outlined in the Siadaty (and associated) studies i.e. trace event → micro-level process → macro-level process/SRL construct. To address RQ2, we will build on the PM techniques explored in our preliminary study (see section 4) and extract temporal relationships from the SRL microprocesses. To address RQ3, we will cluster students in strategy groups, using the methods employed by Bannert et al [4] and Fincham et al [18]. We will perform pair-wise comparative analyses, using appropriate PM algorithms, to articulate the differences in learner strategies from a temporal micro-level perspective. Finally, will consolidate these methods to measure the impact of feedback interventions on learner behaviours, mediated by strategy type (RQ4).

4 Preliminary Results

We undertook a preliminary study, using the 2014 cohort LMS data; this was submitted to the EC-TEL 2018 conference as a full research paper [21]. We employed the R package pMineR [22] to train process models using first order Markov chains. For this study, we did not undertake micro-level analysis, but employed a coarser, tactic-level granularity in our definition of PM activity. The focus of the study is the analysis of tactical cognitive processes in a temporal/stochastic context, and how it informs learning strategy and performance. We found that certain temporal activity traits relate to performance in the summative assessments attached to the course, mediated by strategy type. In addition, more strategically minded activity, embodying learner self-regulation, generally proves to be more successful than less disciplined reactive behaviours.

5 Future Agenda & Publication Plan

It is anticipated that this doctoral submission will be a linked collection of published journal articles. These publications will be interspersed by conference paper submissions to future Learning Analytics and Knowledge (LAK) conferences. Building on the preliminary study, we hope to expand and formalise a process mining/microprocess methodology and replicate it across the remaining LMS cohorts (2015-2017). Finally, it is hoped that data harvested from the London study will provide a means of testing the replicability and scalability of a consolidated LA methodology.

References

1. Winne, P.H., Hadwin, A.F.: Studying as Self-Regulated Learning. In: Hacker, D.J., Dunlosky, J., and Graesser, A.C. (eds.) Metacognition in educational theory and practice. pp. 277–304. Lawrence Erlbaum Associates, Mahwah, New Jersey (1998)
2. Hattie, J., Timperley, H.: The Power of Feedback. Rev. Educ. Res. 77, 81 (2007)
3. Zimmerman, B.J.: Self-Regulated Learning and Academic Achievement: An Overview. Educ. Psychol. 25, 3–17 (1990). doi:10.1207/s15326985ep2501
4. Bannert, M., Reimann, P., Sonnenberg, C.: Process mining techniques for analysing patterns and strategies in students' self-regulated learning. Metacognition Learn. 9, 161–185 (2014). doi:10.1007/s11409-013-9107-6
5. Greene, J.A., Azevedo, R.: A macro-level analysis of SRL processes and their relations to the acquisition of a sophisticated mental model of a complex system. Contemp. Educ. Psychol. 34, 18–29 (2009). doi:https://doi.org/10.1016/j.cedpsych.2008.05.006
6. Siadaty, M., Gašević, D., Hatala, M.: Trace-Based Microanalytic Measurement of Self-Regulated Learning Processes Moray House School of Education and School of Informatics. J. Learn. Anal. 3, 183–214 (2016)
7. Winne, P.H.: A metacognitive view of individual differences in self-regulated learning. Learn. Individ. Differ. 8, 327–353 (1996). doi:10.1016/S1041-6080(96)90022-9
8. Zimmerman, B.J.: Attaining Self-Regulation: A Social Cognitive Perspective. In: Boekaerts, M., Pintrich, P.R., and Zeidner, M. (eds.) Handbook of Self-Regulation. pp. 13–39. Academic Press, San Diego (2000)
9. Winne, P.H.: Learning Analytics for Self-Regulated Learning. In: Lang, C., Siemens, G., Wise, A.F., and Gaševic, D. (eds.) The Handbook of Learning Analytics. pp. 241–249. Society for Learning Analytics Research (SoLAR), Alberta, Canada (2017)
10. Lust, G., Vandewaetere, M., Ceulemans, E., Elen, J., Clarebout, G.: Tool-use in a blended undergraduate course: In Search of user profiles. Comput. Educ. 57, 2135–2144 (2011). doi:https://doi.org/10.1016/j.compedu.2011.05.010
11. Lust, G., Elen, J., Clarebout, G.: Regulation of tool-use within a blended

course: Student differences and performance effects. Comput. Educ. 60, 385–395 (2013). doi:https://doi.org/10.1016/j.compedu.2012.09.001

12. Kovanović, V., Gašević, D., Joksimović, S., Hatala, M., Adesope, O.: Analytics of communities of inquiry: Effects of learning technology use on cognitive presence in asynchronous online discussions. Internet High. Educ. 27, 74–89 (2015). doi:10.1016/j.iheduc.2015.06.002
13. Cleary, T.J.: Emergence of Self-Regulated Learning Microanalysis. In: Zimmerman, B.J. and Schunk, D.H. (eds.) Handbook of Self-Regulation of Learning and Performance. pp. 329–345. Routledge, New York, NY, USA (2011)
14. Siadaty, M., Gašević, D., Hatala, M.: Measuring the impact of technological scaffolding interventions on micro-level processes of self-regulated workplace learning. Comput. Human Behav. 59, 469–482 (2016). doi:10.1016/j.chb.2016.02.025
15. Siadaty, M., Gašević, D., Hatala, M.: Associations between technological scaffolding and micro-level processes of self-regulated learning: A workplace study. Comput. Human Behav. 55, 1007–1019 (2016). doi:10.1016/j.chb.2015.10.035
16. Kovanović, V., Gašević, D., Dawson, S., Joksimović, S., Baker, R.S., Hatala, M.: Does time-on-task estimation matter? Implications for the validity of learning analytics findings. J. Learn. Anal. 2, 81–110 (2015). doi:10.18608/jla.2015.23.6
17. Jovanović, J., Gašević, D., Dawson, S., Pardo, A., Mirriahi, N.: Learning analytics to unveil learning strategies in a flipped classroom. Internet High. Educ. 33, 74–85 (2017). doi:10.1016/j.iheduc.2017.02.001
18. Fincham, O.E., Gasevic, D. V., Jovanovic, J.M., Pardo, A.: From Study Tactics to Learning Strategies: An Analytical Method for Extracting Interpretable Representations. IEEE Trans. Learn. Technol. 1–13 (2018). doi:10.1109/TLT.2018.2823317
19. Abelardo, P., Jelena, J., Shane, D., Dragan, G., Negin, M.: Using learning analytics to scale the provision of personalised feedback. Br. J. Educ. Technol. 0, (2017). doi:10.1111/bjet.12592
20. Pardo, A., Mirriah, N.: Design, Deployment and Evaluation of a Flipped Learning First-Year Engineering Course. In: Reidsema, C., Kavanagh, L., Hadgraft, R., and Smith, N. (eds.) The Flipped Classroom: Practice and Practices in Higher Education. pp. 177–191. Springer Singapore (2017)
21. Saint, J., Gašević, D., Pardo, A.: Detecting Learning Strategies Through Process Mining. ([submitted])
22. Gatta, R., Lenkowicz, J., Vallati, M., Rojas, E., Damiani, A., Sacchi, L., De Bari, B., Dagliati, A., Fernandez-Llatas, C., Montesi, M., Marchetti, A., Castellano, M., Valentini, V.: pMineR: An Innovative R Library for Performing Process Mining in Medicine. In: Teije, A. ten, Popow, C., Holmes, J.H., and Sacchi, L. (eds.) 16th Conference on Artificial Intelligence in Medicine, AIME 2017. , Vienna (2017)