A learning analytics approach towards monitoring co-regulation by human tutors, in a virtual classroom environment

Madiha Azeem Khan

PhD thesis

University College London

I, Madiha Azeem Khan, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis

Acknowledgements

To Professor Rose Luckin. Thank you for your support since the very beginning of this journey. Your expertise, patience and wisdom have been a great source of support and inspiration along the way.

To Professor Mutlu Cukurova. Thank you for the meticulous supervision. You've challenged and broadened my approach towards my research, and made it very enjoyable in the process.

To Tom Hooper and Candida Crawford, and the Third Space Learning team. Thank you for being such an open and collaborative team, and for making this research possible.

.

Acknowledgements

This thesis is dedicated to my parents.

To my mum. Thank you for making me believe I can fly.

To my dad. Thank you for making me believe that you'll be there to catch me if I fall.

To Faisal, Suleman, Iman and Laila. Thank you for your patience and understanding while I worked all the hours in the day (and night). I hope to make you proud.

Abstract

This thesis explores how learning analytics can be used to monitor the co-regulation of upper primary school-aged learners, by human tutors in a virtual classroom environment. Co-regulation is an important route through which young learners acquire SRL skills. However, limited research has been conducted towards understanding how co-regulation occurs in practice. The proliferation of online learning environments offers new research opportunities that discover favourable and maladaptive co-regulation patterns. Tracking patterns of co-regulation using the data generated from online environments can enable timely interventions to support co-regulation, boosting the development of learner SRL

In this thesis, I propose a theoretically grounded approach towards using learning analytics to monitor co-regulation in a Virtual Classroom Environment (VCE). An established model of SRL is contextualised to enable an investigation into how human tutors co-regulate learners in a VCE. A framework of signifiers is designed using a mixed methods approach before being implemented through a combination of learning analytics for datasets comprising high and low-performance tutors. Process mining is implemented to identify holistic and structural features of tutor-learner interactions, while pattern mining and decision trees are implemented to identify statistically significant differences. I experiment with using average progress scores as an alternative criterion for tutor performance to evaluator rankings for student-centred teaching, to explore the differences in the results obtained. My results show that high-performing tutors are more likely to use a diverse and complex range of tutoring practices which extend the learning activities set out in the VCE. Open-ended forms of engagement that encourage learner participation and agency feature prominently. On the other hand, low-performing tutors tend to use a more restricted range of tutoring practices which follow the script of the VCE platform more closely and rely on directive forms of engagement. There are similar differences between the co-regulatory practices of high versus

low-performing tutors, regardless of the performance criterion used. However, when average progress scores are used as the criterion for tutor performance, the differences between the co-regulatory practices of high and low-performing tutors are less stark. While co-regulatory practices are a strong predictor of tutor performance when evaluator rankings are used as the performance criterion, they are not able to reliably predict tutor performance to attainment. My work has a range of operational implications, including the development of intelligent approaches towards tutoring training and evaluation, and timely scaffolding of tutor practices to co-regulation.

Impact statement

This research has multiple implications for research and the operations of a virtual classroom environment. From a research perspective, this work proposes a contextualised model of self-regulated learning, which can be used as the basis for a theoretically grounded framework of signifiers, and to investigate co-regulation in a virtual classroom environment. This research helps to develop a granular pedagogical understanding of the distinct tutoring practices employed by high versus low-performing tutors to co-regulation. The performance measure includes both evaluator rankings for student-centred teaching, and average progress scores, enabling an exploration of the distinct tutoring practices associated with each measure. In addition, this work also helps to build our understanding of how tutors build tutoring practices in a virtual classroom environment (VCE), and the differences in the role of the VCE for high versus low-performing tutors.

From an operational perspective, this research can potentially be used to inform the learning design of virtual classroom environments through the inclusion of scaffolds for tutor practices. It can also potentially be used to inform tutor training and evaluation, through highlighting significant patterns of tutor-learner interaction which can be used as the basis for intelligent interactive dashboards. Further research is required to explore the operational implications of the work set out in this thesis.

Abbreviation	Term
APS	Average Progress Score
HLE	Hyper-media Learning Environment
ITS	Intelligent tutoring system
SCT	Student Centred Teaching
SRL	Self-regulated learning
SSRL	Socially shared regulation of learning
TSL	Third Space Learning (industrial supplier of VCE)
VCE	Virtual Classroom Environment

Contents

Chapter 1: Introduction and Context **22**

Introduction 22

1.1 Study Context, Approach and Preview of Findings 25

1.2 Structure of thesis 27

Chapter 2 Self-Regulated Learning **33**

2.1 Conceptualisation of the regulation of learning 33

Introduction 33

2.1.1 First wave of research into SRL (1980's till late 1990's) 35

2.1.2 Second wave of research (late 1990's till 2015) 38

2.1.3 Third wave of research into self-regulated learning (2015-to date) 47

2.1.4 The theoretical role of context in SRL: areas for further research 50

Section summary 50

2.2 Regulation of learning in collaborative settings 51

Introduction 51

2.2.1 Constructs of regulation of learning in collaborative learning settings 52

2.2.2 Theory based frameworks for co-regulation 55

2.2.3 Empirical research on regulation of learning in collaborative learning settings 61

2.2.4 Empirical research on regulation of learning in collaborative learning settings: areas for further research 73

Summary 75

2.3 Measurement of the regulation of learning: Data types and signifiers 76

Introduction 76

2.3.1 First wave of research 77

2.3.2 Second wave of research 79

2.3.3 Third wave of SRL research 83

2.3.4 Gaps in the literature review 88
Summary 88
2.4 Measurement of the regulation of learning: Learning analytics for the regulation of learning 90
Introduction 91
2.4.1 Overview of analytical techniques (2010-2022) 91
2.4.2 Sequence Mining 95
2.4.3 Process Mining 100
2.4.4 Predictive analytics 110
2.4.5 Gaps identified from literature review 113
Summary 114
2.5 Research questions 116
Introduction 116
2.5.1 Formulation of research questions 117
2.5.2 Presentation of research questions 118
Chapter 3: Methodology **120**
3.1 Studied context and design-based research methodology 120
3.1.1 Study context 120
3.1.2 Design-based research methodology 124
3.1.3 DBR cycle mapped to the research questions, outcomes, validation and contribution 126
3.1.4 DBR cycle mapped to phases of work 128
3.1.5 Ethics 134
3.2: Methodology (data sampling) 136
Introduction 136
3.2.1 Data-sets used, mapped to stages of work 137
3.2.2 Data set A 138

3.2.3 Data set B 139
3.2.4 Data set C 140
3.2.5 Data set D 140
3.2.6 Data set E 140
3.2.7 Data set F 147
3.2.8 Data set G 152

Chapter 4 Developing signifiers for SRL **157**

4.1 Contextualised model of SRL 157
Introduction 157
4.1.1 Theory-based frameworks implemented in previous studies 158
4.1.2 Approach for building contextualised model of SRL 159
4.1.3 Contextualised model of SRL 180
Summary 182
4.2 Theoretically based Framework of Signifiers 183
Introduction 183
4.2.1 Restructure the contextualised model of SRL 184
4.2.2 Review the data generated from the VCE 187
4.2.3 Map the literature to the framework to identify potential signifiers 189
4.2.4 Integrate empirical data into the framework of signifiers 195
4.2.5 Learner SRL processes 198
4.2.6. Empirical review for learner framework of signifiers 200
4.2.7 Observations of learner operations 201
Summary 208

Chapter 5: Analysing SRL and co-regulation **210**

5.1 Process maps for sessions conducted by high-performing versus low-performing tutors 210

Introduction....210
5.1.1 Data sampling....211
5.1.2 Selecting and applying the appropriate algorithm....213
5.1.3 Data processing and configuration of parameters....220
5.1.4 Producing process maps....221
5.1.5 Process mining, using data sample E (using evaluator rankings as the criterion for tutor performance)....222
5.1.6 Process mining using data sample F (using average learner progress scores as the criterion for tutor performance)....242
5.1.7 Comparison of analysis from process mining, using different performance measures....253
Summary....258
5.2 Pattern mining....260
Introduction....260
5.2.1 Identifying relevant pattern mining algorithms for my study context....261
5.2.2 Pattern mining using data sample E (using evaluator rankings as the criterion for tutor performance)....264
Implementation of CM- SPAM....264
Implementation of EMMA algorithm....270
5.2.3 Pattern mining using data sample F (using average progress scores as the criterion for tutor performance)....272
Implementation of CM-SPAM....272
Implementation of EMMA algorithm....275
Summary....276
5.3 Decision trees....279
Introduction....279
5.3.1 Selection of classifiers....280

5.3.2 Implementation of predictive analytics for dataset E ((evaluator rankings used as criterion for performance) 280

JRIP algorith 281

J48 algorithm 284

5.3.3 Implementation of predictive analytics for dataset F (using APS as performance criterion) 289

JRIP algorithm 290

Implementation of J48 algorithm 292

Summary 297

Chapter 6 Discussion **298**

Introduction 298

6.1 Theoretical implications 299

6.2 Implications for framework of signifiers 305

6.3 Approach towards the use of learning analytics 308

6.4 Implications of results 314

Conclusion 328

6.5 Limitations 331

Chapter 7 Conclusions **335**

Introduction 335

7.1 Recap of aims of research 335

7.2 Research summary and conclusions 337

7.3 Key Contributions 346

7.4 Future Research Directions 348

Appendix

353

Appendix A 353
Appendix B 354
Appendix C 354
Appendix D 358
Appendix E 359
Appendix F 360
Appendix G 365
Appendix H 367
Appendix I 369
Appendix J 369
Appendix K 370
Appendix L 371

References **373**

List of Tables

Table 1. Zimmerman cyclical phases of self-regulation 40

Table 2. Constructs of regulation in collaborative learning settings 54

Table 3. Studies on co-regulation of learners, mapped to research setting, types of data collected and analytical frameworks used 57

Table 4. Examples of ITS with pedagogical agents 69

Table 5. Examples of pedagogical agent systems and measures of SRL applied 85

Table 6. Systematic review of the learning analytics used in SRL papers between 2011-2019 (ElSayed et al., 2019) 92

Table 7. Systematic review of the types of data used in LA/EDM papers on SRL between 2011-2019 (ElSayed et al., 2019) 93

Table 8. Process Mining tools used in Learning Analytics research 103

Table 9. Types of data captured by TSL 123

Table 10. Mapping data types to ethical considerations 135

Table 11. Data sets mapped to research questions 137

Table 12. Breakdown of data sample A (used to conduct exploratory data analysis, and build the contextualised model of SRL 139

Table 13. Breakdown of data sample B (used to build the framework of signifiers) 139

Table 14. Sessions observed to validate the framework of signifiers (from the learner perspective) 140

Table 15, Data-set D (sessions analysed to evaluate inter-rater reliability) 140

Table 16. Criteria for human evaluators to assess tutor skill at promoting SCT (student-centred teaching) 141

Table 17. Data-set E (used for implementation of learning analytics) *based on human evaluator rankings for SCT 144

Table 18. Data-set F (used for implementation of learning analytics) *based on average learner progress scores 150

Table 19. Overview of the tutor interview programme 153

Table 20. Themes underlying the tutor interviews, and the analysis conducted of the interview output 153

Table 21. Interview analysis 155

Table 22. Potential tutoring practices which can co-regulate learners in a VCE 164

Table 23. SMART operations of SRL as defined by Winne (Winne, 2010) 166

Table 24. Potential co-regulatory practices in a VCE mapped to literature 167

Table 25. ITS pedagogical agents which promote generation of SRL products 175

Table 27. Co-regulatory tutoring practices categorised by type of practice 187

Table 28. Literature review mapped to framework to identify potential signifiers ... 189

Table 29. Iterated framework of signifiers (tutor practices) 197

Table 30. Framework of signifiers for learner operations (based on literature review only) 199

Table 31. Framework of signifiers for learner operations 200

Table 32. Observation protocol 202

Table 33. Breakdown of the data captured using the observations protocol 206

Table 34. Total number of operations signified, across all sessions observed 207

Table 35. Stages of process mining mapped to chapter sections 211

Table 36. Process Mining algorithms used in Learning Analytics research 213

Table 37. Types of Pattern Mining 262

Table 38. S frequent patterns for high-ranked tutors *Note: there are 29 sessions in total conducted by high-ranked tutors 265

Table 39. S frequent patterns for low-ranked tutors *Note: there are 30 sessions in total conducted by low-ranked tutors....266

Table 40. Significant s- frequent patterns for dataset E (* statistically significant p<0.05)....266

Table 41. Mean and standard deviation of the episode support of patterns in high and low-ranking tutor groups....272

Table 42. S frequent patterns for high APS tutors *Note: there are 20 sessions in total conducted by high APS tutors....273

Table 43. S frequent patterns for high APS tutors *Note: there are 21 sessions in total conducted by high APS tutors....273

Table 44. Significant s- frequent patterns for dataset F....274

Table 45. Mean and standard deviation of the episode support of patterns in high and low APS tutor groups....276

Table 46. Stratified cross-validation (JRIP, dataset E)....282

Table 47. detailed accuracy by class (JRIP, dataset E)....282

Table 48. Stratified cross-validation (JRIP, dataset E, normalised and binned data) 284

Table 49. detailed accuracy by class (JRIP, dataset E, normalised and binned data) 284

Table 50. Stratified cross-validation (J48, dataset E)....285

Table 51. detailed accuracy by class (J48, dataset E)....285

Table 52. Stratified cross-validation (J48, dataset E, normalised and binned data)..286

Table 53. detailed accuracy by class (J48, dataset E, normalised and binned data)..287

Table 54. Stratified cross-validation (J48, dataset E, normalised and binned data, minNo=15)....287

Table 55. detailed accuracy by class ((J48, dataset E, normalised and binned data, minNo=15)....288

Table 56. Stratified cross-validation (JRIP, dataset F) 290

Table 57. detailed accuracy by class (JRIP, dataset F).....290

Table 58. Stratified cross-validation (JRIP, dataset F, normalised and binned data) 291

Table 59. detailed accuracy by class (JRIP, dataset F, normalised and binned data) 291

Table 60. Stratified cross-validation (J48, dataset F) 292

Table 61. detailed accuracy by class (J48, dataset F) 293

Table 62. Stratified cross-validation (J48, dataset F, normalised and binned).....294

Table 63. Detailed accuracy by class (J48, dataset F, normalised and binned).....294

Table 64. Stratified cross-validation (J48, dataset F, normalised and binned, remove pattern B) 295

Table 65. Detailed accuracy by class (J48, dataset F, normalised and binned, remove pattern B) 295

Table 66. TSL rubric for student-centred teaching, with suggested extensions based on my research findings…….318

List of Figures

Figure 1. Research questions in the context of SRL 25

Figure 2. Structure of thesis 28

Figure 3. Zimmerman model of self-regulated learning 40

Figure 4. Winne and Hadwin's model of self-regulated learning (1998) 44

Figure 5. Research questions 119

Figure 6. Screenshot of the interactive Third Space Learning Whiteboard……...121

Figure 7. Overarching DBR cycle mapped to the research questions, outcomes, validation and contribution 128

Figure 8. The DBR cycle mapped to phases of work 129

Figure 9. Tutor scores for promoting student-centred learning, awarded by human evaluators *tutor ID's have been anonymised 143

Figure 10. Learner diagnostic scores for high-ranked and low-ranked tutor dataset. 146

Figure 11. Learner levels of prior attainment, as indicated by school teachers..... 147

Figure 12. Distribution of average learner progress scores for tutors *tutor ID's have been anonymised..... 150

Figure 13. Learner diagnostic scores for high APS and low APS tutor dataset 151

Figure 14. Learner levels of prior attainment, as indicated by school teachers..... 152

Figure 15. Winne and Hadwin (1998) model of SRL..... 162

Figure 16. Contextualised Winne and Hadwin's model of SRL, for exploring how tutors co-regulate learner operations in a VCE..... 181

Figure 17. Petri model of low-ranked tutor dataset, mined with Inductive Miner algorithm 216

Figure 18. Process model of low-ranked tutor dataset, mined with Disco algorithm 219

Figure 19. Process map 1, for high-ranked tutor sessions (sliders set to 0%) 225

Figure 20. Process map 2 for low-ranked tutors (sliders set to 0%) 226

Figure 21. Process map 5 low-ranked tutors (signifiers view, sliders set to 0%) 232

Figure 22. Process map 6 high-ranked tutors (signifiers view, sliders set to 0%) 233

Figure 23. Process map 9 High-ranking tutor sessions (filtered by tutor behaviours) 237

Figure 24. Process map 10 Low-ranking tutor sessions (filtered by learner behaviours) 238

Figure 25. Process map 11 High-ranking tutors (filtered by tutor behaviours) 238

Figure 26, Process map 12 High-ranking tutors (filtered by learner behaviours) 239

Figure 27. Process map 13 high APS tutors (event categories, sliders at 0%) 244

Figure 28. Process map 14 low APS tutors (event categories, sliders at 0%) 245

Figure 29. Process map 15 high APS tutors (event category view, event sliders at 10% and path sliders at 25%) 246

Figure 30. Process map 16 low APS tutors (event category view, event sliders at 10% and path sliders at 25%) 247

Figure 31. Process map 19 high APS tutors (filtered by tutor behaviours) 250

Figure 32. Process map 20, high APS tutors (filtered by learner behaviours) 251

Figure 33. Process map 21, low APS tutors (filtered by tutor behaviours) 251

Figure 34. Process map 22 low APS tutors (filtered by learner behaviours) 252

Figure 35. Decision tree for dataset E, using normalised, un-binned data 285

Figure 36. Decision tree for dataset E, using normalised, binned data 286

Figure 37. Decision tree for dataset E, using normalised, binned data (minNo parameter tuned to 15) 287

Figure 38. Decision tree for dataset F, using normalised, un-binned data 292

Figure 39. Decision tree for dataset F, using normalised, binned data 293

Figure 40. Decision tree for dataset F, using normalised, binned data (pattern B removed a attribute) 295
Figure 41. Contextualised model of SRL 300
Figure 42, Relative frequency of events in sessions conducted by low-ranked tutors 309
Figure 43. Relative frequency of events in sessions conducted by high-ranked tutors 310
Figure 44. Research questions mapped to goals and contributions 336
Figure 45. Contextualised Winne and Hadwin model of SRL, for exploring how tutors co-regulate learner operations in a VCE 339
Figure 46. Intelligent tool to support QA 351
Figure 47. Intelligent tool to support tutor training 352

Chapter 1: Introduction and Context

Introduction

From ancient education systems to the modern classroom, natural language dialogue has been a fundamental component of teaching and learning. Aristotle's approach towards teaching his disciples stressed the importance of oral communication as a specific way of creating knowledge. Similarly, Socrates emphasised the art of questioning, with the Socratic method continuing to be used in classrooms around the world.

While the exploration of natural language dialogue can provide rich insight into teacher-learner interactions, corresponding research has been limited. Understanding how teachers can impart knowledge and secondary skills to students through dialogue has traditionally happened at a relatively macro level, due to difficulties in collating the required data. As online learning environments have developed, new opportunities present for a granular, dynamic and contextualised understanding of learning.

My research explores how tutors co-regulate learners in an online learning environment, by analysing natural language dialogue and online trace data, using a theoretically grounded approach. I adapt the Winne and Hadwin (1998) model of SRL to enable the development of a theoretically grounded framework of signifiers. Learning Analytics is implemented to discover the differences in tutor-learner interaction, for high-performing tutors versus low-performing tutors. Through implementing a combination of learning analytics, I aim to build a holistic and nuanced understanding of the differences between high-performing and low-performing tutors.

The development of virtual classrooms presents new opportunities from a research and operational perspective. Work which explores effective and maladaptive patterns of interaction using natural dialogue amongst other data sources, can build a rounded and contextualised understanding of how oral communication drives learning. In this thesis, I

have explored how learning analytics can be used to identify effective and maladaptive patterns of tutor-learner interaction, to SRL. I have developed a theoretically grounded framework of signifiers, and operationalized this through a combination of different forms of learning analytics, to explore how human tutors co-regulate learners in a VCE. This work has the potential to be operationalised, through machine learning techniques and foundational LLMs. Potential operational implications are far-reaching, and could, for example, include the intelligent scaffolding of oral communication to enhance the quality of teaching and learning outcomes.

In the remainder of this introductory chapter, I will explain the aims of my research, provide a brief overview of my approach, and set out the structure of this thesis

Background context and aims of research

As previously discussed, teachers have used oral communication to build both specific domain knowledge and transferable skills over the ages. This includes transferable skills such as Self-Regulated Learning (SRL). It is well-established that co-regulation is an important means of acquisition of regulatory skills, particularly for younger learners (see, for example, Chan, 2012; Gillespie, 2015). However, while there is a vast body of research on SRL, most research focuses on either the learner working alone, or as part of a larger group. Limited research has been conducted in contexts where the learner is working on a one-to-one basis with another person. A deeper understanding of how young learners are co-regulated by their teachers is needed, both from a research and operational perspective, to highlight effective co-regulatory practices. This work can build our understanding of how co-regulation occurs in practice, and highlight effective and maladaptive co-regulatory practices. In addition, analysing contextualised natural dialogue patterns between a teacher and a learner can build an understanding of co-regulation as a dynamic process, which unfolds during a learning activity.

Separately, there is also a need for further research regarding one-to-one human tutoring. Research within the field of online tutoring has been largely focused on intelligent tutoring systems (e.g., Graesser et al., 2012). While intelligent tutoring systems (ITSs) are a promising area of research in their own right, it can be argued that one-to-one human tutoring is more likely to cater for the complex socio-cognitive needs of disadvantaged learners. While the impact of one-to-one human tutoring on learning outcomes is well established (Cohen et al., 1982), there is a need to build theoretical frameworks to highlight granular effective and maladaptive tutoring practices. This research can be built upon, for example, through intelligent tools which support the human tutor to enhance learning outcomes, rather than replacing the tutor altogether in ITSs. Previously, research in one-to-one human tutoring has been limited by a lack of access to sessions conducted by qualified tutors (cf. Lepper & Woolverton, 2002); however, the development of online tutoring platforms presents new opportunities for research and analysis.

The research described in this thesis was conducted within the above context and includes an extensive literature review (set out in chapter 2 of this thesis) to crystallise the research gap and formulate research questions. Figure 1 presents the research gap, and research questions identified. In this research I set out to study tutor-learner interactions in a Virtual Classroom Environment (VCE), with respect to SRL. My overarching research question is:

How can we use learning analytics to detect the influence of human tutoring practices on learner SRL processes, for high and low-performing tutors in a Virtual Classroom Environment (VCE)?

- *RQ1: How should SRL be conceptualised to explore the influence of human tutors on learner SRL processes?*

- *RQ2: How can we derive theoretically meaningful signifiers, to represent the influence of human tutors on learner SRL processes in a VCE?*
- *RQ3: What are the key differences in the patterns of tutor-student interaction, for high and low-performing tutors?*

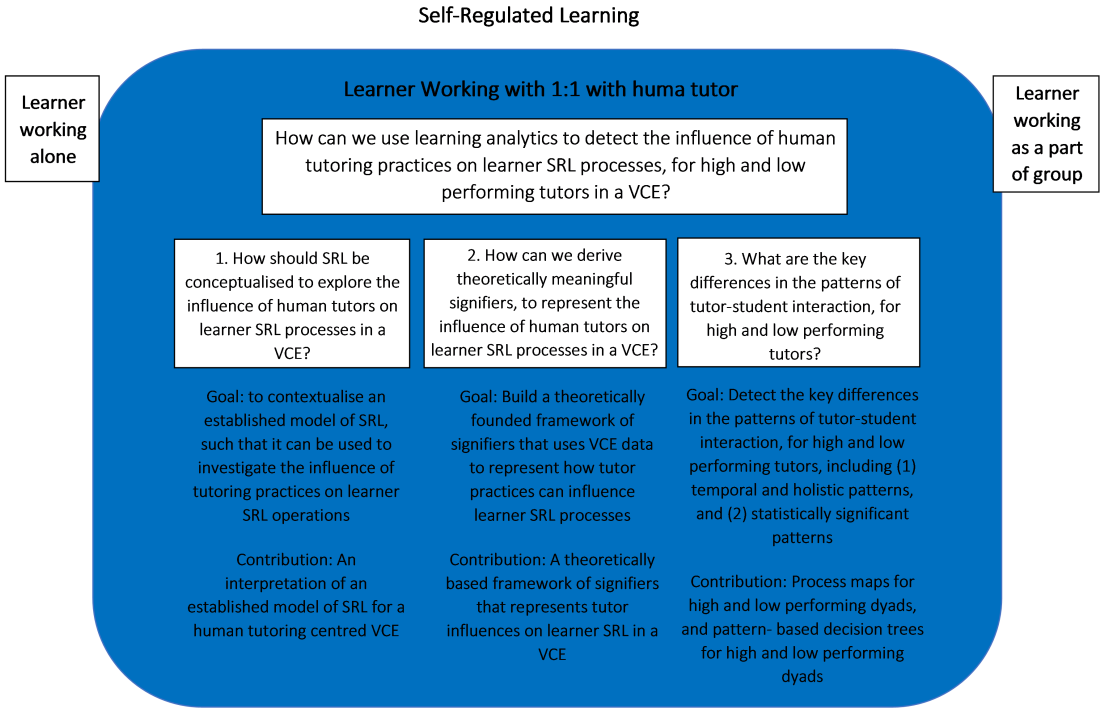


Figure 1. Research questions in the context of SRL

1.1 Study Context, Approach and Preview of Findings

To address my research questions, research was conducted using the virtual classroom environment of a commercial supplier called TSL. Using human tutors, TSL supplies mathematics tutoring to upper primary school-aged children from disadvantaged schools. The tutoring relies on a one-to-one model and is mainly driven by natural dialogue between the tutor and the learner, using an interactive whiteboard and VCE tools.

My research adopts a theory- based approach towards analysing patterns of tutor-learner interaction in the VCE. I adopt and contextualise an established model of SRL, to enable the study of co-regulation in a VCE. I use a mixed methods approach to build a theoretically grounded framework of signifiers, that represents how tutors co-regulate learners in a VCE. Finally, I implement a combination of learning analytics to develop a layered understanding of co-regulation in a VCE, and the distinct practices of high-performing versus low-performing tutors. This includes:

- Process mining to develop a holistic understanding of co-regulation as a dynamic process which unfolds over the course of a learning activity.
- Pattern mining to identify statistically significant and specific patterns of tutor-learner interaction, which occur across and within tutoring sessions, for high-performing versus low-performing tutors
- Decision trees, to highlight the predictive power of specific patterns and combination of patterns

My research demonstrates that high-performing tutors demonstrate more complex patterns of interaction with learners, relying on a diverse range of tutoring practices. They tend to build on open ended forms of tutor-learner engagement, and encourage learner autonomy. In contrast, low-performing tutors tend to adopt simpler patterns of interaction with learners, which focus around tutor-led forms of engagement. They also tend to follow the structure of the learning activity set out in the VCE more closely, while high-performing tutors appear to be confident to build fluid practices on top of the activities set out in the VCE.

In addition to highlighting the distinct practices of high-performing versus low-performing tutors, my research will demonstrate the value of using different forms of learning analytics. Learning analytics can be descriptive in nature (such as process mining), or

generate results which can be statistically validated (e.g., pattern mining and decision trees). My research has relied on a combination of different forms of learning analytics. I discuss the value of each of these forms of analysis independently from each other, and on a combined basis, in light of my research findings.

Finally, my research will also highlight promising new areas for research. For example, there are opportunities to build intelligent tools to support tutor training and quality assurance. In addition, the design of Virtual Classroom Environments is a ripe area for research. Future research can examine the specific role of VCE design in enhancing various tutoring practices. Finally, advances in foundational large language models (LLMS) have significant operational implications for the use of natural language dialogue, with greater opportunities to automate signifiers from this data source. Future research can explore the opportunity to automate the signifiers identified through this research using foundational LLMs. The thesis will provide the basis for future research utilising such state-of-the-art language models to diagnose and support tutor-student interactions in VCEs.

1.2 Structure of thesis

The remainder of this thesis follows the structure set out in Figure 2.

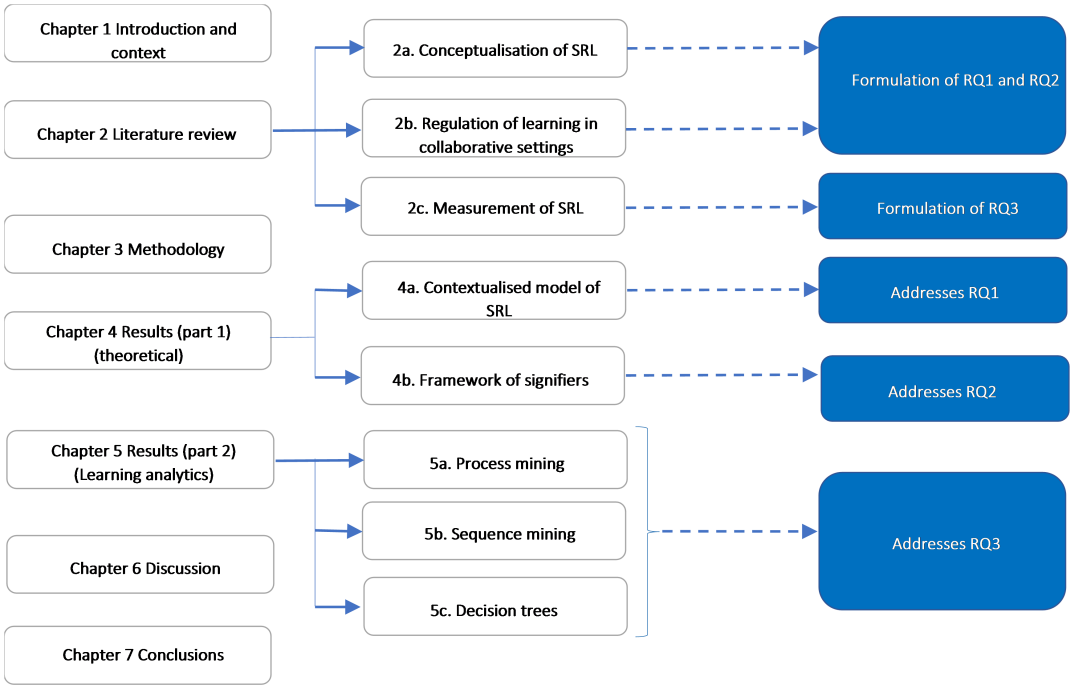


Figure 2. High-level structure

In order to introduce the remainder of the thesis, I have provided brief abstracts below of the thesis chapters.

- Chapter 1: Introduction.
- Chapter 2: Literature review (Self-regulated learning)
    - *Chapter 2.1: Conceptualisation of SRL.* This chapter reviews how research into self-regulated learning has evolved to conceptualise the regulation of learning in increasingly complex and contextualised ways. I demonstrate how research into the regulation of learning has shifted from emphasising individual self-regulation and internal processes towards integrating the influence of context. This includes a discussion of research into the influence of social context on self-regulation, with researchers exploring the different types of regulation that occur when the learner works alone, versus working with others. Throughout the chapter, I describe the evolution of the

conceptualisation of SRL across the past four decades, specifically examining the changing theoretical role of context.

- *Chapter 2.2: Regulation of Learning in collaborative settings.* I move on to review the constructs used to explore the regulation of learning in the context of collaborative learning settings, before analysing the findings from research conducted in this context
- *2.3: Measurement of the regulation of learning: Data types and signifiers*: In this chapter, I review the development of research into the measurement of SRL. This chapter subsection deals with the evolution of the types of data used to measure SRL, followed by a discussion of the research gaps.
- *2.4: Measurement of the regulation of learning: Learning analytics for the regulation of learning:* In this chapter, I review the analytical techniques which have been applied to explore SRL over the last decade. I start with an overview of the different techniques that have been used, prior to selecting a few techniques for a deep dive. For the selected techniques, I review the implementation methodologies, and explore contexts within which the analytics have been applied. I identify gaps in the literature.
- *2.5: Research questions:* I present the key research questions underpinning this thesis, based on the gaps in the literature identified.
- Chapter 3 Methodology
    - *3.1 Methodology: Studied context and design-based research methodology.* In this chapter, I will detail the study context and provide an overview of the design-based research approach taken.
    - *3.2 Methodology: Data sampling.* In this chapter, I will set out the characteristics of the data sampled for my thesis. I will start by setting

out the different sets of data used, aligned to stages of work. I will then review the structure and characteristics of each data set.

- Chapter 4: Results (part 1): Developing a framework of signifiers:
  - *4.1: Contextualised model of SRL.* In this chapter, I will further address RQ1: *How should SRL be conceptualised to explore the influence of human tutors on learner SRL processes in a VCE?* To fully address RQ1, I will propose a contextualised model of SRL which can be used to explore the influence of human tutors on learner SRL. This chapter begins by analysing the theoretical frameworks used in past studies, for their suitability to the studied context, and justifies the need to develop a contextualised model and theoretical framework for my research. I review established models of SRL to judge their suitability for the studied context, and select a model for contextualisation, prior to refining this model through exploratory data analysis and interpretation for my study context. I validate the contextualised model by mapping relevant literature to the elements of the model, prior to presenting the final iteration of the contextualised model of SRL.
  - *4.2: Theory-based framework of signifiers.* In this chapter, I address *RQ2 How can we derive theoretically meaningful signifiers, to represent the influence of human tutors on learner SRL processes in a VLE?* I developed a theoretically grounded framework of signifiers which can be used to represent how tutors co-regulate learners in a VCE. I set out the approach adopted to build a framework of signifiers, before identifying signifiers for tutor practices and learner operations

in the context of a VCE. An observation study is conducted to validate the signifiers of learner operations

- Chapter 5: Results (part 2): Analysing SRL and coregulation
  - *5.1: Implementation of Process Mining.* In this chapter, I will implement the framework of signifiers (developed in chapter section 4.2) and analyse the patterns of labelled signifiers using process mining. Following the implementation of process mining, I will discuss my findings in the context of the output from tutor interviews. I will analyse the implications of my findings for the framework of signifiers and consider the merits and limitation of process mining, in contrast to a frequency-based approach. Finally, I will conclude by summarising my findings and further analysis planned.
  - *5.2: Implementation of Pattern mining*. In this chapter, I implement pattern mining for my datasets. I start off by reviewing the range of pattern mining algorithms, and their relevance for my study context. This includes evaluation of which algorithm or combination of algorithms will be best suited to address my research questions. I implement the relevant algorithms for my datasets and highlight significant sequences. Algorithms are implemented for both dataset E (where tutor performance is determined using evaluator rankings), and dataset F (where tutor performance is determined based on objective measures of learner progress). Finally, I analyse the results obtained and explore the differences between the results obtained for dataset E and dataset F. I also consider how the output from pattern mining can

be interpreted in light of the results obtained from process mining, and implications for further research.

- *5.3: Decision trees.* Following the identification of statistically significant sequences in 5.2, I explore the use of these sequences to build predictive analytics. I start by selecting predictive analytics, and pre-processing my data sample. I implement the algorithms for my dataset, experimenting with using binned and un-binned data. Classification models are built for both dataset E (with tutor performance determined by human evaluator rankings), and dataset F (with tutor performance determined by average progress scores). I analyse the results obtained for each data-set, and explore the differences between the classification models for each dataset. I also interpret the classification models in light of the process maps generated in 5.1. I primarily aim to explore how effectively decision trees can classify tutors into high and low-performing groups, using sequences of tutor-learner interactions.

- Chapter 6: Discussion. In this chapter, I discuss the implications of my research findings from theoretical, analytical and operational perspectives, and highlight the limitations of my research
- Chapter 7: Conclusions. In this chapter, I summarise my findings, and draw some conclusions. In addition, I identify new areas for future research.

Chapter 2 Self-Regulated Learning

2.1 Conceptualisation of the regulation of learning

Introduction

This chapter of my thesis contains my literature review. As per the introduction to my thesis (chapter 1), my broad area of focus is self-regulated learning (SRL). SRL is the learner's active management of their learning processes, and includes cognitive, metacognitive and affective elements (e.g., Zimmerman, 2000).

This chapter considers how research into SRL has evolved to conceptualise the regulation of learning in increasingly complex and contextualised ways. In particular, I review the occurrence of three distinct waves of research into SRL, with differing emphasis placed on the role of time and context in each wave (Panadero et al., 2016). I demonstrate how research into the regulation of learning has shifted from emphasising individual self-regulation and internal processes towards integrating the influence of context.

This chapter follows the structure set out below:

- A brief overview of the three waves of research into SRL, highlighting the key differences in the changing theoretical role of context in each wave
- A detailed review of the first wave of research into SRL (1980s and 1990s), including an exploration of the fundamental theoretical components of SRL, and key differences between established instruments used to measure SRL
- A detailed review of the second wave of research into SRL (late 1990s till 2015), including reviews of established models of SRL and the key differences between these models
- A detailed review of the third wave of research into SRL (2015 – present), including reviews of researcher-designed platforms being used to measure SRL

- Finally, I consider interesting areas for further research from a theoretical perspective. The highlighted areas are the focus of the literature review that is set out in subsequent chapters.

Waves of research into SRL

Overview of the waves of research into SRL

Panadero et al. (2016) reviewed past research into SRL and defined the occurrence of three distinct waves of research, each with a differing emphasis placed on the role of time and context. Context refers to both static factors (e.g. the study domain, resources), as well as the dynamic factors in the learning environment (e.g. teacher cues, degree of interaction with peers). The waves of research are briefly outlined here, prior to additional detail being provided in the sections below:

- **1980's and early 1990's**: The **first wave** of research was dominated by early models of SRL. In this wave, research adopted a static conceptualisation of SRL as an aptitude, and measurement relied on self-report data, mainly collected using interview schedules and questionnaires (Zimmerman, 2008), for example, the Motivated Strategies for Learning Questionnaire (MSLQ; Pintrich et al., 1993). The role of context was recognised at a macro level (e.g. the influence of domain) but did not accommodate the influences of contextual changes during a learning activity. Conceptualisations of SRL focused on the building blocks of SRL, rather than the development of an architecture which represents the relationship between the SRL components.
- **Late 1990's till 2015:** The **second wave** of research emerged, in which a dynamic, process-oriented view of SRL began to dominate research. The focus shifted towards understanding how time and context could influence learner SRL throughout an activity. Updated models of SRL were developed which could act as the foundation

for research which investigated SRL from a temporal perspective (e.g., Winne & Hadwin, 1998; Zimmerman, 2000)

- **2015-present**: Current research is often regarded as the **third wave** of research into SRL. In this wave, researchers explore how real-time measurement of SRL can be combined with intervention. The role of time and context have moved beyond being guiding factors, towards fundamentally determining how SRL is measured and supported.

In parallel to the role of context being emphasised, new constructs have emerged in the past few decades, to fully explore the regulation of learning in settings where a learner is working with others. Researchers have identified *co-regulation* and socially *shared regulation* as potentially occurring alongside self-regulation, in such settings (Hadwin et al., 2011). Further research is needed to elaborate and refine these concepts in relation to each other and as stand-alone constructs in a range of social settings. In particular, more work needs to be done to build theoretically grounded frameworks through which these concepts can be operationalised.

2.1.1 First wave of research into SRL (1980's till late 1990's)

Research on SRL is rooted in diverse theoretical origins (Zimmerman, 2000), however, a consensus slowly emerged in the first wave of research on the broad definition of SRL. Researchers agreed on a common conceptualisation of SRL as being *a learner's active management and adaptation of learning processes to meet their learning goals, and overcome challenges* (Corno, 1986; McCombs, 1986; Schunk, 1986; Wang & Peverly, 1986; Zimmerman, 2000). SRL consists of cognitive elements, metacognitive elements and motivational elements. Cognitive elements are the *intentional implementation of learning processes*, such as learning strategies. Metacognitive elements are the components which *require the learner to consider the process of learning*, such as task analysis, goal setting,

choice of learning strategies, and monitoring and reflections on learning. Motivational elements are motivational *factors which impact on, and in turn are affected by SRL*. For example, learner motivation impacts a learner's will to persist despite challenges (which is a hallmark of SRL), and similarly, working through challenges effectively can boost confidence and thus motivation.

Initial conceptualisations of SRL were primarily focused on the components of SRL. For example, Zimmerman and Martinez-Pons (1986) proposed the Self-Regulated Learning Interview Schedule, which conceptualised self-regulated learning as consisting of 14 learning strategies, spanning the cognitive, metacognitive and motivational dimensions. The strategies identified included self-evaluation (Bandura & Cervone, 1983), organizing and transforming (Baird, 1983; Corno & Mandinach, 1983), goal-setting and planning (Bandura & Schunk, 1981), seeking information (Baird, 1983; Wang, 1983), keeping records and self-monitoring (Spates & Kanfer, 1977), self-consequences (Mace & Kratochwill, 1985), rehearsing and memorizing (McCombs, 1984), seeking peer, teacher, or adult assistance (Zimmerman, 2000), and reviewing notes, tests, or textbooks (Wang, 1983). While SRL was regarded as inherently being a non-social process, researchers did recognise that learning occurs through modelling and observation (based on the prior research of social learning theorists), and help seeking was thus included as a potential regulatory strategy. In addition, the researchers recognised that different learning strategies would be applicable depending on the learning context (e.g. in classroom, outside of classroom). The schedule was thus tailored to different contexts at a macro level. This construct was tested through empirical studies, which correlated student scores on the Self-Regulated Learning Interview Schedule to their scores on standardised achievement tests (Zimmerman & Martinez-Pons, 1986), and in later empirical studies, to teacher observations of student strategy use (Zimmerman & Martinez-Pons, 1988).

In contrast, Pintrich and Zusho (2002) placed a greater level of emphasis on the role of motivation and beliefs in mediating the processing of information by an individual and influencing the subsequent formulation of cognitive and metacognitive strategies. This general cognitive architecture distinguishes the MSLQ from generic study skill inventories The researchers proposed a new self-report instrument which mapped learning strategies to motivational factors. The motivation scales were aligned to three broad areas:

(1) value (intrinsic and extrinsic goal orientation, task value),

(2) expectancy (control beliefs about learning, self-efficacy); and

(3) affect (test/performance anxiety).

The learning strategies were categorised into three different types of strategies: cognitive, metacognitive and resource management strategies. Cognitive strategies included (1) rehearsal (e.g., repetition), (2) elaboration (e.g., paraphrasing), (3) organization (e.g., outlining), and (4) critical thinking (e.g evaluate or apply to a new context). Metacognitive strategies are assessed by one large scale that includes planning (e.g., set goals), monitoring, and regulating strategies. Resource management strategies include (1) managing time and study environment; (2) effort management, (3) peer learning, and (4) help-seeking. This instrument allowed researchers to map the cognitive, metacognitive and resource management strategies used by learners to their motivational states.

As already described, SRL was primarily conceptualised in the first wave of research as a set of strategies spanning the cognitive, metacognitive and motivational dimensions. This conceptualisation of SRL can be characterised as one in which SRL is a static aptitude, as SRL was not considered to evolve over the course of an activity. Self-report and observational data were used to retrospectively measure SRL as a fixed variable. However, the conceptualisation of SRL did account for the impact of context on SRL behaviours at a macro-level. For example, interview schedules were adapted depending on whether the

students were learning inside or outside of the classroom (Zimmerman & Martinez-Pons, 1988). In addition, the first wave of research into SRL established a general cognitive architecture, which accounted for motivational factors mediating the formulation of cognitive and metacognitive strategies (e.g. Pintrich & Zusho, 2002). This cognitive architecture laid the foundation for the granular, process-oriented conceptualisation of SRL which emerged in the second wave of research.

Subsequent research into SRL is built on the fundamental concepts of SRL identified in the first wave of research and the general cognitive architecture established. While later models refine the basic cognitive architecture proposed in the first wave of research to detail the interactions between multiple components of SRL, the SRL components themselves were identified in early research. For example, a significant amount of research was conducted in the first wave on types of cognitive strategies that can used by learners, ranging from basic to more sophisticated types of cognitive engagement. The research from this period remains highly relevant for researchers working to build frameworks of SRL behaviours, and operationalise theoretical models of SRL.

2.1.2 Second wave of research (late 1990's till 2015)

Early conceptualisations of SRL acknowledged the role of context in shaping learner SRL, but were not sufficiently granular to enable researchers to explore SRL as an evolving process. During the second wave of research into SRL, researchers became increasingly interested in granular, holistic representations of *SRL as a process*. These conceptualisations allowed researchers to explore the evolution of SRL across the course of a learning activity, and understand the influence of specific changes in the context (e.g., instructional changes/environmental changes) during the course of the activity.

I have reviewed below some of the major models to emerge from this period, including examples of how they have been operationalised and areas for future work.

Zimmerman (2000) model of SRL

The construct of SRL proposed by Zimmerman and Martinez-Pons (1988) was updated to reflect the complex, cyclical and processual nature of SRL in the model proposed by Zimmerman (2000). Zimmerman developed a social-cognitive model of SRL. This model has three distinguishing features; (1) it is triadic in nature, (2) it conceptualises SRL as a cyclical process, and (3) it acknowledges the key influences on SRL from physical and social resources in the environment.

The Zimmerman model is triadic in nature, that is, it builds on past research and recognises three reciprocal self-regulatory influences, rather than simply considering self-regulation to be driven by personal processes (Bandura, 1986; Zimmerman, 1989). The three influences are personal, behavioural and environmental processes. Personal self-regulation is the learner being aware of, and exercising control over their motivations and mindset with the aim of improving learning outcomes. Behavioural self-regulation is the learner monitoring and adapting their behaviour e.g. choice of learning strategies. Environmental self-regulation is the learner monitoring and adapting the environment to enhance their learning e.g. by reducing distractions, creating incentives or seeking appropriate levels of help. The forms of regulation have reciprocal influence, for example, learner behaviour such as self-recording can reciprocally impact both personal processes such as self-efficacy, and the environment resources available to the learner (Zimmerman, 1989).

In addition to recognising the triadic form of SRL, Zimmerman (2000) also emphasises the cyclical nature of SRL. He sees SRL as a complex process which is impacted by personal, behavioural and environmental factors and evolves in line with these, rather than being a fixed attribute (Zimmerman, 2000). The process of self-regulation is structured across three phases; forethought, performance control/volition, and self-reflection and reaction (see Figure 3). Forethought refers to the processes that precede a learning task, and prepare for it.

Performance or volitional control is processes that occur during execution of the learning task and require learner attention and actions. Self-reflection involves processes that occur after task execution, and shape a learner's response to that experience. These self-reflections, in turn, influence forethought regarding subsequent attempts to engage with the task, completing a self-regulatory cycle (Zimmerman, 2000).

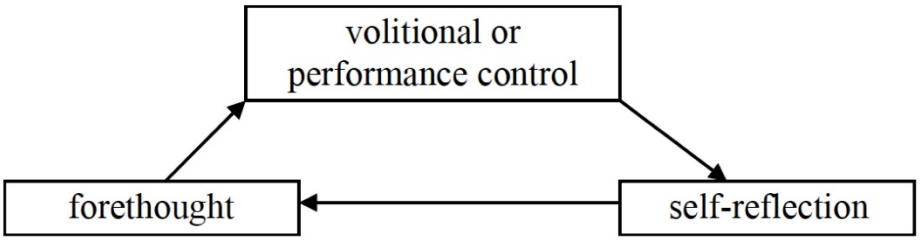


Figure 3. Zimmerman model of self-regulated learning

Each of the phases of SRL includes a number of variables, which are listed in table 1 below (Zimmerman, 1989, 2000). Although a learner is generally presumed to move through the phases in a liner manner, this is not necessary and SRL may engage the variables in any number or order (Zimmerman, 2000).

Table 1. Zimmerman cyclical phases of self-regulation.

Forethought	Performance/volitional control	Self-reflection
Task analysis Goal setting Strategic planning Self- motivation beliefs Self-efficacy Outcome expectations Intrinsic interest/value Goal orientation	Self-control Self-instruction Imagery Attention focusing Task strategies Self-observation Self-recording Self-experimentation	Self-judgment Self-evaluation Causal attribution Self-reaction Self-satisfaction/affect Adaptive/defensive

Significant research has been conducted by Zimmerman and others on each of the variables and the role they play across phases of SRL. For example, research has been conducted on the use of learning strategies as a form of self-regulation. Zimmerman (1989) stated that for a learner to be considered self-regulated, the learning must "involve the use of specified

strategies to achieve academic goals based on self-efficacy perception" (p. 329). This means that for an activity to be considered self-regulated, there must be evidence that the learning strategy was used intentionally, and is related to the learner's perception of self-efficacy or goal setting. Examples of learning strategies include organising and transforming information, self-consequences, seeking information, and rehearsing or using memory aids (Zimmerman & Martinez-Pons, 1988).

Research on learning strategies has focused on the metacognitive decisions to use a strategy (such as rehearsal, organisation or elaboration), or to switch between strategies. In addition to this, researchers have also explored how learners think about learning strategies in the self-reflection phase. For example, Bandura (1986) has identified self-judgment as a key self-reflective process. Self-judgment involves self-evaluating one's performance and making attributions for the results, such as whether a failure in performance is due to ability or effort. Attributions of failures to learning strategies are very effective in boosting motivation and volition in periods of low performance, as these attributions enable a learner to maintain a perception of efficacy until all possible strategies have been tested (Zimmerman & Kitsantas, 1996).

The third key element of the Zimmerman (2000) social cognitive model is the strong influence of the social and physical environment on individual self-regulation. The social and physical environment is seen as a resource for the individual to enhance forethought, performance control, and self-reflection. For example, modelling and certain types of instruction can be used to socially convey self-regulatory skills, such as motivational techniques (e.g. self-praise), or self-corrective methods (e.g. adaptive self-reactions). Conversely, social models can also inhibit the development of self-regulation, for example by modelling maladaptive forms of self-regulation such as defensive self-reactions (Zimmerman, 2000).

The above review of the Zimmerman model highlights the importance of conceptualising SRL as a dynamic, multi-layered and cyclical process, which is shaped by contextual factors such as environmental and social conditions. The interaction between multiple variables, and the shaping role of context on learner SRL, are key features of the Zimmerman (2000) model. This model enables researchers to conceptualise SRL as an evolving process, and examine the impact of changing conditions on the development of SRL over the course of the learning activity.

Boekaerts's (2002) model of self-regulation

Boekaerts's (2002) socio-cognitive model of self-regulation emphasizes the role of context in constructing the theories of self-regulation that guide and inform practice. She argued that theories of instruction were often content-driven, and thus teachers found them hard to transfer to new contexts. To enable principles of instruction to be embedded within teacher practice, it is important to move away from content-driven theories of instruction towards context-driven theories. i.e. theories where the whole self is kept in context. For example, evidence suggested that classrooms should move towards student-centred environments (O'Neill & McMahon, 2005) with Powerful Learning Environments (De Corte et al., 2004) being designed that encourage students to self-regulate, including mastering, evaluating and updating, and being collaborative with their peers. However, teachers have struggled to get specific guidelines from researchers on how to implement principles in new contexts, or deal with challenges e.g., students that demand external regulation or refuse to participate in group work.

Boekaerts also pointed out that conceptualising SRL as an aptitude versus a process was not necessarily a binary decision and often, SRL could be conceptualised as both. For example, one of the key drivers of self-regulation is personal goals. A hierarchy of goals is harboured by an individual, ranging from abstract higher-order goals (e.g., 'be successful') to

specific personal goals (e.g.; ‘do well in my A levels’). Personal goals are implemented through a sequence of actions termed a script (e.g., do homework every day), with the specific nature of a script depending on the demands of the environment (e.g., such as homework being a large component of A-level grades).

Self-regulation can be mapped to this network of goals, with 2 complementary aspects of self-regulation highlighted. Firstly, learners demonstrate an aptitude for SRL, by engaging in a value generating activity, when they choose a script based on their personal goals. Secondly, learners generate an outcome for their learning process, by extending their knowledge of and ability to differentiate between and implement different scripts. For teachers to encourage SRL, they need to make time for learners to engage in both aspects of SRL, including planning and mapping to personal goals at the beginning, and following the activity, reflecting on what scripts have been learnt and their effectiveness. Boekarts argued that most learning settings are collaborative, and learners are influenced by not just academic goals, but also social goals and scripts. There is therefore a double self-regulation process occurring which teachers need to be aware of, and help to facilitate.

Winne and Hadwin model of self-regulated learning

Winne and Hadwin (1998) structure their SRL model in a fluid manner which enables a detailed examination of the interaction between the various components of SRL, including contextual factors. The model acknowledges that SRL occurs across phases but differs from other models by examining the set of information processes that occur within each phase, with the processes being driven and shaped by the components of SRL (Winne & Hadwin, 1998). Researchers have noted that the various components of the Winne and Hadwin (1998) model synthesise components of SRL referred to in existing literature (Azevedo et al., 2011; Greene & Azevedo, 2007) and are referred to using the acronym COPES; Conditions,

Operations, Products, Evaluation, Standards. The model is depicted below, followed by a description of the components of the model.

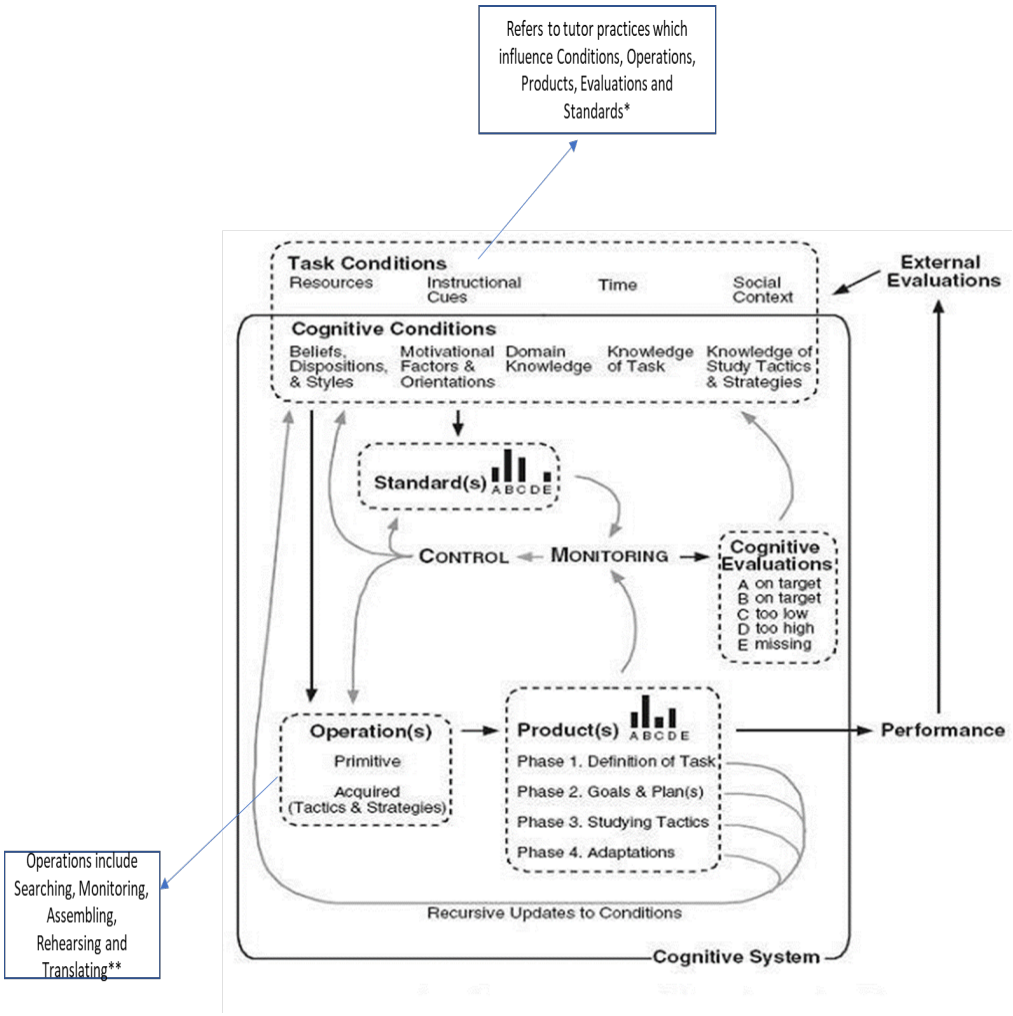


Figure 4. Winne and Hadwin's model of self-regulated learning. From "Studying as Self-Regulated Learning," by P. H. Winne and A. F. Hadwin, 1998, in D. J. Hacker, J. Dunlosky, and A. C. Graesser (Eds.), *Metacognition in Educational Theory and Practice* Routledge (https://doi.org/10.4324/9781410602350-19).

The 'C' in COPES refers to 'conditions'. In a similar manner to Zimmerman and Pintrich models, the Winne and Hadwin conceptualisation of SRL acknowledges the influence of context (learner and task conditions) on SRL. However, the model is unique in that it incorporates conditions in the structure of the SRL model, and examines the information processes which are shaped by an interaction of conditions with other components of the model.

The ‘O’ in COPES refers to ‘operations’. In contrast to the models conceptualised by Zimmerman or Pintrich, the Winne and Hadwin model examines how the information generated from conditions is processed by the learner through primitive abilities (memory, verbal), as well as acquired abilities (e.g. reasoning), with the output being the product of the SRL phases. The processing can be done actively by the learner, or a subconscious process (Winne, 2017a). Researchers have noted that by making the operations that a learner engages in explicit, the Winne and Hadwin (1998) model accommodates the learner moving back and forth within a specific phase, as the context of the task evolves (e.g., noted by Azevedo et al., 2011). For example, the learner may start a task with a certain understanding of the aims of the task, based on which the learner formulates a task definition (product of phase 1 of SRL). As additional information is provided on the purpose of the task (i.e. the task conditions are updated), the learner may process this information and update their own definition of the task.

The ‘P’ In COPES refers to ‘products’, or the output from the phases of SRL. There are 4 phases; task definition, goal setting, task performance and reflection. These phases are similar to the phases in the Zimmerman and Pintrich model, with the key difference being the addition of the ‘task definition’ phase.

The ‘E’ in COPES refers to ‘evaluation’, and ‘S’ refers to standards. The model proposes that the products of SRL phases are monitored and evaluated against internally held values, called standards. If there is no fit between the products and the standards, then the information is reprocessed, and/or learner conditions are adapted until there is a fit. For example, a learner may be given a maths problem to do, which they initially judge as being very difficult to do. This assessment of task difficulty is one of the standards for the first phase of SRL, task definition. As the learner starts the task, the learner will be given additional information on task difficulty, from learning resources, peers, tutors etc. They process this information to make further assessments on task difficulty, and if the learner

finds that the new information creates an assessment of task difficulty that does not match their initial assessment, then the learner may make adaptations to their own conditions (e.g. understanding of own ability, motivation, mindset), and reprocess information until there is a fit.

The key difference between the Winne and Hadwin model, versus the Zimmerman and Pintrich models, is the nature of the central role that monitoring and evaluation play. While monitoring and evaluation is seen as a key element of SRL in all three models (Greene & Azevedo, 2007; Hadwin & Winne, 2001; Pintrich, 2000; Zimmerman, 2000), the Winne and Hadwin model is distinct in that it considers the impact of monitoring and evaluation at a component level, rather than examining impact by phase of SRL (as noted by Azevedo et al., 2011). It provides a foundation for granular analysis of how changes on one component of SRL (e.g., contextual conditions) can lead to changes in other components (e.g., type of operations used by learner). Further, the Winne and Hadwin (1998) model allows SRL processes to be examined in a more fluid manner than other models. While all models are weakly sequenced, the Winne and Hadwin model provides a framework for analysing how a learner can move back and forth within a phase of SRL, as well as between phases (as noted by Greene & Azevedo, 2007).

Summary of the conceptualisation of models in second wave

In summary, the second wave of research built on previous work to represent SRL as a process, and explored the key role of context in the evolution of SRL processes. The models that emerged during this period explored the interaction between contextual factors and other variables such as cognitive and metacognitive behaviours to track the development of SRL products.

While a broad consensus emerged during this period on the key role of context in the evolution of SRL, models of SRL can still be distinguished by the specific role of context

within the architecture of the model. For example, the Zimmerman model structures a cyclical SRL process based on the phases of SRL: forethought/planning, performance control and self-reflection. Social and environmental factors influence the SRL processes, but remain outside the architecture of the SRL process. In contrast, the Winne and Hadwin model proposes a cognitive architecture which tracks the interaction of various components of SRL, including contextual factors, such as task conditions and learner conditions. This comprehensive architecture allows SRL to be analysed as a dynamic, evolving process which unfolds over the course of a learning activity and is rooted in context.

The models of SRL that emerged during this second wave can also be distinguished by the variables emphasised. For example, the Pintrich model emphasises the role of motivational variables in influencing SRL processes. It can be useful to adopt this model of SRL when exploring the influence of contextual factors such as social dynamics, on motivation and therefore on SRL processes. Alternatively, the conceptualisation of SRL proposed by Boekaerts (2002) emphasises the role that individual factors such as goal setting and goal mapping, play in SRL, and explores how these can be most effectively supported through the environment.

2.1.3 Third wave of research into self-regulated learning (2015-to date)

Research from 2015 onwards is often regarded as the third wave of research into SRL. In this wave, researchers explore how real-time measurement of SRL can be combined with intervention. The role of time and context have moved beyond being guiding factors, towards fundamentally determining how SRL is measured and supported. For example, researchers have developed software which records learners' actions as they plan, monitor and complete their work (Azevedo et al., 2011; Winne & Hadwin, 2013). Learners can review the traces collated by the software, which may prompt further monitoring and adaptations.

This latest period of research builds on work conducted in earlier phases. For example, Hadwin and Winne (2001) called for research into how instruction can drive the development of SRL. They recognised the need to examine the various sources of data that emerge from natural classroom environments to explore the impact of contextual factors, and consider how SRL can be promoted in real classroom settings. For this purpose, they developed an adaptive electronic textbook, called CoNoteS2. This aims to enable research into how students engage in SRL, to feedback into instructional design and textbook design. The platform provides an environment that includes implicit and explicit scaffolds for SRL. For example, the implicit scaffolds include cues for learners to organise their work, including planning out the task, and considering strategies for learning (e.g., note taking, different types of notes, and types of topics that can be addressed in the notes). The platform provides a list of objectives which learners can review, and plan their use of tools and strategies against, thus encouraging learners to set and refine goals over the duration of a task. The platform also allows learners to see how their studying tactics have evolved over time, thus supporting forward reaching adaptations. Explicit scaffolds include pre-labelled notes and indexes, encouraging specific forms of information processing. By collecting granular time-stamped data on learner activities, CoNoteS2 acts as both a measurement and intervention tool for SRL.

In 2013, Winne and Hadwin developed the Study platform, to enable the collection of time-stamped fine-grained data, which can be analysed to explore the evolution of SRL processes and changes in SRL in response to changes in context. The features of n-study include:

- Notes, Terms, Glossaries that learners are prompted to label by type, thus encouraging deeper information processing
- Tags, through which text that is being reviewed can be labelled with a type of tag

- Chat tools, with embedded prompts to scaffold learners as they chat. Learners can select a type of role from a drop-down list next to the chat (e.g. planner), and based on that, prompts will appear to help to scaffold the discussion from a co-regulatory perspective. For example, there may be prompts such as ‘how shall we approach this task? What should the next step be? ‘I think we may be a little off track’. Learners can build concept maps, libraries, and documents within n-study using hyper-media features, tags, and linking to notes, terms and glossaries

Study allows researchers to gather fine grained time stamped trace data, on learner actions, which are then organised into human level events. These events can be organised into a timeline, to show the order and context of activities. Trace data sequences can be analysed to identify transition matrixes and transition graphs, which show how learning strategies evolve, and the shapes of learning strategies at various points (e.g., linear, cyclical, branching), and in response to interventions.

Winne and Hadwin (2013) argue that there are lots of information sources on the internet, but much of it does not follow instructional guidelines, regarding organisation and layout of information or use of adjuncts. It can therefore be cognitively demanding to process this information. In addition, information can be unreliable, and make it difficult to monitor the effects of adaptations thus impeding self-regulation. Researchers have previously suggested that learners may not use learning tactics in new contexts, even when they are of benefit, because they are unable to manage the additional cognitive load of applying these tactics in addition to processing new information. This is particularly true when the new information is not laid out according to best practice in instructional design. Researcher designed platforms can thus have the advantage of laying out the information in a manner which reduces cognitive load, so that the impact of interventions on the use of learning tactics can be more sensitively measured.

2.1.4 The theoretical role of context in SRL: areas for further research

As described above, the theoretical role of context has evolved to become an integral component in SRL research, which fundamentally shapes how SRL is measured and scaffolded. 'Context' is a broad term and refers to a range of factors, such as the nature of the learning resources, social dynamics, environmental factors and domain (Butler & Cartier, 2017). While the evolution in theory to place greater importance on context is well established, models of SRL often require significant levels of interpretation to align them to the context being researched. There is a need for work which contextualises models of SRL to under-researched environments. For example, research on the influence of others on learner SRL in online environments is still in its early stages (e.g., exploring the influence of human tutors on learner SRL in a virtual classroom environment). It would be useful to interpret established models of SRL to investigate this context, drawing on related theories of learning where relevant.

Section summary

In this chapter, I have reviewed the three waves of research into SRL, and explored the evolution of its conceptualisation. In particular, I have highlighted how the conceptualisation of SRL has developed to both acknowledge its dynamic and process-oriented nature, as well as the integral role of context in influencing how SRL processes unfold. As part of my review, I have evaluated the differences between established models of SRL, as well as researcher designed platforms and measurement instruments. While the significant influence of context on SRL processes is now widely acknowledged, there is need for research which contextualises established models to under-researched contexts. For example, there is a need for theoretically grounded research into the influence of others on a learner's self-regulatory processes. Online learning environments such as virtual classroom environments can provide rich sources of data to enable this type of research.

Given that the influence of others on a learners SRL has been identified as an interesting area for research from a theoretical perspective, a literature review of the regulation of learning in collaborative settings is set out in the following sub-chapter. Following the detailed self-regulation learning, I identify specific areas for further research, and formulate research questions.

2.2 Regulation of learning in collaborative settings

Introduction

This chapter reviews the regulation of learning in collaborative settings. SRL has traditionally been explored in settings where a learner works alone. However, as learners increasingly work with others in online and hybrid settings, research is needed to explore the regulation of learning in social learning settings. Accordingly, researchers have explored new constructs of regulation in social learning settings to account for the influence of others on SRL.

These new constructs enable an exploration of the influence of the social context on learner SRL, and are therefore aligned with the conceptualisation of SRL as a process, which emerged in the second wave of research on SRL. Self-regulation, in settings where the learner works alone, is a construct which is primarily focused on self-regulation. It can be conceptualised either as an aptitude (as per the first wave of research on SRL), or as a process (as per the second wave of research), or both.

This chapter follows the structure set out below:

- Theoretical constructs which explore the regulation of learning in collaborative settings are reviewed
- Empirical research focusing on regulation of learning in collaborative settings is reviewed
- Gaps in the literature review, which require further research, are identified.

2.2.1 Constructs of regulation of learning in collaborative learning settings

Effective collaboration requires (1) individual self-regulation to advance the common goal (2) co-regulation to support the self-regulatory capabilities of others and enable them to make their unique contribution towards the common goal and (3) socially shared regulation, which consists of metacommunicative awareness and coordination of strategies to build consensus. All 3 regulation modes are central to collaborative learning occurring (Hadwin et al., 2017). In order to effectively collaborate, group members need to guide and support regulating as well as regulate collectively as a group (Järvelä et al., 2010).

The constructs of regulation in collaborative learning settings are outlined in greater detail below:

- Self-regulation consists of cognitive, metacognitive and affective information processes. It includes metacognitive monitoring and control over an individual's own cognitions, motivations, attributions and other thought processes (Pintrich, 2000; Winne & Hadwin, 1998). Control comprises fine grained adaptations made to cognitions within the task, and large grained adaptations made to the cognitive strategies, as well as cognitive and metacognitive information (e.g., beliefs about a task, self-efficacy). It is directed inwards and occurs at the 'I' level. SRL in collaborative learning settings manifests itself as learners being in charge of their individual contributions to the group, and contributing towards his/her own learning goals. Learners may or may not make SRL activity visible to other group members (Järvelä et al., 2016).
- Co-regulation is when one person supports another's self-regulatory processes, and occurs at the 'you' plane (Hadwin et al., 2011). It consists of one person modelling or prompting another person's cognitions, metacognitions or supporting their motivational/affective processes. Coregulated learning (CoRL) occurs when learners'

activities are influenced by other members in the group (Hadwin et al., 2011). Miller and Hadwin (2015) specified 2 forms of CoRL:

  - When group members prompt contributions from each other, e.g., learning goals
  - When an individual's SRL is gradually influenced by the activities of the wider group, e.g., when individuals adopt the learning goals of the wider group

- Socially shared regulation is when individuals are fully collaborating, and coordinating self-regulatory processes to build consensus on cognitive, metacognitive and affective information (Hadwin et al., 2011). It occurs at the 'us' or 'we' plane. Socially shared regulation is when the group negotiates shared goals and perceptions, and then coordinates group wide strategies to enact, including collective monitoring and adaptation. For example, reciprocal negotiation of a group goal that can be adopted by all group members without a conflict with personal preferences. The main distinction between CoRl and SSRL is that SSRL does not mean that all group members are the 'same', but rather that they actively contribute to the discussion, and negotiate the same goals and approach (Järvelä & Hadwin, 2013). In the context of collaboration, monitoring has a social aspect, and includes monitoring collaborative progress and ensuring peers understanding of context (Rogat & Linnenbrink-Garcia, 2011).

Collaborative tasks exist on a spectrum of solo tasks done in cooperation with each other, which typically trigger forms of co-regulation, in which metacognition may not be shared through to fully collaborative tasks which require individuals to complete tasks together and share metacognition, as in Hadwin et al.'s (2010) study. Often tasks which

trigger full collaboration are more complex and exist in the individuals Zone of Proximal Development (Vygotsky, 1978).

Malmberg et al. (2017) developed a framework to compare the constructs against, aligned to the three main questions within the regulation of tasks: (1) what is being monitored, (2) what goals/standards is it being monitored against, and (3) what is being controlled/regulated as a consequence. This framework is presented in the table 2 below. It is worth noting that the different constructs of regulation operate on different types of information. Self-regulation is primarily focused on information about the self. On the other hand, co-regulation and socially shared regulation include information types held by the self, about themselves and also about others. For example, co-regulation includes information about other's task knowledge, and strategies, but this is monitored against standards set by the individual for others. In addition, new forms of regulatory processes and strategies may be relevant for the different constructs. For example, new types of metacognitive events may emerge in socially shared regulation, such as metacognition that is visible but not shared, or joint activity leading to joint metacognition and regulation.

Table 2. Constructs of regulation in collaborative learning settings

	What is being monitored	What goals/standards it is being monitored against	What is being controlled/regulated as a consequence?
Self-regulation	An individual monitoring their own perceptions of a task, their own goals, the strategies they are familiar with, their affective state, and their beliefs/opinions	Standards set by an individual for themselves	Individual exerting control over their own state and actions, and environment
Co-regulation	An individual monitoring the task perceptions, goals and affective/cognitive state of others, the	Standards set by an individual for others	Individual exerting control over the state and actions of others

	strategies that others know of, and how their actions/interactions can influence the actions of others.		
Socially shared regulation	An individual monitoring the common or negotiated task perception, the negotiated team goal, the alignment between the individual goals and the team goal, the strengths and weaknesses of the team as a whole, strategies for collaborative working, and ways actions/interactions can influence the progress of the team as a whole.	Jointly agreed standards	Individuals exerting control over the collaboration of the group as a whole

2.2.2 Theory based frameworks for co-regulation

Following the outlining of constructs of regulation in collaborative learning settings, I explored the literature to identify theory-based frameworks, specifically with reference to co-regulation. The aims of the review were to:

1. Identify relevant studies. Studies were considered as relevant if the research question addressed co-regulation by a more knowledgeable 'other', either as a primary research focus, or as a sub-topic, and regardless of whether the learning environment was online or offline. A broad definition of 'relevance' was adopted given the likelihood of limited research having been conducted into co-regulation in online learning environments. Papers which explored how tutoring or teaching practices could

influence learner SRL were therefore regarded as relevant to developing my theoretical understanding of co-regulation in a VCE

2. Identify the frameworks for analysis used in the relevant studies.
3. Identify the theoretical basis of the frameworks used, if any.

I followed the below steps to carry out my exploratory literature review:

- I searched the UCL databases, and results were filtered to exclude any papers that had not been peer reviewed. The UCL databases can be accessed through UCL library services, and are indexed through the 'Explore' function. The UCL databases were selected as a source for my research due to their coverage of a diverse range of publications. Examples of relevant databases which can be accessed through UCL include:
    a. Education Resource Information Center (ERIC): Provides access to educational literature and resources, including information on technology in education
    b. Digital Education Resource Archive: A digital archive of documents published electronically by government and related bodies in the area of education, which may include resources on education technology
    c. Education Database (ProQuest): Gives access to over 900 educational publications, some of which likely cover education technology topics.
    d. Other relevant databases which provide access to a range of multi-disciplinary research include JISC Journal archives, SAGE and SCOPUS.
- I further evaluated the relevance of the paper by reading the paper abstracts.
- Papers that were judged to be relevant to my research question were read in depth, and the quality of the paper was assessed based on the robustness of the research design, methods of data analysis, validation of results, and depth and breadth of the

bibliography. Any papers that were judged to be of insufficient quality were excluded from the review.

- I reviewed the bibliographies of the selected papers for additional references which appeared relevant. These papers then passed though review steps (3) and (4) as stated above.

Table 3 below presents papers judged as relevant, setting out the research context, types of data collected and theoretical approach adopted. As demonstrated by table 3, the majority of studies conducted on co-regulation have been in the context of an offline, traditional classroom environment. With regard to scaffolding by online tutors, work has explored the influence of pedagogical agents in scaffolding learner SRL, but limited work has been done in the context of human-to-human tutoring.

Table 3. Studies on co-regulation of learners, mapped to research setting, types of data collected and analytical frameworks used

Research setting	Relevant studies	Types of data collected and analysed	Theoretical approach adopted
Traditional classroom environment	(Braund et al., 2021)	Qualitative data is collected from semi structured interviews with kindergarten teachers and running observations of kindergarten classrooms	The data is coded on an inductive thematic basis, with reference to the relevant literature
	(Andrade & Brookhart, 2020)	Literature review of studies in Classroom Assessment	The Pintrich and Zusho (2002) model of SRL is adopted and expanded to consider the ways on which Classroom Assessment may co-regulate learning. A literature review is conducted and mapped to the theoretical framework
	(Allal, 2020)	Classroom Assessment	Proposed a new model of co-regulation
	(Saariaho et al., 2019)	Video data from student teacher practicums	This study codes video data to observe how student teachers co-regulate students, using both verbal and non-verbal data. There is no explicit theoretical framework adopted, but the

			study uses categories of co-regulation based on well-known components of SRL
	(Dignath & Büttner, 2018; Ewijk et al., 2013)	Video observations of primary and secondary school mathematics classrooms, and interviews with primary and secondary school mathematics teachers	The ATES (Assessing How Teachers Enhance Self-Regulated Learning) framework is based on Boekaerts's (2002) model of SRL. It is used to code observation data collected from a traditional maths classroom, to identify direct and implicit teaching practices which boost learner SRL, as well as development of a learning environment that promotes SRL
	(Bell & Pape, 2014)	Observations of 7th grade mathematics students in a classroom environment	The paper examines how teachers scaffold self-regulatory development through classroom discourse, A theoretical framework and strategy observation tool is developed, based on Schunk and Zimmerman (2000) model
	(Meyer & Turner, 2002)	Discourse analysis of primary school mathematics classroom	Coding scheme for instructional and non-instructional scaffolding based on theories of scaffolding in the classroom
	(Perry et al., 2002)	Observations of teacher interactions with the whole class for primary school	The output from observations and teacher interviews was analysed in light of SRL research, but no theoretical framework was used
Intelligent tutoring system Intelligent tutoring system	(Graesser & McNamara, 2010)	Computer tutors with pedagogical agents that converse with learners in natural language. Data collected and analysed includes learner actions and discourse moves	Relevant literature explored. No new framework proposed
	(Trevors et al., 2014)	Adaptive hyper media learning environment, with embedded pedagogical agents. Data collected and analysed	This study explores the influence of pedagogical agents on the quantity and quality of notetaking by learners. The results from the study are interpreted in light

		includes agent prompts, and learner actions	of the model of SRL set out in Winne and Hadwin's (1998) model of SRL

The majority of the studies reviewed either didn't apply a theoretical framework for analysis, (preferring to explore findings in light of the relevant literature), or developed theoretical frameworks for the context of a traditional classroom environment. In the latter case, frameworks have been tailored to capture macro events and processes, rather than fine-grained data about teacher-learner interactions. For example Ewijk et al. (2013) used an observation instrument (ATES) based on (Boekaerts, 2002). According to this framework, teaching practices may be coded as follows:

- Direct instruction on SRL – where the teacher explicitly advised the student that this was a SRL strategy and the value of it
- Implicit instruction of SRL – where the teacher prompted or modelled the strategy, or provided feedback on the implementation of a strategy
- Development of a learning environment conducive to SRL – constructivist, collaborative, situated in authentic learning

The framework applied in this study indicated whether direct or implicit instruction of SRL had occurred, and if the teacher had developed a learning environment conducive to SRL. However, the specific types of teaching practices utilised (for example, whether the teacher used modelling, prompts or hints) were not coded. A more detailed framework can be developed and applied in online learning settings, where it is possible to collect finer grained data. The greater detail would drive an understanding of the influence of different types of teaching practices on the SRL processes demonstrated by learners. Similarly, Saariaho et al. (2019) coded video data of classrooms, by extending established macro-components of SRL for co-regulation; teaching practices were coded as either being co-planning, co-strategy use, and co-reflection, with detailed examples of each of these practices provided to facilitate

coding. Again, the codes themselves were limited to identifying the function of practice, rather than also signifying the specific practice itself.

Researchers are beginning to recognise the need for theoretical models and frameworks which enable the concept of co-regulation to be operationalised and explored in practice. To this end, there have been some recent studies which have proposed more granular frameworks of co-regulation. For example, Andrade and Brookhart (2020) developed a theoretical framework to assess co-regulation through Classroom Assessment (CA), and conducted an exercise mapping the literature to each element of the framework. They began by broadly defining CA to include assessments conducted in class, as well as assessment through student-teacher dialogue. The Pintrich and Zusho (2002) model of SRL was adopted and expanded to consider the potential ways in which CA may co-regulate learning. Andrade and Brookhart (2020) chose to adopt the Pintrich and Zusho (2002) model as the basis for their work, as it takes a comprehensive approach (phases of SRL), and also considers SRL by area (cognition, motivation, behaviour, context). This would enable the influence of CA on SRL to be considered from a range of perspectives. Previous studies were then reviewed and mapped to each part of the framework to validate the theoretical approach. Examples of the literature mapped to the co-regulation phases in the cognition area are provided below, (focusing on teaching practices as a form of CA):

1. For **goal setting** in the **cognition** area, Andrade and Brookhart (2020) identify two categories of studies which may be relevant:
    a) The influence of teacher expectations on CA and student learning outcomes
    b) The influence of teachers on the goals that students set for themselves, by helping them formulate success criteria
2. Similarly, Andrade and Brookhart (2020) propose that CA can be regarded as co-regulating learners in the **monitoring** phase of **cognition** by providing feedback on

progress. They identify research which has investigated the relationship between CA and feedback (e.g., Hattie, 2008a).

3. In the **control** phase of the **cognition** area, CA can be regarded as co-regulating learners by considering the use of informal assessment information by teachers to individualise instruction to learners, and guide the next steps in learning that they may take (Heritage & Heritage, 2013).

2.2.3 Empirical research on regulation of learning in collaborative learning settings

In the sections below, I have reviewed research on the regulation of learning in collaborative learning settings, including traditional classroom environments, human tutors, pedagogical agents and small peer groups/dyads. I have analysed the research findings in relation to the various types of regulation that occur in collaborative learning settings, and considered areas for further research.

Traditional classroom environments

Researchers began to increasingly explore the role of teachers in co-regulating learners, as socio-cultural approaches emphasising the shared responsibility between teachers and students for building knowledge gained prominence (McCaslin & Good, 1996; McCaslin & Hickey, 2001). A fundamental assumption emerged that self-regulation is attained through social interactions, and SRL outputs should be perceived in context (McCaslin & Hickey, 2001). This assumption aligns with a shift in perceiving SRL as a social process rather than as an individual process.

A significant chunk of research for co-regulation in classrooms has been conducted in offline environments, using observation tools. Observation tools have enabled researchers to investigate the influence of teaching practices in the classroom, for example by exploring how teachers use different types of strategies (cognitive, versus affective versus metacognitive). However, the types of data that can be collected by observation tools can be

relatively large grained (for example, researchers may be able to collect data on the type of strategy employed but would be unable to specify the strategy itself). In addition, most studies have been conducted in traditional settings where the teacher is working with a group of students, and the focus tends to be on the actions of the teacher or the environment they have designed; studies are not able to monitor student actions and responses to teacher practices. Further, while observation tools have allowed researchers to explore how teachers co-regulate leaners, they can be difficult to scale.

An example of a study examining how teachers co-regulate learners in a classroom environment is the work conducted by Meyer and Turner (2002). They coded video and audio observation data for primary and secondary mathematics classrooms, using a framework based on the Boekaerts (2002) model of SRL. Teaching practices were coded, based on the types of SRL strategies encouraged (spanning cognitive, metacognitive and motivational strategies). An additional layer of coding was used to indicate whether the practice related to direct instruction, implicit instruction or if it was relevant to the learning environment. Direct instruction on SRL referred to practices where the teacher explicitly advised the student that this was a SRL strategy and the value of it. Implicit instruction of SRL indicated that the teacher prompted or modelled the strategy, or provided feedback on the implementation of a strategy. Finally, development of a learning environment conducive to SRL referred to teaching practices aimed at promoting an environment which is constructivist, collaborative and situated in authentic learning.

Meyer and Turner (2002) found that direct instruction of SRL was very rare amongst the sample studied, and that teachers found it difficult to explain to researchers what was meant by cognitive, motivational, and metacognitive strategies. In the context of implicit instruction, the observation data demonstrated that primary teachers relied on motivational strategies significantly more than secondary teachers, and secondary teachers used

metacognitive strategies significantly more than primary teachers. With regard to the design of the learning environment, primary teachers were more likely to create opportunities for their students to engage in SRL through constructivist learning environments (open ended activities, peer discussions, real life problems) with a good balance of teacher led versus student led activities, while secondary teachers developed environments that skewed towards being teacher-led.

Meyer and Turner (2002) argued that teachers need to show a combination of direct and implicit instruction to have a lasting and significant effect on learner SRL skills; this was proven to be rare based on the results from the study, indicating a need for teacher training on co-regulation of learner SRL. The potential requirement for training teachers on SRL strategies was further supported by interviews of secondary school teachers, which demonstrated that teachers over-estimated the extent to which they instructed learners on SRL strategies. The findings from this study align with findings from a meta-analysis of the effectiveness of training programs to foster SRL by Dignath and Büttner (2018). This previous work found that self-regulation training was most successful when led by researchers, rather than the students' regular teachers.

A similar approach towards exploring how teachers can co-regulate learners in the classroom was proposed by Ewijk et al. (2013), who developed the ATES instrument. This instrument measures the promotion of SRL through direct and indirect means i.e., considering instruction of SRL, and the design of the learning environment. The researchers argued that teacher promotion of SRL measured by teacher ratings can be subject to bias, or misunderstandings of terminology by the teacher. Accordingly, the ATES instrument was developed to code the instruction of learning strategies and design of learning environment based on observed data. With regard to learning strategies, three dimensions were coded: the frequency of whether instruction took place or not, whether it was implicit or explicit, and

strategy type (cognitive/metacognitive/motivational). The specific strategy itself (e.g., note-taking, or questioning) was not coded, and as the coding was relatively large grained, it was termed as being low inference coding. In addition to coding dimensions of learning strategies, the design of the learning environment was also coded. Environments which enabled authentic problem situation, and collaborative learning in which students could practice their SRL skills, were coded as promoting SRL. The researchers argued that students can only develop SRL skills if they are given opportunities to learn them (through direct explicit and implicit instruction), and practice them (through indirectly promoting SRL through the nature of the learning environment).

As well as evidencing the need to train teachers on how to scaffold SRL through direct and indirect practices, studies have also demonstrated the value of co-regulation by teachers. For example, teacher co-regulation can encourage a more diverse range of sophisticated forms of regulation, benefiting student learning. Walker et al. (2004) conducted a study in which primary school students (grade 5) worked on a joint society/environment project in an online environment scaffolded by teachers. They found that the intervention group was advantaged in its planning and regulatory activities, as well as the quality of the final product and ability to use ICT. Observations of the products and discourse provided evidence of socially shared regulation as well as co-regulation. As well as encouraging complex forms of regulation, effective teacher co-regulation may also limit help seeking and improve learning outcomes. For example, Makitalo-Siegl et al. (2011) examined how high and low levels of teacher scaffolding for secondary students affected help seeking and learning behaviours, and found that students in the 'high' group asked for help less frequently, and demonstrated better understanding of the material. This indicates that support enabled the students to regulate themselves more effectively, including monitoring where their understanding was weak, and implementing effective plans to remedy this.

While the above studies highlighted the importance of teaching practices in the development of learner SRL, they were limited in scale by the use of observation data. The emergence of online trace data offers the opportunity for new frameworks to be built, which can examine the influence of instructors on learner SRL at scale. Researchers have also identified the importance of coding instructor-learner interactions from both the perspective of the teacher as well as the learner (Meyer & Turner, 2002). This will provide a holistic and temporal view of SRL development, driving deeper insight into how teachers can influence learner SRL.

Human tutors

One to one tutoring is noted as being suited to the promotion of SRL as tutors can track student knowledge and SRL capability more accurately, and adapt accordingly. However, research exploring how human tutors co-regulate learners is limited, potentially due to expense associated with using human tutors on research projects (Graesser & McNamara, 2010). Further, the past research that has been conducted has mainly focused on offline environments. There is a strong need for additional research into how human tutors scaffold learner SRL in online environments. This is particularly true, given the expansion in human tutoring following the Covid-19 pandemic (Cullinane & Montacute, 2023).

One of the challenges posed is determining how to conceptualise tutor performance, prior to being able to track what 'good' tutoring practices look like. Researchers have employed several approaches to conceptualise the performance of human tutors. In some studies. researchers have considered the learning gains attained by learners through human tutoring to other instructional methods, such as computer-based tutoring systems, classroom instruction, or no tutoring (e.g., reading materials alone). For example, VanLehn's (2011) meta-analysis compared the effect sizes of human tutoring, intelligent tutoring systems, and

no tutoring, finding that human tutoring had a lower effect size (d = 0.79) than commonly believed, but still higher than no tutoring.

In other studies, researchers have used human evaluator rankings as a proxy for tutor performance. For example, Cukurova et al. (2022) explored the relationship between tutoring practices and tutor performance in a Virtual Learning Environment. Tutor performance was determined through a weighted measure developed using human evaluator rankings for tutors on a range of dimensions. Cukurova et al. (2022) developed a tagging tool to observe human tutor-learner sessions online, and code tutoring behaviours for high and low-performing tutors. The resulting dataset was analysed for statistically significant patterns of behaviour. The researchers found that there were statistically significant differences in the patterns of behaviour exhibited by human tutors who had high performance rankings (from human evaluators), versus human tutors who had low rankings

In addition, past research has considered the nature of the learning environment developed by the tutor, as well as the tutoring practices employed by them. Chi et al. (2008) have distinguished between tutor centred, student centred and interaction centred tutoring environments:

- Tutor centred environment refers to where the tutor lectures to the student, who passively receives the information, even if the student does not understand it
- Student centred environment refers to the student taking control by choosing work, planning, asking questions, self-evaluating. While this is an environment which is SRL rich, student learning outcomes may be impacted by the student inaccurately calibrating their learning or not asking enough questions
- Interaction centred tutoring (Chi et al., 2008; VanLehn et al., 2007) refers to the dialogue interactions providing optimal levels of scaffolding, with the student being

actively engaged and being given a significant amount of control, while the tutor still ensures that learning activities are productive

Besides Interaction centred tutoring, other intermediate approaches include:

- Just in time Instruction (Schwartz & Bransford, 1998). This is an intermediate position which targets when the student is open for instruction (e.g., after they have failed at a problem)
- Open learning environments (Bull & Kay, 2007) where the teacher and student both have open access to the knowledge and skills the student has mastered and the landscape to be mastered, and the student can select the new topics/skills for mastery

Past research has found that human tutors struggle to promote sophisticated forms of SRL, instead driving their own agenda through the conversation (e.g. Graesser & McNamara, 2010). For example, Graesser et al. (1995) analysed video data from 100 hr of unskilled tutors and classified the utterances into discourse patterns and analysed patterns of discourse. Cade et al. (2008) did a similar analysis for outstanding tutors in middle and high school. Both studies found that tutors did not often implement pedagogical techniques such as Socratic tutoring techniques, modelling-scaffolding-fading, reciprocal teaching, building on pre-requisites, or diagnosis/remediation of misconceptions. Tutors often corrected errors straight away rather than waiting for the student to self-correct.

However, research has found that human tutors do encourage students to talk and engage rather than merely lecturing to them. They often follow an expectation and misconception tailored dialogue, which refers to a list of anticipated good answers and misconceptions associated with each problem, and use 3 dialogue moves to elicit information from learners – pumps (‘what else’ ‘keep going’), hints and prompts for the student to fill in important words (with hints and prompts usually being expressed as questions; Graesser & McNamara, 2010).

Further work is needed to explore how human tutors co-regulate learners in online environments. This work could be used to inform the design of online environments to best support tutors, as they scaffold learners, and potentially inform the development of adaptive tools. Online environments, such as virtual classroom environments, also tend to control the design of the learning activity and environment; another interesting research question would be to explore the extent to which tutor practices vary despite the controlled online environment.

Pedagogical agents

Intelligent tutoring systems may use pedagogical agents to scaffold learner SRL through social interactions. Graesser and McNamara (2010) argue that pedagogical agents can more consistently promote SRL versus human tutors, as it would be very hard to train a human tutor on a conversational strategy which conflicts with their natural style. Further, pedagogical agents can accurately track dialogue moves and compute the optimal next steps. While these benefits of pedagogical agents versus human tutors make logical sense, I am not aware of any studies exploring the differences between the two types of tutors. Further work is needed on human tutors coregulating learners, as well as studies comparing human tutors with pedagogical agents, to explore the differences between the forms of co-regulation.

Pedagogical agents take on a range of forms including acting as mentors, peers, tutors or avatars. They help students learn SRL skills by modelling them or by holding an interactive conversation. They may take a direct approach towards co-regulation where the agent train students on metacognitive and SRL skills, tracking and measuring progress. e.g., Meta tutor. Alternatively, they may take an indirect approach, such as the approach adopted by Auto-tutor; this consists of discourse patterns modelling or encouraging learner SRL.

I have below provided examples of ITS with pedagogical agents that co-regulate learners through a direct approach, describing the agent, the strategies they provide training on, and whether the training is direct or indirect.

Table 4. Examples of ITS with pedagogical agents

System	Type of SRL training	Pedagogical agent	Learning strategies
Meta-tutor (Azevedo et al., 2009)	Direct	There is a main agent Gavin that controls the overall learning environment, and three satellite agents which control the phases of SRL, planning, monitoring and applying learning strategies. Each phase can be further broken down under the guidance of the agent. For example, monitoring can be broken down into judgements of learning, feeling of knowing, content evaluation, monitoring the effectiveness of a learning strategy, and monitoring progress towards a goal.	Examples of learning strategies are re-reading, taking notes, searching in a goal-oriented manner, drawing diagrams or tables, making inferences, coordinating information from various sources.
iStart (Interactive Strategy Trainer for Active Reading and Thinking) (McNamara et al., 2004)	Direct	Agents model reading strategies. In the 'Practice' module, students type self-explanations into the system, for which feedback is provided by the agent, and requests	It focuses on 5 reading strategies; monitoring comprehension, paraphrasing explicit text, making bridging inferences between one sentence and another, making predictions, and elaborating text.

		made for the student to modify inadequate answers.	
SEEK (Source, Evidence, Explanation and Knowledge) (Graesser et al., 2007)			Aims to improve critical thinking skills for college students, using a 'hint' button which has suggestions for how to guide google search, a 'pop up ratings and justifications' which prompts students to assess the reliability of the information, and finally a pop-up journal, which asks 5 questions about the reliability of the site the learner visited.
iDRIVE (Instruction with Deep-level Reasoning questions in Vicarious Environments) (Gholson & Craig, 2006)	Direct	Dyads of agent	iDrive has dyads of agents modelling deep reasoning questions in question-answer dialogues.
Auto-tutor (Graesser et al., 1999)	Indirect	Conversational agent	This system encourages SRL skills and strategies using conversational dialogue. It is organised around difficult problems that need reasoning to work out. Questions require the learner to articulate 3-7 sentences (expectations), and coverage of expectations usually requires 30-100 conversational turns. It simulates human tutor-learner dialogue by generating pumps, hints, prompts and feedback, eventually using assertions if required.
Operation ARIES (Acquiring Research Investigative and Evaluative Skills) (Wallace et al., 2009)	Indirect	Conversational agents	Learners interact with two pedagogical agents to develop their critical thinking skills. Guide agent acts as the tutor with expertise in the subject, while the other agent takes on the role of a non-expert peer (e.g., a fellow student). The three-way dialogue is sensitive to the learner performance – for example, if the learner is performing poorly, then the 2 agents will interact with each other to model learning. If the student is performing very well, the guide agent

			will ask the learner to teach the non-expert agent, supported by the tutor agent. When learner performance is intermediate, there will be a trialog structured in line with Auto-tutor.

As stated above, very limited work has been done in the field of human tutors working in online environments. In this thesis, some lessons learned, from the field of pedagogical agents (regarding the identification of signifiers of SRL) are explored to be potentially used to monitor how human tutors co-regulate learners in online environments.

Dyads and small peer groups

With the expansion of online and blended learning, it has become increasingly common for peers to work together online in small groups or dyads. These learning settings offer the opportunity to study the different forms of regulation, both as stand-alone constructs, and with regard to the transitions between them.

Research on transitions between different forms of regulations in small group settings is still emerging. For example, Hurme et al. (2009) explored the relationship between SRL and SSRL in an asynchronous learning environment in which students collaboratively solved problems. Text messages were coded for metacognitive activity which either (1) referred to earlier activity, (2) interrupted, changed or proposed an alternative to joint problem solving activity, and (3) provided explicit reasoning for considering an alternative route to problem solving.

The researchers showed that metacognition was shared when an individual was trying to control or regulate the group planning through presenting the results of their individual metacognition. The findings from this study contrast with the study by Hadwin et al. (2010). This study looked at collaborative work by graduate students using n-study software, particularly focusing on task perceptions and goals. A co-constructed statement was

submitted by the groups at the beginning of the task and compared with individual reflective statements submitted at the end of the task. In contradiction to the earlier study completed by Hurme et al. (2009), it was found that shared metacognition was rarely attained and did not evolve over the course of the task. Further research is needed to explore the relationship between the different forms of regulation in small group settings, and resolve contradictions.

Malmberg et al. (2017) conducted a study which contributed to our understanding of the temporal aspects of regulation of learning in small group settings. They used state lag sequential analysis to explore the most prominent transitions between types and processes of regulated learning for a small group of peers. Rather than using every event which signals regulated learning, this study took an approach which identified only the most significant types and processes of regulated learning that occur in the context of collaborative learning. The researchers found that:

- Co-regulation played an important role by prompting the interaction and alignment essential for collaboration, and encouraging participation in constructing joint knowledge and activating SRL. This is helpful, especially in the early phases of collaboration. Co-regulation also seemed to provide the potential grounding for SSRL.
- SSRL usually occurred in the intermediate phase, indicating that on-going interactions, and task execution were necessary before SSRL planning could be initiated. SSRL monitoring was linked to co-regulatory planning, indicating that SSRL monitoring builds collaborative knowledge which is then the basis for co-regulation.
- SRL planning was significantly greater in the intermediate phase versus the starting phase. Overall, SRL processes occurred less often than co-regulatory or SSRL

processes. This is possibly because SRL is less visible than other forms of regulation, and also because SRL processes are partly embedded in SSRL and CoRL.

An important finding from this study was an enhanced understanding of the role of SSRL monitoring in small group settings. Shared monitoring is very important in reducing the gap between the current state and the ideal state, in the context of collaborative learning (Molenaar & Chiu, 2014). It can be regarded as being beyond a trigger for adaptation and control of SRL processes, and is rather a process itself which builds knowledge that can be used to activate other forms of SRL. Further research is needed to explore the role of monitoring and how it relates to other regulated learning processes, particularly as understanding the role of monitoring further could develop theoretical understanding of how and when small scale versus large-scale adaptations is made (Hadwin et al., 2017).

Research has also indicated the need for scaffolding for effective regulatory processes to emerge in small group settings. For example, Hurme and Järvelä (2005) found that dyads of 13-year-old students needed support to develop effective processes to regulate their learning, while they collaboratively worked on maths problems. Students used the chat function to discuss, propose strategies and suggest alternatives, and were instructed to make discussions visible via the chat. The researchers found that students demonstrated metacognitive knowledge, but less metacognitive monitoring, either because content was not challenging enough or because they needed prompting from the teacher or software. Authors characterised the chat as 'every day talk' rather than mathematical reasoning, which aligns with Zimmerman (2008) research that students need support to develop effective self-regulation. Further work is needed to understand the types of scaffolding that would be useful to help students develop SRL and SSRL skills.

2.2.4 Empirical research on regulation of learning in collaborative learning settings: areas for further research

There is a need to build granular frameworks to investigate co-regulation of learners in online learning settings. Chan (2012) noted coding schemes used for co-regulation tend to be based on models of SRL, and are brief. Further work is needed to develop more granular schemes which can be operationalised in online settings, and help develop our understanding of this construct. A more extensive theoretical understanding of regulated learning processes is needed along with empirical evidence, to support teachers and learners in identifying strengths and weaknesses of regulated learning processes. Instructional strategies (Paris & Paris, 2001) or pedagogical scripts could be built using these. Winne (2015) proposed using process discovery as a tool to visualise the process of collaborative knowledge construction but this relies on more extensive theoretical frameworks of regulated learning processes.

Further empirical work is needed to explore how teachers co-regulate learners, and the specific types of co-regulatory processes used. For example, Malmberg et al. (2017) identified the co-regulated learning processes as being planning, monitoring, strategy use, but did not distinguish between the nature of the co-regulatory processes – for example, is planning or monitoring co-regulated through directive statements, or through open-ended prompts?

A deeper investigation into the nature of co-regulation would be merited in studies which explore how teachers co-regulate learners, and the teacher is working to transfer responsibility for regulation towards the student. The Malmberg et al. (2017) study explored regulation in a group of peers, and was therefore focused on the transitions between co-regulation and SSRL, rather than exploring co-regulatory processes, and optimal ways of scaffolding SRL.

Another question that requires investigation is how the design of the online learning environment can influence the scaffolding of SSRL. Synchronous learning environments appear to offer an ideal learning environment for the development of shared metacognition

through the use of tools, and interactive whiteboards (Carroll et al., 2003). However, more research needed to explore how teachers leverage the features of online environments to co-regulate learners and build SRL. In particular, it would be interesting to understand the variations in how teachers are influenced by online environments when co-regulating learners.

Addressing the gaps in research that are described above is of value, as it will inform future research on software that imitates or builds on effective ways of scaffolding SRL.

Summary

In this chapter, I have reviewed the main theoretical constructs employed to describe the regulation of learning in collaborative settings. The theoretical constructs include self-regulation, co-regulation and socially shared regulation. I have identified the differences in the types of information processed for each construct, and what is regulated as a consequence. For example, I have described how co-regulation occurs at the 'you' plane, regulating the actions and behaviours of the other learner, while socially shared regulation occurs at the 'us' plane, regulating the actions and behaviours of the group.

I have followed my theoretical review by considering the body of empirical research conducted on regulation of learning in collaborative settings, and highlighting gaps for additional research. Studies have been conducted on the regulation of learning in a variety of settings, including traditional classroom environments, human tutors, pedagogical agents and small peer groups/dyads. However, this area remains broadly under-researched, and there are multiple opportunities for additional work. For example, while research has been conducted on the co-regulation of learners in offline environments (e.g., tutoring or traditional classroom environments), the research has relied on large grained observation data. In these environments, the focus tends to be on the actions of the teacher or the environment they have designed; studies are not able to monitor student actions and responses to teacher

practices. The development of online classrooms offers new study contexts, which could potentially utilise fine grained data, to develop a richer understanding of how teachers co-regulate learners.

A number of gaps have been identified for further exploration. For example, there is a need to build granular frameworks to investigate co-regulation of learners in online learning settings. A more extensive theoretical understanding of regulated learning processes is needed along with empirical evidence, to develop an understanding of how co-regulation occurs in practice. In addition, further empirical work can explore the strengths and weaknesses of specific types of co-regulatory processes employed. For example, is planning/monitoring more effectively co-regulated through directive statements, modelling or open-ended prompts? Another question that requires investigation is how the design of the online learning environment can influence the scaffolding of SRL. For example, how does the design of the online learning environment impact teaching practices, and how does this vary by teacher type? The gaps identified in this literature review inform the formulation of my research questions, set out in Section 2.5. However, prior to formulating the research questions, I explore the literature on the measurement of the regulation of learning. This includes consideration of the data types and signifiers used (Section 2.3), and the application of learning analytics to measure SRL (Section 2.4).

2.3 Measurement of the regulation of learning: Data types and signifiers

Introduction

The measurement of SRL has long been a topic of research interest for several reasons (see, for example, Azevedo et al., 2010; Bannert et al., 2017; Panadero et al., 2016; Winne, 2010). Firstly, an evaluation of SRL in various contexts enables the testing of assumptions in theoretical models, and the further development of these models. Secondly, the measurement of SRL is methodologically challenging, because SRL is an internal process; exploration of

these methodological challenges offers rich avenues for research. Finally, a better understanding of the development of SRL in various contexts enables the operational development of effective scaffolds to support SRL.

In this chapter section, I review the development of research into the measurement of SRL. This chapter section deals with the evolution of the types of data used to measure SRL, and is followed by another chapter section (Section 2.4) considering the analytical techniques and tools used.

This chapter subsection follows the structure below:

- The evolution of the types of data used to measure SRL, aligned to the waves of research into SRL (as defined by Panadero et al., 2016)
- Gaps identified through the literature review i.e., the study contexts and data types which are under-researched, with respect to the measurement of SRL

Evolution of the types of data used to measure SRL

The types of data used to measure SRL need to be aligned to the conceptualisation of SRL used by the researcher (Panadero et al., 2016). The evolution across theoretical models of SRL have therefore influenced the types of data that researchers gather to measure SRL. I have reviewed below the implications of theoretical developments in SRL on the types of data required to measure SRL, in each of the waves of research (as defined by Panadero et al., 2016).

2.3.1 First wave of research

During the first wave of research (1980's and early 1990's), SRL was conceptualised as an aptitude, and it was considered to remain in the same state over the course of a learning activity. The components of SRL included cognitive, metacognitive, and motivational, and

the learner's aptitude for each of these components was measured outside of the duration of the learning activity. Richard Snow (1991) defined a learner aptitude for SRL as follows:

"*. . . individuals differ in their readiness to profit from a particular treatment at a particular time; aptitude constructs are theoretical concepts fashioned to interpret these observed differences in person-situation interaction terms . . . aptitude is not limited to intelligence or some fixed list of differential abilities but includes personality and motivational differences along with styles, attitudes, and beliefs"* (p. 205)

During the first wave of research, learner aptitude for SRL was primarily measured using self-report data, generated from inventories or think aloud protocols. Inventories ask learners to imagine a learning event, and to estimate the frequency, intensity, likelihood, difficult or capability of performing a process or generating a product (Winne, 2010). Context may be described at various levels of granularity in the instruction. Learner responses may be constrained by a Likert (1932) scale or open-ended in oral protocols such as interviews. Inventories may be domain specific, such as the Motivated Strategies for Learning Questionnaire (MSLQ) developed by Pintrich et al. (1993). Alternatively, schedules may be domain generic (e.g. Learning and Study Strategies Inventory (LASSI) questionnaire developed by Weinstein et al., 1987).

Another source of self-report data is thinking aloud protocols. Think aloud protocols refer to the learner vocalising their thoughts during the learning activity. Protocols are developed to train the learner on how/when to vocalise their thoughts, and to aid the process of coding the output. The data gathered from think aloud protocols aligns with the conceptualisation of SRL as an aptitude rather than an event, because the data gathered is the learner's perception, as opposed to the event in action itself. The self-report data gathered from Think Aloud protocols differs from the data gathered from inventories in fundamental ways (Winne, 2010). Firstly, the data is gathered on the fly as the learning task occurs and cognitive operations are performed. This can result in a dataset which is more sensitive to evolutions in context. Secondly, the condition prompting the cognitive operation is the

learner choosing to report it, as opposed to the researcher asking a question. Learner training on think aloud protocols, and willingness to engage, will influence the dataset collected.

2.3.2 Second wave of research

During the second wave, researchers began to recognise multiple challenges with using self-report data to measure SRL (Boekaerts & Corno, 2005). One challenge was related to the accuracy of data from self-report measures of SRL; there was a significant risk that this data could be compromised by bias or poor recall. Another challenge was that the self-report measures often referred to student performance over several learning tasks and required students to report their perceptions of their aggregate performance. These perceptions were often inaccurate (Hadwin & Winne, 2001). Similarly, the use of think aloud data poses similar challenges with regard to the sample collected. The researcher has no information about how the learner samples experiences and choice of which to 'think aloud' about. Researchers are therefore required to analyse the context and make strong assumptions about the choices the learner made in order to think aloud, which may be hard to validate (Winne, 2010).

While there are accuracy and sampling challenges with using self-report data, a more fundamental challenge was triggered by the theoretical evolution in the second wave. SRL was increasingly conceptualised as a *dynamic process* rather than a static aptitude and new types of data were required to align with this conceptualisation (Aleven et al., 2010). New versions of SRL models were underpinned by the fundamental assumption that students monitor and make adaptations when they find their strategies are ineffective for a particular task. For example, students may monitor the use of note taking during a particular task, and find that the strategy is maladaptive, thus switching to another strategy such as concept mapping. Inventory self-report data is misaligned with this assumption, because it is not specific to a task, and thus does not allow students to report on how their behaviours are task

and context specific. In addition, inventories are often administered before the task thus decontextualising the responses.

Researchers explored the use of event data to measure SRL as a dynamic process that unfolds over the course of a learning activity. **Events are the very actions that learners perform** rather than descriptions of the actions or the mental states generated. Researchers cannot access the internal cognitive operations that are occurring, but use the events to **map the information** that is being operated upon and **make inferences** about the cognitive operations that are applied. Event data can be represented in 3 successively more complex layers; occurrence, contingency, and patterned contingency. Occurrence refers to the event itself without any reference to the preceding context, contingency refers to the event which has occurred after a specific occurrence in the context, and patterned contingency refers to regular and repeated arrangements of contingencies, and are more complex representations of events (Winne, 2010). In addition to offering a complex, contextualised measure of SRL, event data is not potentially compromised by issues with bias, recall or sampling.

Examples of event data are video and audio data, eye tracking data, log files, physical sensors of body posture, sensors that measure physiological indicators etc, and software sensors which detect if a learner has or has not performed a certain action, and trigger a system query. The query and the resulting response from the student can be used as measures of SRL

When learners are working offline, researchers have proposed techniques through which trace data can be generated. For example, Winne (2010) suggests that researchers create opportunities to trace the mental operations in which learners engage. Learners can be trained to demonstrate the proximal products of cognitive operations, and researchers can use the presence/frequency of traces as indicators of whether the operations occurred or not. Similarly, Graesser and McNamara (2010) hypothesize that learners must build a mental

model of a causal structure to develop deep comprehension. To obtain trace data, learners can be trained to draw a mental model. The standards which need to be met for generating trace data to evidence the learner capability to conduct a cognitive operation are:

- Learners must be **alert to the conditions** which prompt particular cognitive operations
- Learners must **understand which cognitive operations to apply** when conditions are perceived
- Learners must be able to **execute the trace** e.g., generate a flow chart
- An issue may be that learners lack the **necessary motivation** either to enact the operation or generate the trace
- The **resources to enact the operation** and generate the trace must be available (e.g., dictionary, or sufficient time or clearly laid out information)

In addition to the traditional sources of event data, researchers are currently exploring new forms of physiological event data, such as eye tracking and sensor data. Combining multiple forms of data provides a more accurate picture of the events that have occurred, and enables a more comprehensive, granular understanding of SRL (Bannert et al., 2014). This can contribute to the empirically based refinement of models of SRL, and the development of micro assumptions for further testing.

While event data has become increasingly popular, a comprehensive understanding of SRL requires going beyond the operations of SRL. In order to fully understand the development of SRL, analysis must also account for the conditions which trigger the use and adaptation of strategies. SRL can be modelled in terms of IF-THEN-ELSE (Winne, 2010). The ‘IF’ relates to an element of the task context, which prompts learner monitoring and enactment of a certain strategy (the ‘THEN’) subject to certain conditions being met at that point in time, or ‘ELSE’ another strategy. Learners choose their operations based on their

external and internal environment, with the internal environment including factors such as knowledge, developed skills, beliefs, expectations, interests and other mental states. Data collected to measure SRL such as self-report survey items administered before/after the learning event (and thus outside the arena of when learning decisions are taken), or think aloud protocols, or products generated (e.g. a solution to a problem or a correction to an error) only measures the operations themselves, and to get a full, accurate measure of SRL, it is essential to also account for the 'IF' which is shaping the operations.

Researchers have sought to understand the 'IF' shaping learner operations by gathering data on learner aptitudes. For example, Moos and Azevedo (2008) sought to collect data on learner self-perceptions of their plans, expectations, feelings and other factors, while they coded think aloud data. Similarly, Greene et al. (2010) reviewed research which considered how learner characteristics (such as prior knowledge and theory of intelligence) shape learner SRL. Aleven et al. (2010) and Graesser and McNamara (2010) logged patterns of behaviour in everyday tutoring sessions and in CBLE tutoring. They theorised about how these patterns of behaviour are grounded in SRL, considering how the standards for metacognitive monitoring (underpinned by learner aptitudes) informed the choice of study tactics and strategies.

In contrast to the static conceptualisation of aptitude adopted in the first wave of research, an understanding has emerged that aptitude is malleable and can be developed through training/education, or through the learner exercising their own agency and building this aptitude solely by themselves. As learners self-regulate, they may choose to modify the level or nature of an aptitude over the course of a learning activity, thus changing the context within which the leaner is operating (Greene et al., 2010). This has important implications for the measurement of aptitudes as they cannot be presumed to be static, and the output from self-report surveys taken at the beginning of a learning task, cannot assumed to be constant

throughout the task. The changing nature of aptitudes means that it is important to capture data during an intervention to record how aptitudes may change. In addition to the malleable nature of aptitudes, it is also increasingly understood that aptitudes may be theorised as categories, but often, a learner will not neatly fit into one category or another. Rather a profile of a learner can be developed depending on which selection of categories they most closely identify with (Zhou & Winne, 2012).

Apart from using self-report data to measure learner aptitude, researches can also take account of learner aptitudes through theorising based on the environment and learner responses. For example, when a CBLE is designed to offer learners various prompts for a range of study tactics and strategies, inferences can be made about learner aptitude depending on the types of strategies chosen, and how this evolves across the duration of the task (Winne, 2010). Further research is needed to explore how contextualised event data can be used to draw inferences about learner aptitude.

2.3.3 Third wave of SRL research

In the third wave of SRL research, measures are used which go above and beyond simply tracking SRL, and support the development of SRL (i.e., act as an intervention). For example, learners may use learning diaries in which learners record the events related to their learning activity. Researchers can later analyse these observations to measure/track SRL, and learners can also re-read their entries, which may trigger reflection and adaptations to SRL. Schmitz and Perels (2011) have explored the use of learning diaries to both measure SRL and trigger reactivity. They propose 4 mechanisms to explain how the monitoring triggered by SRL diaries can influence behaviours:

- **Creates awareness** of the importance of SRL and various components, and can thus trigger behaviours such as setting goals

- Diaries may **explicitly prompt SRL behaviours** such as goal setting, planning etc
- Reflection on learning is encouraged through the process of recording
- Finally, if the **whole cycle of SRL is embedded** by the instrument asking for variables corresponding to each phase, learners are made more aware of the relationship between the different phases of SRL

Another line of research is computer-based scaffolding of SRL. Researchers have designed software which can be used to both measure and scaffold SRL. The software collates traces of learner behaviour which can be analysed by researchers (the 'measurement' of SRL). The software also provides prompts to learners through hints from various agents (this is the 'intervention'). Examples of software are n-study (Winne et al., 2019) and Meta-tutor (Azevedo et al., 2010).

Software has been designed to capture various forms of data and support both self-regulation and socially shared regulation. In addition, systems have been designed in which pedagogical agents offer co-regulatory support to learners. However, I am not aware of any systems which support and measure human co-regulation of learners, and this is an area which requires further research. Table 5 below sets out examples of software systems, the types of regulation supported by the system, the datapoints captured and the measures of SRL. Multiple measures have been identified to measure learner SRL, in systems where learners are co-regulated by pedagogical agents. However, these measures focus on tracking the development of learner SRL, rather than investigating the effectiveness of various co-regulatory practices. In addition, the measures only apply to the highly structured environment in which pedagogical agents co-regulate a learner and would not apply for the varied contexts in which an expert human co-regulates a learner (e.g., human tutor co-regulating learner).

Table 5. Examples of pedagogical agent systems and measures of SRL applied

Examples of software systems	**Types of regulation supported**	**Data measures**
Meta-tutor (Azevedo et al., 2010)	Co-regulation and SRL (Agent-learner)	Each of the SRL skills are measured using log file data on the students' actions, decisions, ratings and verbal input
iStart (McNamara et al., 2007)	Co-regulation and SRL (Agent-learner)	Measures of SRL skills are the quality of the self-explanations, relative frequency of positive feedback from the agents, and relative frequency of requests to modify self-explanations
Radar, OurPlanner, and OurEvaluator (Järvelä et al., 2015)	SSRL	Log file data Thinking aloud protocols Eye tracking and physiological measures
Auto-tutor (Graesser & McNamara, 2010)	Co-regulation and SRL (Agent-learner)	Relative frequency of positive feedback versus negative feedback from Auto Tutor is a measure of the accuracy of student knowledge High proportion of student questions is a measure of student initiative Feeling of knowing and meta comprehension judgements are measures of metacognitive skills Measuring the quality of student contributions has require special models to be developed, as student contributions are often not grammatically or semantically well formed, and thus simple computations of the number of content words (nouns, main verbs, adjectives) is not enough. Statistical models of meaning, world knowledge and subject matter content are used. Does not currently allow students to select questions or topics. This would be a good measure of the students seeking control
Operation ARIES (Graesser et al., 2010)	Co-regulation and SRL (Agent-learner)	Student choice on what to study

		Student choice on whether to take a pre-test Can check whether students choose to go through chapters of an e-book in a linear fashion or to skip through Dialogue based metrics can be computed in a similar way as for Auto-tutor Relative rates of student questions, help seeking, problem selection, topic changes Proportion of tutor pumps and hints rather than prompts and assertions Proportion of expectations covered by students Patterns of tutor-student dialogue moves

When creating an instrument to measure and intervene re: SRL, the SRL cycle should be embedded in the instrument to ensure that it will promote SRL. Researchers need to choose variables from each phase of SRL, focusing on those which are of most research interest. An example is the PROSPECTS project (Järvelä et al., 2015) which is aligned to the theoretical model of SRL developed by Zimmerman (2000). In this project, 3 different instruments are used for different phases and processes of SRL:

- RADAR is used at the group level and individual level, to raise group awareness about SRL and SSRL, and about other groups members readiness to complete the task.
- OurPlanner is used by the group for forward facing planning, including goal setting and strategy selection .

- Finally, OurEvaluator is used by members to retrospectively evaluate performance, the achievement of established goals, and the retrospective use of strategies – this is again completed at the group level

Researchers must also consider whether measurement is occurring before, during or after the learning activity. If measurement and intervention is intended to take place during the learning activity, researchers must be mindful of the cognitive load theory (Paas et al., 2003) which states that in complex learning activities the amount of information and interactions that have to be processed can overload the finite working memory. Researchers must be careful to ensure that the additional cognitive load placed on the learner does not skew their performance and the study results. Further, while implementing a measurement and intervention tool, researchers must consider whether training should be provided to learners on how to use the instrument, and whether the student's commitment to work with the instrument could be a challenge.

As described above, research has moved away from sole reliance on self-report measures towards contextual measures such as event data. Researchers are interested in using event data, and triangulating it against other forms of data to build richer, more nuanced views of SRL. Different measures and tools continue to be used in parallel to each other, including self-reporting (Schellings & Van Hout-Wolters, 2011). For example, a joint project was conducted between 4 universities in which 4 measures, which consider contextual and situational factors and partly rely on self-reporting, were triangulated against each other (Dresel et al., 2015). Similarly, new self-report instruments continue to be developed (e.g., Andrade & Dugan, 2011). The main shift is that the measurement of SRL now aims to fully understand the impact of context and uses different measures triangulated against each other to do so. In addition, research has moved towards using data measures to both monitor SRL, and to intervene to scaffold it.

2.3.4 Gaps in the literature review

While researchers are interested in developing contextualised views of SRL, most research has focused on CBLEs where the learner is either working alone, or with a group of peers. The data measures used have ranged from tagging, to eye-tracking, and dialogue moves. Limited research has concentrated on scenarios where the learner is working under the guidance of an expert human in a virtual classroom environment (VCE), although there has been some research on the use of pedagogical tutors to co-regulate learners. We need to do more work to understand how we can measure the influence of another human on a learners SRL. How can we use the data generated from the VCE to signify co-regulation and SRL? This is an important question as it has implications for how we scaffold co-regulation in online environments, a topic that is particularly relevant given the boom in online tutoring (*Online Tutoring Market by Courses, Platform, and Geography - Forecast and Analysis 2023-2027*, n.d.; *Online Tutoring Services Market Size Report, 2022-2030*, n.d.) Researching in this new context would also potentially mean working with new types of data, such as natural unstructured dialogue between human tutor-learner. These new data sources could potentially offer a rich new source of insight into the development of learner SRL.

In addition to identifying signifiers in new research contexts, there is a need to further explore the use of theory as a foundation for the development of frameworks of signifiers. For example, Matcha et al. (2020) found that there is limited research on frameworks, which are both theoretically grounded, and sufficiently granular to investigate self-regulation in online learning environments.

Summary

I have reviewed the implications of theoretical developments in SRL on the types of data required to measure SRL, in each of the waves of research (as defined by Panadero et al., 2016).

During the first wave of research, SRL was conceptualised as a static aptitude. Learner aptitude for SRL was primarily measured using self-report data, generated from inventories or think aloud protocols. The influence of context on SRL was usually explored at fairly high level (e.g., by considering the domain specified whilst measuring aptitude for SRL).

In the second wave of research, SRL was increasingly conceptualised as a *dynamic process* rather than a static aptitude and new types of data were required to align with this conceptualisation (Aleven et al., 2010). Researchers explored the use of event data to measure SRL as a dynamic process that unfolds over the course of a learning activity. In addition to the traditional sources of event data, researchers also began to explore new forms of physiological event data, such as eye tracking and sensor data to enable a more comprehensive, granular understanding of SRL (Bannert et al., 2014).

While the use of event data has become increasingly popular to understand the influence of context on SRL, researchers have continued to acknowledge the importance of measuring SRL as an aptitude. In order to fully understand the development of SRL, analysis must also account for the conditions which trigger the use and adaptation of strategies, which can be understood with reference to learner aptitudes. In contrast to the static conceptualisation of aptitude adopted in the first wave of research, an understanding has emerged that aptitude is malleable and can be developed outside of, or within, a learning activity. This has important implications for the measurement of aptitudes as they cannot be presumed to be static, and the output from self-report surveys taken at the beginning of a learning task, cannot assumed to be constant throughout the task.

In the third wave of SRL research, measures are used which go above and beyond simply tracking SRL, and support the development of SRL (i.e., act as an intervention). For example, learners may use learning diaries in which learners record the events related to their

learning activity. Researchers can later analyse these observations to measure/track SRL, and learners can also re-read their entries, which may trigger reflection and adaptations to SRL. In addition, software has been designed to capture various forms of data and support various forms of regulation. For example, systems have been designed in which pedagogical agents offer co-regulatory support to learners, and data points are gathered on the SRL skills demonstrated by learners. While these platforms gather data on the SRL skills demonstrated by learners, further research is required which collects data-points on the effectiveness of various co-regulatory processes.

As described above, research into the measurement of SRL has evolved to utilise data points which account for both the influence of context, and shaping conditions (such as learner aptitude). However, there remain multiple gaps which require further work. For example, limited research has concentrated on scenarios where the learner is working under the guidance of an expert human in a virtual classroom environment (VCE). This study context would develop our understanding of how to measure the influence of another human on a learners SRL. Researching in this new context would also potentially mean working with new types of data, such as natural unstructured dialogue between human tutor-learner. These new data sources could potentially offer a rich new source of insight into the development of learner SRL. In addition to identifying signifiers in new research contexts, there is a need to further explore the use of theory as a foundation for the development of frameworks of signifiers. This would enable a theoretically grounded exploration of how co-regulation occurs in practice. These gaps in the literature are used to help formulate my research questions, set out in 2.5. Prior to formulating my research questions, I explore the use of learning analytics for the measurement of the regulation of learning.

2.4 Measurement of the regulation of learning: Learning analytics for the regulation of learning

Introduction

As the conceptualisation of SRL has evolved into a dynamic construct, new analytical techniques have been developed and applied. In order to analyse the sequential and temporal aspects of SRL, researchers have experimented with sequence mining and process mining techniques amongst others. Each of these techniques includes sub-sets of techniques. In addition, the techniques may be applied in combination with each other, or informed through other forms of analysis, such as decision trees, cluster analysis and dotted chart analysis.

In this chapter section, I review the analytical techniques which have been applied to explore SRL over the last decade. I start with an overview of the different techniques that have been used, prior to selecting a few techniques for a deep dive. For the selected techniques, I review the implementation methodologies, and explore contexts within which the analytics have been applied. I identify gaps in the literature and opportunities for additional work.

2.4.1 Overview of analytical techniques (2010-2022)

During the course of the past decade, researchers have experimented with analytical techniques which enable them to investigate the sequential and temporal characteristics of SRL. This aligns with the conceptualisation of SRL as a dynamic and evolving process which is shaped by context, in the second and third waves of research into SRL. In order to fully investigate SRL from a sequential and temporal perspective, researchers have explored the use of new techniques, either using analytics which are specially developed for learning sciences or by borrowing from other fields. These techniques are often used in combination with self-report analytics to provide a fuller view of SRL.

There is significant research interest in using cluster and classification analytics, as well as sequential and temporal analytics. Cluster and classification analytics are often used in combination with forms of sequence and process mining, by helping researchers to

discover groups which show similar behaviours, and may thus share common patterns. ElSayed et al. (2019) conducted a systematic review of research in learning analytics and educational data mining from 2011-2019. They identified 109 relevant papers, out of which 59 used some form of data analytics in combination with self-reporting, while the remainder purely relied on data analytics. Table 6 below presents the analytics used in the study, and table 7 presents the types of data analysed. As evidenced, clustering, classification, sequence mining and process mining have been popular over the course of the last decade, and there has been a strong interest in applying these to log data. Most studies (61.82%) have been conducted in the context of an undergraduate course, 22.73% were conducted in context of school education, 7.27% in open online education, 5.45% in the context of post-graduate education, and 2.73% in the context of the workplace.

Table 6. Systematic review of the learning analytics used in SRL papers between 2011-2019 (ElSayed et al., 2019).

Analytics	Number of papers in which the technique has been used (2011-2019)	Description
Clustering	42	Unsupervised analytical techniques, which can be used to categorise behavioural profiles into subset clusters (Romero & Ventura, 2010; Segedy et al., 2015). The most commonly used clustering technique is Agglomerative Hierarchical Clustering (22 studies) followed by K-means clustering (13 studies).
Predictive analytics	19	Supervised algorithms which predict membership of individual data instances to a group based on quantitative information relating to various features of those data instances using a training set of previously labelled items (Phyu, 2009; Romero & Ventura, 2010). The most frequently used predictive analytics is regression analysis (used in 17 studies) followed by Bayesian classifiers (6 studies), neural networks (5 studies), 4 studies used decision tree classification, and another 4 used support vector machines. 1 study used K-Nearest Neighbour algorithm.
Temporal Data Mining techniques (sequence mining	26	Examines the connections between recurrent SRL events according to order and temporality. Includes sequence. mining and process mining (Reimann et al., 2014; Sonnenberg & Bannert, 2015)

and process mining)		Process Mining creates a model to describe a process which can either be time stamped or fully ordered. Process Mining tools and approaches include fuzzy miner, ProM, PM squared, bupaR, Markov models, heuristic miner, and conformance checking. 13 studies have used process mining. Sequence mining examines event patterns, analysing the frequency of the occurrence of patterns (Taub & Azevedo, 2019). Techniques can be implemented separately or combined as in the differential sequence mining methodology (Kinnebrew et al., 2013). 15 studies have used sequence mining

Please note that the analytics referred to in table 6 may be used in combination with each other in the referenced papers.

Table 7. Systematic review of the types of data used in LA/EDM papers on SRL between 2011-2019 (ElSayed et al., 2019).

Log data (77 studies)	Self-reported data (59 studies)	Assessments (27 studies)	Demographic data (9 studies)	Video recordings (5 studies)	Chat & forum data (5 studies)

Please note that the data types referred to in table 7 may be used in combination with each other in the referenced papers.

While there is increasing interest in the application of sequence mining and process mining to SRL, the contexts and data types which have been explored remain relatively narrow. For example, most studies have been conducted in an undergraduate setting, using hypermedia and think aloud protocols. Applying analytics to real world online settings, and exploring the use of online trace data, remains relatively rare. It is important to further our understanding of the development of SRL in real world settings, so the evolution of SRL in these contexts can be optimally supported. In addition, trace data offers a valuable source of insight which may be less prone to bias, and can be gathered in an unobtrusive manner.

Additional research is required which explores how sequence mining and process mining techniques can use trace data (Bannert et al., 2014).

An important consideration is the type of analytics that is most suitable for the research, and whether sequential or temporal analytics should be used, or both in combination with each other. Sequential analysis identifies which events have a strong relationship, and tend to follow each other. Temporal analysis identifies when events occur and how they influence each other over the course of the process (Reimann, 2009). Researchers can decide which type of analytics to use depending on the type of question that is asked in relation to the order and timing of events, and what types of analytics are thus most useful. For example, do we want to know how frequently an isolated range of events occur in a particular order? Or are we interested in how frequently each of the events occurs and what transitions occur to and from other events? An additional layer of detail would be to understand the probability of moving from one event to another. Researchers align the scope of their research with the conceptualisation of the learning process, and this model will inform the type of analytics selected. For example, research which utilises a divergent/convergent model conceptualises ideas as occurring more frequently in the divergent state. Analytics on the probability of events and transitions between states, such as process maps, would be appropriate to test the hypothesis in this model. Similarly, an approach towards understanding the sequence of actions in problem solving (e.g. cycles of initiation-response-feedback) requires understanding of frequency of events and transitions. Both process mining or lag sequence analysis may be useful here. Alternatively, a researcher may simply need to know how regularly a sub-sequence repeats within a larger dataset, without needing information on the transitions from and to that sub-sequence. In this case, optimal sequence mining can be applied.

Given the dynamic conceptualisation of SRL adopted by my research (see section 2.1), sequence and process mining analytics appear to be appropriate analytical techniques. In addition, these can be extended through forms of predictive analytics, such as decision trees. In the following section, I therefore review sequence mining, process mining, and predictive analytics in more detail. In particular, I consider examples of the studies in which these analytics have been applied and the insights/questions raised for future work and new contexts. I then summarise the gaps in the literature and formulate my research question.

2.4.2 Sequence Mining

Sequence mining approach

As discussed above, sequence mining refers to the identification of sequences or sub-sequences within a set of data. Various algorithmic approaches can be taken, where 'algorithmic approach' refers to the relation between different events being unknown and meaning being created by investigating statistical relationships between them. Sequence mining can be differentiated from process mining, or hidden markov models. These approaches take a whole sequence as a model approach, and aim to represent the holistic process, rather than isolated sub-sequences (Bannert et al., 2014). Process mining is further discussed in the section below, along with further discussion of the differences between sequence mining and process mining, and the benefits/considerations of applying each or in combination.

There have been a wide variety of sequential analysis approaches over the past 10 years. For example, theory driven approaches recode log files to theoretically meaningful events, define the length of the sequence, or set the threshold for significance (Roll & Winne, 2015; Winne, 2010), while data driven approaches extract the most common sequences from log files, regardless of content or length (Bannert et al., 2017; Beheshitha et al., 2015). Some

researchers focus on the variance of sub sequences from learner to learner and apply multi-level analysis techniques (Azevedo et al., 2019; Taub et al., 2014) while others focus on clusters of learner and apply chi square analysis.

Regardless of the specific approach, most sequence mining studies can be structured as three phases of work:

1. Pre-processing phase:
2. Mining and characterisation phase
3. Analysis phase

The pre-processing phase consists of 2 steps. Firstly, researchers need to consider whether the raw data needs to be re-coded. This can be done using an action library, which specifies the links between the observed events and the recoded conceptualised label. For example, researchers may recode data based on theories of SRL (Winne, 2017b), or apply a tool related coding scheme to code computer log files (e.g., Siadaty et al., 2016). Alternatively, researchers may decide that the raw data can be mined in its existing form (e.g., Van Laer & Elen, 2018). The second step of the pre-processing phase is to assign each sequence of events to a learner or case. Chronological ordering is adequate for investigating sequential characteristics, but temporal characteristics also require the exploration of time elapsed between consecutive events.

The mining and characterisation phase explores characteristics of sequences and sub-sequences. Various algorithmic approaches can be taken, where 'algorithmic approach' refers to the relation between different events being unknown and creates meaning by investigating statistical relationships between them. Algorithms can be applied using off the shelf software, for example, SPMF (Fournier-Viger et al., 2017), ProM (van der Aalst, 2012), SPAM (Sequential Pattern Mining) by Ayres et al. (2002), or TraMineR (Gabadinho et al., 2011). Parameter settings need to be configured before an algorithm can be applied. These include

time constraints, and thresholds for the identification of events/sub sequences, and methods for counting events/sub-sequences.

Finally, the analysis phase consists of an exploration of the results from the data mining, in light of the research questions. For example, researchers may explore how internal and external variables influence the sequences or sub-sequences (Winne & Baker, 2013), or the impact of interventions on learning outcomes, or patterns of activity (Duffy & Azevedo, 2015; Jovanović et al., 2017). To check the strength of the correlation between the presence of sub-sequences and a variable (e.g. learning outcomes), statistical tools such as Chi square may be used or Kinnebrew et al. (2013) alternatively used a t-test, which identifies the most interesting sub-sequences in an exploratory analysis. Researchers may choose to examine the influence of variables on sequences, and/or the influence of sequences on variables (e.g. learning performance).

While there has been a diversity of approaches towards sequence mining, there is no established methodological framework which can guide the application of sequence mining analytics, and enable comparison of results between studies. Researchers have called for future work in sequence mining to contribute towards methodological discussions (Azevedo, 2014; Molenaar & Järvelä, 2014), including considerations on which techniques to apply. Van Laer and Elen (2018) note that most research in this field does not explain the choices behind algorithms, units of analysis or parameter settings. They propose initial guidelines for the transparent and systematic analysis of sequential data, which are detailed below:

1. Researchers should consider the **time frame** over which variables being analysed are stable and use this as the unit of analysis
2. When recoding raw data, researchers should ensure the **link** between data forms of **different levels of granularity** is made clearly

3. Researchers should clarify the reasons for their **choice of parameter settings**. It is not always possible to choose parameter settings based on theory, and where it is not possible, researchers should make it clear that the choice of parameter settings is arbitrary.

Examples of past work implementing sequence mining

Recent studies have demonstrated the importance of the specific type of pattern mining analytics used, in addressing the research questions posed. For example, Kuvalja et al. (2014) used a range of techniques to analyse children's self-regulatory behaviours and self-directed speech, including frequency analysis, sequence mining and periodic pattern mining. Frequency analysis was used to determine the prevalence of behaviours, but showed no differences in the self-regulatory behaviours between the two groups. Sequential lag analysis, a type of sequential pattern mining, was used to detect patterns of sequential behaviours, but no statistically significant adjacent sequences were detected. Finally, temporal (t)-pattern analysis, a type of periodic pattern mining, was applied to identify recurring non-adjacent patterns. This type of analysis showed a statistically significant difference between the two groups, with frequent occurrence of more complex temporal patterns of behaviour amongst learners with a specific speech disorder. Kuvalja et al. (2014) showed the value of using periodic pattern mining to detect extended temporal patterns, which could not be detected through sequential and event-level analysis and suggested that further research should be done in this field to explore how t-pattern analysis can be most effectively used, and parameters set.

In another study, Cukurova et al. (2022) analysed tutor behaviours in a virtual classroom environment using the CM-Spam algorithm (a type of sequential pattern mining). Statistically significant patterns were identified for high and low-performing groups of tutors, and used as the basis for a decision tree. Decision trees were able to predict tutor performance

with high levels of accuracy, indicating that sequential mining can be an effective technique to create sequences of actions as input for building decision trees.

Researchers had also experimented with new methodologies combining different types of sequential and episode mining. Kinnebrew et al. (2013) explored the use of sequential and episode mining for comparing learner behaviours, using trace data generated from a classroom study with Betty's Brain. The researchers had previously used hidden Markov models to represent the direct (probabilistic) transitions between states. However, they noted that these models are an aggregation of the underlying events and transitions, and may not depict fine grained events/transitions. In order to detect the patterns of fine-grained events and transitions occurring, a technique such as sequence pattern mining can be used, which provides more precise definitions of interactions. However, the researchers noted that the implementation of sequence pattern mining can be complex, as it may result in the detection of a very high number of patterns, especially if the algorithm is configured to allow for a gap to account for noise. The number of patterns can be reduced by setting the threshold of traces over which the pattern must occur to a high percentage, but this may result in the analysis missing important patterns which would help us understand the interactions/behaviour. To help address these issues, the researchers developed a new methodological approach combining the use of both sequence and episode mining to detect differentially frequent patterns, and analysed how behaviours were correlated to productive and counter-productive phases of student activity.

Gaps identified from literature review

While there have been a number of studies applying sequence mining for SRL behaviours over the past decade, most of these studies have mined student behaviours in settings where the student is working alone (e.g. hypermedia environments). There is limited work that I am aware which applies sequence mining to a context where a student is being

co-regulated by a tutor. Cukurova et al. (2022) examined tutor behaviours in the context of a Virtual Classroom Environment, but no study that I am aware of has applied sequence mining to represent tutor-student interactions. I propose that this is an important area for research, as it will enable the identification of significant patterns of behaviour, which differentiate successful from unsuccessful interactions, and potentially lay the foundation for future work into adaptive scaffolding.

There is also a need for additional research experimenting with the use of various algorithms and parameter settings. For example, what are the benefits of applying sequence mining algorithms in combination, if any? A discussion of these topics will contribute towards the broader methodological discussion around sequence mining, and could potentially help to inform future work on methodological frameworks for sequence mining.

2.4.3 Process Mining

Process mining approaches

Process Mining is a subset of Educational Data Mining (EDM). While most forms of EDM aim to detect specific patterns, process mining tracks holistic end to end processes. It aims to facilitate and deepen an understanding of educational processes, by making unexpressed processes explicit. Traditionally, process mining has been applied most often in business process management, with the aim of improving processes and monitoring compliance, but there is increasing interest in the potential for its application in learning analytics and education in general (Bannert et al., 2014).

Process Mining is applicable when it is assumed that there is a process governing a particular sequence of events (van der Aalst, 2012); this is relevant for learning where learners and tutors can be assumed to engage in particular processes, depending on the specific nature of the task and underlying pedagogy. The process that is uncovered by a process model may be driven by a learning strategy, or an external resource given to the

learner (Sonnenberg & Bannert, 2015). By reflecting the dynamic nature of learning, process mining enables process related issues to be highlighted and addressed. For example, bottlenecks within the process, which prevent learners from progressing their work, have been the focus of previous research (van der Aalst, 2012). Furthermore, by discovering and mapping out learner-resource, learner-learner and learner-tutor interactions, it is possible that process mining can also help to identify and understand what types of interactions are related to positive learning outcomes (Juhaňák et al., 2019). Finally, process mining has the advantage of producing graphical depictions which are relatively easy for the end-user to interpret and understand (Juhaňák et al., 2019; Sedrakyan et al., 2016).

There are three types of process mining: process discovery, conformance checking, and extension/enhancement.

- Process discovery is the detection and visualisation of unknown processes from data in an event log, and is the most relevant for education, where processes are not fully known or established.
- Conformance checking refers to the evaluation of existing practices against an established process, including the identification and possible resolution of bottlenecks.
- Extension/enhancement refers to the development of new processes or refinement of existing processes, using the data found in the event log.

Following data sampling, the first stage of Process Mining is collecting and processing data into an event log, which can be used for process mining. This can be challenging when the data inputs are from various sources, such as log files, audio, video etc. This challenge was acknowledged by the Process Mining Manifesto (van der Aalst et al., 2011) which called on researchers to explore the use of various data sources and ways of resolving challenges. To date, many researchers applying process mining to education have

relied on a single source of data for the event log. For example, Bannert et al. (2014) coded think aloud verbal data, to build an event log. Similarly, Juhaňák et al. (2017) relied on logfile data from an LMS to build an event log, when using process mining to investigate quiz taking behaviour amongst college students.

The second stage of process mining is selecting the most appropriate algorithm for the scope of work. Process Mining consists of a wide range of algorithms, which vary in their ability to deal with complex processes, and noise within data (van der Aalst, 2012). Prior to selecting an algorithm, education researchers must consider how complex the behaviour patterns are likely to be, whether it is possible for learners to engage in a choice of patterns, and how much noise there is likely to be in the data. Alpha algorithms such as Petri nets are deterministic in nature and do not deal well with scenarios in which there is uncertainty over learner behaviour (van der Aalst, 2012). They may therefore not be suited for investigating patterns of self-regulation, where learners can engage in a number of routes. In contrast, heuristic algorithms such as Fuzzy Miner are well suited to analysing situations in which learners may engage in a range of patterns of behaviour, and are able to detect less frequent variants (Juhaňák et al., 2017) These algorithms also deal well with complexity, by allowing thresholds to be set before events are included in the final output, with some events being abstracted and others aggregated. This enables different levels of analysis to be conducted, which may yield additional insights into the process. Where there is a high variety of SRL behaviours in the sample data, the process model may underfit/over-generalise i.e., allow for many more behaviours than actually observed. In this case, mining subsets may be useful approach – this is called trace clustering. A plug in has been implemented in Pro M called Disjunctive Workflow Schema (de Medeiros et al., 2008), which is combined with the Heurestics Miner Algorithm to split the event log into clusters iteratively until high precision process models are achieved.

Table 8 sets out the process mining algorithms that have been used in a selection of studies on process mining for learning behaviours, and process mining for SRL in particular. As shown, the algorithms used include Inductive Miner, Heuristic Miner, Fuzzy Miner, and pMineR, with the latter two algorithms being used most frequently. A systematic review conducted by Saint et al. (2021) stated that Fuzzy Miner was the most frequently used algorithm, used in 9 SRL studies. The pMineR has been used in 6 studies, while the Inductive Miner algorithm has been used in 2 SRL studies.

As discussed by Saint et al. (2021), the dominant use of Fuzzy Miner and pMineR is due to the ability of these algorithms to handle complex, and unstructured processes such as learner SRL, where learner agency means that a large number of variants occur. These algorithms prioritise the clear visualisation of the most important transitions. In contrast, where the learning processes being investigated are highly structured, alternative algorithms such as Inductive Miner and Heuristic Miner are more suitable. These algorithms create process models against all four conventional criteria of process model soundness. The first criteria is replay fitness, which refers to the ability to map an end-to-end process from the data to the model. Secondly, precision, which refers to how closely the model mirrors the transitions found in the data, and excludes any transitions which are not found. Thirdly, simplicity, which refers to the clarity of the process model. Finally, process models are evaluated against their ability to accommodate future patterns of behaviour i.e., the generalisability of the model. The fitness between the mined model and the event log used can be measured using the 'Improved Continuous Semantic Fitness' (de Medeiros et al., 2008); this indicates the number of correct parsed event sequences, and penalises for allowed extra behaviour in the model.

Table 8. Process Mining tools used in Learning Analytics research.

	Inductive Miner	Fuzzy Miner	pMineR
Toolkit	Pro M, BupaR	Pro M Disco	pMineR package

		Celonis	
Metrics	Frequency, time	Frequency, time	Transition probability
Key features	Guarantees process Model soundness	Simplified visualisations, highlighting key events and transitions	First Order Markov Models, enabling quick comparisons
Examples of learning analytics studies	(Bogarín et al., 2018; Cerezo et al., 2020)	(Bannert et al., 2014; Maldonado-Mahauad et al., 2018; Sedrakyan et al., 2016)	(Ahmad Uzir et al., 2020; Matcha et al., 2019; Saint et al., 2020)

Saint et al. (2021) conducted a review of the algorithms set out in Table 8, by executing each algorithm using a common dataset, and evaluating the research utility of the resulting models and metrics. The work found that the process models generated by Inductive Miner and Heuristic Miner could be validated against the established criteria, but did not provide an insightful visual representation of SRL processes. The process models included all micro-level processes, and did not sufficiently differentiate the most frequent and probable transitions. In contrast, the Fuzzy Miner and pMinerR algorithms generated process models which highlighted significant micro processes and transitions. This enabled a rich discussion of the evolution of SRL within the specific learning context. The researchers noted that the Inductive Miner or Heuristic Miner algorithm should be selected in research studies where there is a structured learning process, and where it is important to investigate the soundness of the process model against fixed criteria. In contrast, in research studies where the intended output of the study is to conduct a contextually rich exploration of the evolution of SRL, Fuzzy Miner or pMinerR may be more appropriate. In addition, Saint et al. evaluated the types of metrics generated, and found that the absolute frequency and median time lag metric generated by Fuzzy Miner, and the probability of transition metric generated by pMinerR, were effective in highlighting key micro processes and transitions. They proposed using a combination of metrics from both algorithms, to enable immediate understanding of both the absolute and relative importance of a process and transition.

Examples of past work implementing process mining

Recent work has advanced our understanding of how process mining can effectively be used to monitor SRL processes. Studies have examined process mining (1) from a methodological point of view, and (2) used process mining to explore the relationship between variables (e.g., learning outcomes), and the holistic processes adopted by learners.

Studies with methodological insights

Saint et al. (2020) proposed a methodologically rigorous approach towards the use of process mining for SRL. They developed the Trace-SRL framework, which is a methodology consisting of the following steps:

- Firstly, a theory and model for SRL is adopted by the researcher
- Secondly, the macro level processes of SRL are identified and the micro level processes are mapped onto these macro level processes.
- Thirdly, the online trace data is extracted and coded according to the micro level processes.
- In the next stage, the sample may be clustered using a supervised or unsupervised method.
- Finally, the most appropriate process mining algorithm which suits the scope and nature of the work is selected, and applied to the clusters, which can then be compared and analysed.

In a prior study, Siadaty et al. (2016) proposed a theoretically grounded framework to explore the influence of scaffolds on SRL processes, using micro trace data. The authors asserted that more work is needed to integrate (1) trace data into research methods and (2) align trace data with 'theoretical models of SRL. Their study proposed a trace-based methodology to measure SRL processes at macro and micro levels, as well as the effect of tools and scaffolds on theorized SRL elements. Macro level processes include categories and

phases such as planning and monitoring, while micro processes are more specific activities such as goal setting within planning (Greene & Azevedo, 2007). The trace based micro analytic measurement protocol follows the below process:

1. A 'precise and plain' formulation of a SRL model plus the type of conceptualisation adopted, and the specific constructs/processes which are to be scaffolded.
2. 'Precisely formulate' the context i.e. theorise how scaffolds may influence the micro SRL processes. This follows the IF-THEN modelling suggested by Winne (2010).
3. Finally, the researchers capture the user traces into log files and use these to measure the constructs of interest. This requires:
    - Identifying the micro events which comprise the SRL micro process. To manage the large volume of low-level events in log files, researchers should take a targeted theory informed approach towards collecting data. Identifying the relevant micro events requires researchers to comprehensively review the software system and user interfaces to identify relevant traces
    - Translating the traces to SRL events. Actual traceable actions may be at a lower level of granularity than the SRL micro events, and an entire sequence of traces equals an event. Researchers must first build the sequences which represent the micro SRL events. These can be stored in log files, and traces will be translated through a parsing algorithm into the larger grained SRL events

The theoretically driven approaches proposed by Saint et al. (2020) and Siadaty et al. (2016) guide researchers to consider SRL models in their entirety, reviewing all macro level phases and processes, prior to mapping out the relevant micro processes and events. This

enables researchers to implement process mining to analyse the influence of the scaffold or variable on holistic SRL processes.

Empirical studies demonstrating the value of process mining

In addition to work that has guided researchers from a methodological perspective, studies have been conducted which have demonstrated the unique value of process mining techniques. Analysis of the influence of interventions or variables on the structure of learning processes can offer valuable insights, which cannot be obtained through narrower analytical techniques, such as sequential analysis or frequency analysis. For example, Sonnenberg and Bannert (2015) used process mining to highlight structural differences in the processes of learners who were presented with metacognitive prompts, versus a control group of learners who were not given prompts. Following the implementation of the Heuristics Miner algorithm, the researchers found that ANALYSE (including orientation, planning and goal specification) was well embedded in the processes followed by the experimental group but not the control group, and that students in the experimental group show more loops between metacognitive and cognitive events (so called 'regulation circles'). They also demonstrated the value of using process mining as opposed to simply using measures of frequency; for example, the process maps showed that EVALUATION is a well-connected state, despite being of low frequency in both groups. Similarly, MONITOR is high frequency in both process models but hardly connected in either process map. These findings demonstrate how process analysis can reveal important information that frequency analysis cannot.

Similarly, Sonnenberg and Bannert (2015) considered the temporal and contextual conditions under which instructional support is beneficial for learners using data mining and process mining. They classified the effectiveness of prompts on regulatory processes, explored how the classification correlates with learning outcomes, and identified the conditions through which prompts can encourage regulatory processes. Prompts were

classified as ‘effective’ based on an analysis in the increase of metacognitive events above the baseline following the prompt. After classification, process mining was used to analyse the student activities prior to the prompt, to understand the most appropriate temporal positioning of the prompt. It was found that orientation and monitoring activities increased the effectiveness of prompts. In addition, the process model for non-effective prompts resembled models of SRL, indicating that non effective prompts may be disruptive to learners who are already effectively regulating. In contrast, the process models for effective prompts were less structured and showed weakly connected regulatory events.

Research has also highlighted the value of process mining in detecting adaptive and maladaptive patterns that learners may engage in. For example, Bannert et al. (2014) led a process mining study on the self-regulation of students working with hypermedia. The researchers coded think aloud verbal data from undergraduate students navigating a hypermedia site, using a theoretical framework specifically designed for hypermedia environments. The researchers selected two groups with high and low attainment scores for further analysis of their self-regulatory processes. Extreme groups were selected for analysis, as past research has demonstrated high variation in learning and regulating activities of students, meaning that an aggregation of learning processes from all students would not therefore be appropriate. Application of the Fuzzy miner process mining algorithm (from the open-source Pro-M toolkit) revealed the maladaptive patterns that less successful students engaged in, such as a loop between monitoring, and reading/repeating.

Process mining also enables process analysis at different levels of detail. For example, Sedrakyan et al. (2016) applied an adaptation of the Fuzzy miner process mining algorithm called Disco, when analysing the self-regulatory behaviours of computer science students engaging in complex problem-solving task (namely domain modelling). This study demonstrates the flexibility of process mining tools, which can be used to conduct analysis at

multiple levels. Disco was used to provide a bird's-eye level analysis for the entire group; between best performing and worst performing participants; and between early and late sessions. The researchers also applied a fine-grained analysis at an event level, to further understand the differences between sequences of events. Similarly, Schoor and Bannert (2011) have explored sequences of social regulatory processes during a CSCL task and their relationship to group performance. This study used process mining to detect patterns that distinguished high performance from low performance groups.

Further work which uses process mining to explore holistic learning processes in a range of learning contexts is regarded as important, as it will allows the development of SRL theories on a micro-level (Molenaar & Järvelä, 2014; Sonnenberg & Bannert, 2015). It would also inform future work on the development of scaffolds for learners, which encourage them to execute regulatory skills which they may struggle to execute spontaneously, as per the 'production deficit' (Veenman, 2007; Veenman et al., 2006).

Gaps identified from literature review

Given the relatively nascent nature of the field, researchers have identified several gaps in work on process mining for SRL. Firstly, additional work needs to be done to investigate the correlation between scaffolding of SRL and learning outcomes. While there is a body of work on the relation between metacognitive prompts and learning outcomes, limited work has been done on the influence of the broader co-regulatory practices on SRL. For example, Lin and Lehman (1999) prompted students to give reasons for their actions (e.eg. what is your plan for solving the problem?) in a computer-based simulation environment. The study showed significantly higher transfer performance, but was limited to examining a narrow type of prompt. Similarly, Bannert et al. (2015) examined the influence of self-directed metacognitive prompts on navigation behaviours (e.g. time spent on visiting relevant web-pages) and learning outcomes. They showed that these prompts resulted in more

strategic navigation and higher transfer performance. A promising avenue for further research is to explore the influence of a diverse range of external practices/scaffolds on SRL processes, and learning outcomes.

Another area for research is to examine optimal features and sequential positioning of scaffolds. Previous research has found that the effectiveness of a scaffold may depend on the features of the scaffold (Bannert, 2009), or its positioning in the overall process. For example, Thillman et al. (2009) determined that strategy instruction should be embedded while learning to be effective, while Wittwer and Renkl (1998) also found that strategy instruction should be integrated into on-going cognitive activities to be effective. Researchers could use process mining to validate and refine these findings. Researchers have called for more work to be conducted on further understanding the influence of the sequential positioning of prompts within the overall process (Molenaar & Järvelä, 2014; Sonnenberg & Bannert, 2015). Azevedo and Hadwin (2005) highlighted many questions for additional research such as the appropriate time to scaffold, how to provide the appropriate type and level of support for the phase of SRL, and how to fade support as the student's progress in regulating their learning. This could inform future work on micro models of SRL with assumptions on the conditions of effective scaffolding (e.g., timing, type of scaffold, student expertise).

2.4.4 Predictive analytics

Predictive analytics approaches

The purpose of predictive analytics is to develop a model which will predict the class or value of new data, based on rules developed through mining a training data set. The model is based on the assumption that observed variables in a training data set (known as 'features'), can be used to predict the value/class of new data. Predictive modelling can be distinguished from explanatory modelling, in that the model is used to predict the features of new data; in contrast, explanatory models are used to analyse all available evidence to develop causal

explanations for the observed outcomes. Predictive modelling is therefore forward looking, while explanatory modelling is a reflective process.

Predictive modelling is conducted across three stages, as outlined below:

1. Problem identification: Predictive modelling is often conducted to identify where additional support and interventions are required. For example, predictive modelling may be used to identify students who are at risk of dropping out of their degree programmes, so that these students can be offered the appropriate support. The first step in predictive modelling is therefore to define the situation within which an intervention is required, and consider whether predictive modelling is appropriate. This analytical technique is suitable when there are quantifiable characteristics that can be modelled, a clear outcome, a large set of data, and if the training data is indicative of future data sets.
2. Data collection: Researchers need to select the outcome variable (e.g., learning outcome), and the suspected correlates (e.g., time spent on learning platform), prior to gathering the relevant macro data. Recent research has focused on the collection of event-based data to develop predictive models and the conversion of this into meaningful features for predictive models. This is a promising area for research, which would enable the development of interventions to guide students learning during the course of an activity.
3. Feature selection: The data collected is analysed to identify micro features which correlate with the value to be predicted. The degree of correlation will vary on the feature, and learning algorithms vary in their use of correlates; some algorithms may use all features, while others will use only the most informative correlates.
4. Building the predictive model: Following data collection and feature selection, a predictive model can be developed using the training data. A range of algorithms exist

for building models (Brooks & Thompson, 2017). These vary from decision trees (which are repeated partitions of the data based on a series of single attribute "tests") to linear and logistic regression, and support vector machines and neural networks. The algorithms can be configured using the available parameters; for example, researchers can configure leaf size or depth of tree, when building decision trees, to ensure a level of generalizability.

5. Evaluating the predictive model: The prediction accuracy, precision and recall of a model can be assessed by testing it on a portion of the historical dataset that was held out for testing. Researchers often use the process of k-fold cross validation, in which the dataset is partitioned at random into k segments; k predictive models are developed by training on each of the segments except 1 segment, which is 'held out' and used for testing. The test results are then pooled from each of the models, and an evaluation of model quality is conducted. This process allows (1) all available data points to be used for testing (2) nearly all data points to be used for training.

Examples of studies

Research on predictive analytics has investigated the variables which are predictors of learning performance, engagement, as well as the modelling psychological constructs.

For example, Kizilcec et al. (2017) analysed the relationship between students' SRL strategies in MOOCs and course achievement. They defined student strategies by looking at student interaction with online materials and SRL surveys. They found that goal orientation and strategic planning were predictors of course goal achievement. Similarly, Cho and Yoo (2017) predicted students learning achievement using students SRL online behaviours, with seventeen attributes being used in the classification model. The student SRL online behaviours were discovered through mining log data and survey data.

Researchers have also studied predictors of student engagement. For example, researchers have predicted the likelihood of a student dropping out of a MOOC (Xing & Du, 2018), or whether students will graduate from school in time (Aguiar et al., 2015). Baker et al. (2011) predicted formative achievement for a student based on their interactions with an intelligent tutoring system.

In addition to engagement and performance measures, researchers have studied whether constructs such as affective and emotional states can be modelled using predictive analytics, with textual discourse or facial characteristics used as features (D'Mello et al., 2007; Wang et al., 2015).

2.4.5 Gaps identified from literature review

While there is an established body of research exploring predictive analytics for learning outcomes using state-based data, research using event-based data is still in its early stages. In addition, most event based predictive analytics focuses on learner interaction with the online environment and resources (e.g., time spent watching lectures), with a limited body of work using SRL behaviours as predictors of achievement.

I am not aware of any research which applies predictive analytics to examine the relationship between tutor co-regulation and teaching/learning outcomes in a Virtual Classroom Environment. In particular, I am not aware of research which uses patterns of interaction between tutors and learners to predict teaching and learning outcomes with respect to SRL. Cukurova et al. (2022) used patterns of tutor practices to predict tutor performance, but this study was focused on examining tutor practices (rather than tutor-learner interactions), and the scope of the study was not focused on SRL.

It would be valuable to apply predictive analytics to the field of tutors co-regulating learners in a Virtual Classroom Environment. This would deepen our understanding of what constitutes successful tutor-learner interactions, in light of the findings from the analytics and

regarding theory. It would also help us identify the scenarios in which tutors and learners may need additional support.

Summary

There has been an increased interest over the past decade in implementing learning analytics using event data to explore SRL. This aligns with the process-oriented conceptualisation of SRL that has been widely adopted (as described in section 2.1 of this thesis). In particular, there is interest in using sequential and temporal analytics to explore the dynamic nature of SRL and the influence of the broader context on the evolution of SRL. Predictive analytics are implemented to understand the strength of the association between SRL events and outcomes.

In this section, I have reviewed sequence mining, process mining and predictive analytics. I have outlined the considerations at each stage of implementation for these analytics, and the distinctions between various algorithmic approaches. I have also reviewed examples of empirical studies which implemented these forms of analytics, and highlighted gaps for additional research. I have summarised the gaps identified below.

Regarding sequence mining, there have been several studies applying sequence mining for SRL behaviours, most of these studies have mined student behaviours in settings where the student is working alone (e.g. hypermedia environments). There is a limited research work applying sequence mining to a context where a student is being co-regulated by a tutor. I propose that this is an important area for research, as it will enable the identification of significant patterns of behaviour, which differentiate successful from unsuccessful interactions, and potentially lay the foundation for future work into adaptive scaffolding.

In addition, we also need research which builds our methodological understanding of the sequence mining implementation. For example, this research could explore the potential

benefits of applying sequence mining algorithms in combination or using different types of parameter settings for algorithms. This type of research could contribute towards the broader methodological discussion around sequence mining, and could potentially help to inform future work on methodological frameworks for sequence mining.

About process mining, research which implements process mining in a range of learning contexts is regarded as important, as it may enable the development of SRL theories on a micro-level (Molenaar & Järvelä, 2014; Sonnenberg & Bannert, 2015). This research could also potentially inform the development of scaffolds for learners, which encourage them to execute regulatory skills which they may struggle to execute spontaneously. Another area for research is to examine optimal features and sequential positioning of scaffolds. For example, what types of scaffolds are most effective, and is there a certain point in the overall SRL process when they should be introduced? Previous research has found that the effectiveness of a scaffold may depend on the features of the scaffold (Bannert, 2009), or its positioning in the overall process.

Concerning predictive analytics, research using event-based data is still in its early stages. In addition, most event-based predictive analytics focuses on learner interaction with the online environment and resources (e.g., time spent watching lectures). There is limited research work using SRL behaviours as predictors of achievement. To my knowledge, no work implements predictive analytics using tutor co-regulatory processes. Research highlights significant features of tutor-learner interaction, and using these features to build predictive analytics would be valuable in exploring the relationship between tutor co-regulation and learning outcomes. This type of work is interesting from a research perspective (through identifying key features of tutor-learner interaction, which influence learning outcomes). It could also be potentially valuable from an operational perspective (through identifying when scaffolding/intervention/training are required).

The above literature review has also demonstrated that there are multiple studies which apply forms of sequence mining, process mining, and predictive analytics to explore SRL, but there is no work that I am aware of which employs all three forms of analytics. Each type of learning analytic performs a different function, and further research is needed to apply all three to understand the value (if any) of combining the insight from each form of analysis.

2.5 Research questions

Introduction

While there is a vast body of research on learner SRL, most research focuses on settings in which the learner is working independently, and the influence of context on learner SRL often refers to the influence of environmental factors, such as learning resources. There is a need for additional research, which explores the influence of social context, considering how others can promote or inhibit the development of learner SRL. Specifically, there is a need to build a holistic and granular understanding of how various teaching practices can influence learner SRL. This research would be valuable from a theoretical perspective, enabling the development of micro-models of SRL. In addition, it can be potentially valuable from an operational perspective, supporting teacher training and development.

Given the need for further research into teacher co-regulation of learners, I proposed to conduct my research in the context of a virtual classroom environment (VCE). In this study context, human tutors work on a one-to-one basis with learners in an online setting. This study context is particularly interesting, as it generates rich forms of data, including natural language dialogue, and online trace data. Research can explore a theoretical approach towards analysing tutor-learner interactions, utilising the various forms of data.

In this section, I summarise the gaps identified from the literature review regarding tutor co-regulation of learners and formulate the proposed research questions and contributions. This is followed by a presentation of my overarching research question, and sub research questions.

2.5.1 Formulation of research questions

Firstly, there is a gap in understanding how established models of SRL can be used to explore how human tutors co-regulate learners in Virtual Classroom Environments (VCE). How can established models of SRL be interpreted to act as the basis for research into this area? What are the relevant theories of learning which can be useful in building this interpretation? I propose to address this gap through RQ1:

RQ1: How should SRL be conceptualised to explore the influence of human tutors on learner SRL processes in a VLE?

My goal is to contextualise an established model of SRL, such that it can be used to investigate the influence of tutoring practices on learner SRL operations. I put forward this interpretation in section 4.1 of this thesis.

Secondly, there is a gap in the development of a theoretically based framework of signifiers, which represents how human tutoring practices influence learner SRL in a VCE. How can the contextualised model of SRL be used to develop a framework of signifiers? What are the relevant human tutoring practices and signifiers? What are the relevant learner practices and signifiers? I propose to address this through RQ2:

RQ2: How can we derive theoretically meaningful signifiers, to represent the influence of human tutors on learner SRL processes in a VLE?

My goal is to develop a theoretically based framework of signifiers which will represent how tutor practices can influence learner SRL processes. The data generated from the VCE will be used to signify tutor co-regulatory practices and learner SRL practices, to

build a holistic view of how human tutors can influence learner SRL. My contribution is set out in section 4.2 of this thesis.

Finally, there have been many studies which have highlighted structural differences in the SRL processes adopted by high versus low-attaining learners (e.g. Bannert et al., 2015) and identified significant patterns of behaviour, as well as some early work on the relation between SRL behaviours and learning outcomes. However, no previous work has applied sequential, temporal and predictive analytics to co-regulatory processes. There is a gap in understanding the relation between co-regulatory processes and patterns, and teaching and learning outcomes. A deeper understanding of the relationship between co-regulatory processes and performance would help develop our understanding of optimal teaching practices, including positioning and design of scaffolds.

My literature review has also demonstrated that there are multiple studies which apply forms of sequence mining, process mining, and predictive analytics to explore SRL, but there is no work that I am aware of which employs all three forms of analytics. Each type of learning analytics performs a different function, and it would be interesting to apply all three to understand the value (if any) of combining the insight from each form of analysis.

In order to address the highlighted gaps in the literature review, I propose to address the below research question.

RQ3. What are the key differences between high and low-performing tutors in a VCE, in the patterns of tutor-student interaction, with respect to SRL?

My goal is to implement the framework of signifiers though sequence mining, process mining and predictive analytics to build our understanding of the regulatory processes that occur in a VCE, and how these vary with tutor performance and learning outcomes. My contributions towards this research question are set out in chapter 5 of this thesis.

2.5.2 Presentation of research questions

Following the formulation of my specific research questions, I have presented my study context, overarching research question and specific research questions in Figure 5.

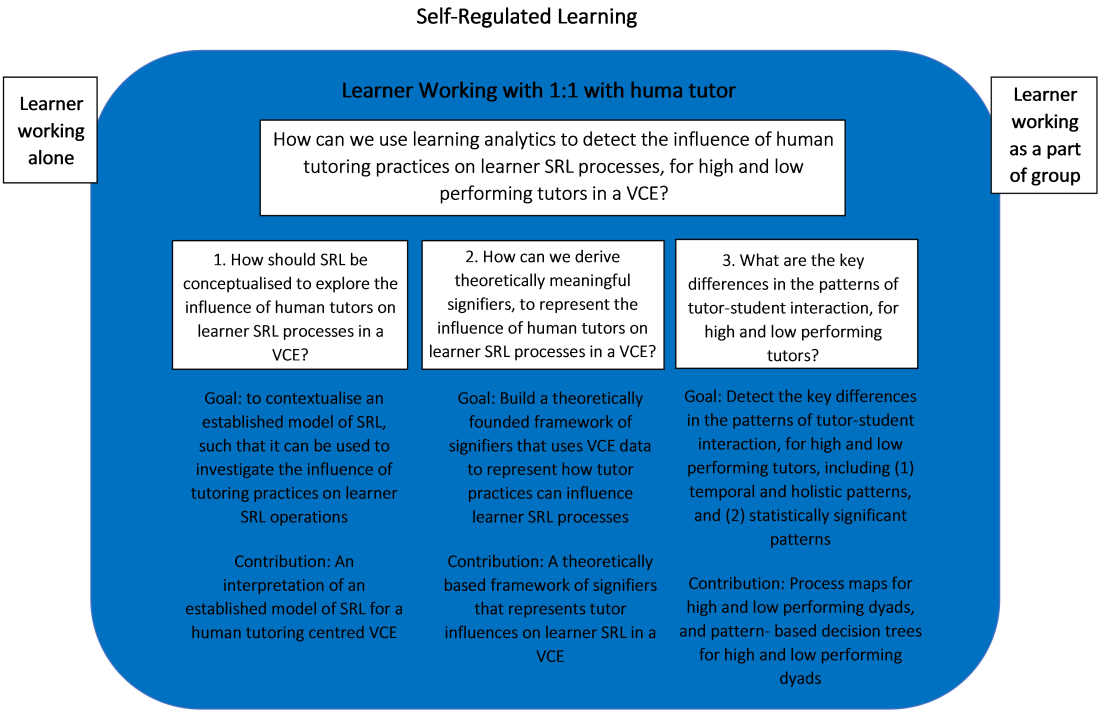


Figure 5. Research questions

Chapter 3: Methodology

3.1 Studied context and design-based research methodology

In this chapter, I will outline the study context in detail and provide an overview of the design-based research approach taken.

3.1.1 Study context

As stated in Section 2.5 of this thesis, I have defined the overarching research question to be

"How can we use learning analytics to detect the influence of human tutoring practices on learner SRL processes, for high and low-performing tutors in a VCE?"

The specific context studied in my research is a Virtual Classroom Environment developed by a commercial supplier, Third Space Learning.

Third Space Learning (TSL) provides one to one maths tutoring for primary school children in the U.K. The primary school learners are in years 5 and 6 (aged 10/11 years old), and the target schools typically have a high percentage of students on Free School Meals. TSL offers low-cost human tutoring in an online environment, with a tutoring centre established in Sri Lanka. The human tutors are centrally trained through the tutoring centre, and come from a range of professional backgrounds (including but not limited to education).

Study context: VCE platform

The Virtual Classroom Environment (VCE) consists of an interactive whiteboard, with a chat function, shared editing tools, and tutor-controlled incentives (e.g., effort points, stickers, emojis). The tutor and learner can hear each other, but cannot see each other. TSL produces lesson content in the form of slides which are displayed on the whiteboard. The lesson content is aligned to a broader curriculum map, designed by TSL. Prior to commencing lessons at TSL, students complete a diagnostic quiz, on the basis of which

topics from the curriculum map are selected for the student. Each lesson is roughly 45 minutes long and follows the structure below:

- Warm up questions: Consists of 2-3 questions, which the learner is expected to solve independently, to enable the tutor to diagnose gaps in their understanding
- Objectives: Tutor introduces the topic that the lesson will focus on
- Problem solving: A set of problems gradually advancing in complexity
- Reflection: A brief plenary, in which the learner is guided to reflect on their learning experience

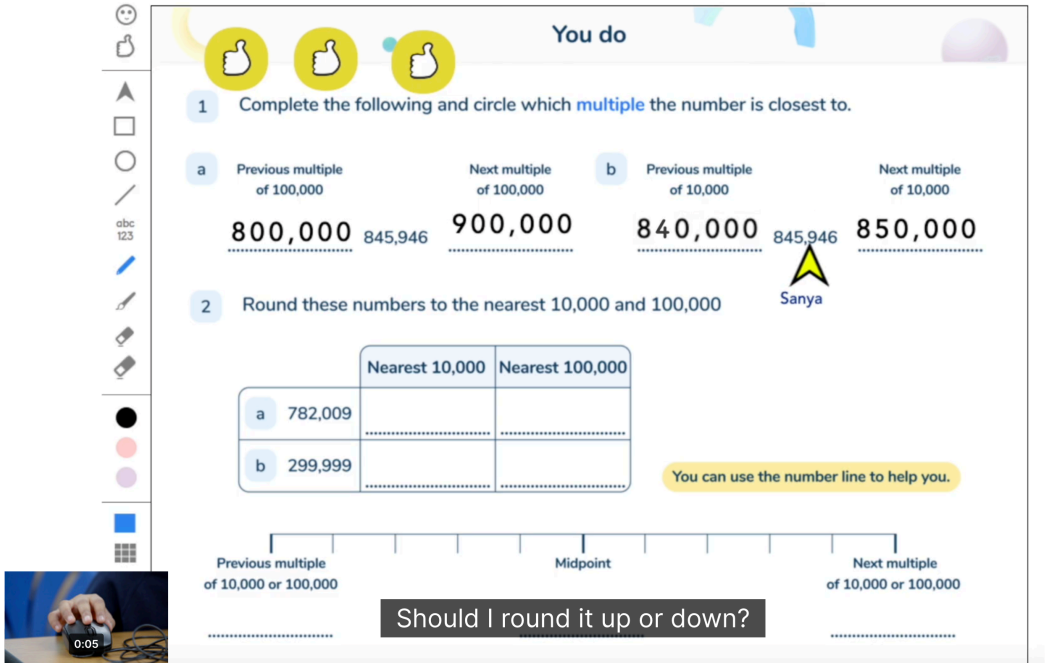


Figure 6: Screenshot of the interactive Third Space Learning whiteboard.

While the lesson objectives, lesson plan and lesson slides are set centrally, tutors have a relatively large degree of control over the flow of the tutoring session itself. For example, tutors control the movement between the lesson slides, and the pace of the session. They can choose to skip lesson slides, and where to start/end a session. In addition, they can introduce self-generated content into a session by starting a blank page within the session, and writing new instructional content on it. Tutors are also

given various tools which they can use in a flexible manner to motivate learners and build rapport. For example, tutors have access to a library of stock images. They can attempt to build a rapport with learners by discussing their interests and hobbies, and presenting selected images from the library to engage the learners.

Study context: tutoring training and evaluation

Tutors are centrally trained, based on a student-centred pedagogy. Student-centred learning refers to an interactive style of teaching, in which the teacher encourages the learner to actively participate and think aloud (O'Neill & McMahon, 2005). The quality of the tutor training is monitored through human evaluators, who review a small percentage of session recordings each term. Formative assessments are based on a rubric developed in collaboration with IOE. The rubric is based on 5 pillars; subject knowledge, student-centred teaching, student satisfaction, familiarity and reflection. The student-centred teaching pillar considers whether the tutor personalises learning to the student and encourages student engagement and autonomy, and is scored on a fixed scale.

Study context: data captured

TSL captures a range of data about tutors, learners, and tutoring sessions, which is listed in Table 9 below. As illustrated, TSL mainly relies on metrics captured outside the tutoring session to understand tutoring performance. For example, TSL assesses the quality of tutoring through the human QA of ~5% of all recorded sessions, or self-report data generated from students or tutors. However, there are issues with using self-report data, such as inaccuracy and bias (Boekaerts & Corno, 2005; Winne & Jamieson-Noel, 2002) and human QA is unable to monitor a sufficiently large proportion of sessions. There is a need to understand the quality of the sessions using data captured while the session is in progress. This will provide a more accurate and comprehensive understanding of the features of sessions which are regarded as being successful.

Table 9. Types of data captured by TSL

Data	**When collected**	**Data type**	**Description**	**Data use case by TSL**
Student profile	Prior to TSL programme (supplemented by tutors after sessions)	Descriptive	Information supplied by school and tutor about the student, including demographics, and learning level/behaviours	Inform tutors, prior to delivery of sessions
Student pre-programme questionnaire	Prior to TSL programme	Qualitative	TSL pre-programme questionnaire, to test attitudes towards learning	To compare to post programme questionnaire and test impact of programme
Student pre-quiz scores	Prior to TSL programme	Quantitative	Learner quiz when they start at TSL to assess understanding of topics and generate curriculum map.	To compare to post programme quiz and test impact of programme
Tutor and student audio from sessions	During the session	Qualitative	Sessions	N/A
Session recordings	During the session	Qualitative	Sessions	~5% used for human evaluation
Tutor rankings (based on QA)	Updated at the end of each half term	Quantitative	Rankings based on QA (occurs for 1-2 sessions per term, per tutor). QA uses a rubric designed alongside IOE	To monitor quality of tutoring and support struggling tutors
Tutor ratings (based on learner surveys)	Post tutoring sessions	Qualitative	Learner survey carried out at the end of a session	To monitor quality of tutoring and support struggling tutors
Student post-quiz scores	After LO completed	Quantitative	Learner quiz after each learning outcome completed	To compare to pre programme quiz and test impact of programme
Post session tutor feedback	At the end of each tutoring session	Qualitative	Includes what steps were covered in session	Inform tutors, prior to delivery of sessions

Student post diagnostic assessment	At the end of TSL programme	Quantitative	To compare to pre diagnostic, and assess attainment	To compare to pre diagnostic scores, and assess attainment

My research explores *"How can we use learning analytics to detect the influence of human tutoring practices on learner SRL processes, for high and low-performing tutors in a VCE?"* in the context of the VCE supplied by TSL. My work aims to build an understanding of how in-session data can be used to support tutors in co-regulating learners effectively. As such, I have adopted an iterative design-based research approach to my work which is outlined below.

3.1.2 Design-based research methodology

Design-Based Research (DBR) methodology is appropriate when a theoretically grounded solution is required to be applied to authentic learning contexts. DBR aims to iteratively develop a theoretically grounded solution for the practical context (Reeves et al., 2005). The iterative nature of DBR acknowledges the complexity of shaping a theoretically based solution that addresses the regulation of learning as it unfolds in authentic learning contexts. The iterative nature also ensures that the feedback from stakeholders such as tutors, learners and management is incorporated into the design, and that the contribution is relevant to practice.

DBR prioritises the development of a solution for a specific setting through iterative design, rather than promoting a solution which can be generalised across contexts. In addition, the intended outputs from DBR are diverse and can range from full to partial solutions for the specific setting (Edelson, 2002).

The hallmark of DBR is its iterative nature. The exact research processes followed are flexible, and the structure of the methodology may vary somewhat from one researcher to another. For example, van den Akker (1999) conceptualised the DBR methodology as

occurring across 4 phases: (a) analysis of the educational problems identified by researchers/practitioners, (b) development of solutions, informed by theory and availability of technology innovations, (c) implementation and initial evaluation of the solution in the target setting, and (d) reflection on the overall processes to iterate the solution for the subsequent cycle. On the other hand, McKenny and Reeves (2012) outlined three core processes of DBR:

- The **'discovery'** stage (sometimes termed 'analysis') aims to understand all aspects of the problem, and "seek out and learn from how others have viewed and solved similar problems" (McKenny & Reeves, 2012, p. 85). Analysis is a fundamental aspect of design-based research and is applied continuously throughout DBR, to inform the design of the intervention.
- The **'design'** stage aims to develop interventions such as specific technologies or "less concrete aspects such as activity structures, institutions, scaffolds, and curricula" (Design-Based Research Collective, 2003, pp. 5–6). This may involve evaluating a range of potential designs, prior to selecting the most promising option for development.
- The **'evaluation'** stage aims to evaluate iterations of the design to detect areas for further improvement. Evaluations can be conducted through assessing the impact of the intervention on standard attainment measures, such as tests and quizzes. Interventions may also be evaluated through insights from interviews and observations.

In the section that follows, I discuss how the structure of my study maps to the overarching design-based research cycle, and the key outcomes and contributions of each stage. In the next section, I will discuss each stage of the DBR cycle in more depth, and in particular highlight the iterative nature of my work.

3.1.3 DBR cycle mapped to the research questions, outcomes, validation and contribution

As illustrated by Figure 7 below, the three stages of DBR are aligned to the three research questions posed by my study:

- **Discovery stage:** As described above, the 'discovery' stage consists of exploring of the problem to understand how others have viewed and solved similar problems. In the context of my study, the 'problem' sought to be addressed is developing an understanding of the influence of tutoring practices on learner SRL. In RQ1 I developed a conceptualisation of SRL for the studied context, exploring the literature, and studied context, to consider how a tutor could co-regulate learners. The research was conducted to address RQ1 which could be considered a 'discovery'. This research stage produces a review of theoretical constructs used in similar studies, and an analysis of the data available in the studied context. The contribution from the 'discovery' stage is the contextualisation of an established model of SRL, integrating it with relevant learning theories, such that it can act as the foundation for research into the co-regulation of learning by human tutors.
- **Design stage:** The Design stage aims to design a framework of signifiers which represent how human tutors co-regulate learners in a VCE. This stage of research aligns with RQ2 *('how can we derive theoretically meaningful signifiers to represent the influence of human tutors on learner SRL processes in a VCE?').* The 'design' stage involves a consideration of the range of approaches towards building a framework of signifiers, informed by a review of past studies in similar contexts, and empirical analysis of a sample of VCE

sessions. The contribution from this stage is a theoretically grounded framework of signifiers for tutor co-regulatory practices and learner SRL operations.

- **Implementation and evaluation stage:** In this stage, I implement the framework of signifiers using learning analytics, and evaluate its ability to detect the differences in how high-performing versus low-performing tutors co-regulate learners in a VCE. Process mining and sequence mining are implemented to identify the holistic and sequential differences in the processes adopted by the groups of tutors, while decision trees are executed to analyse the ability to predict tutor performance by using significant tutor-learner sequences as features. The resulting process maps, significant sequences and decision trees are analysed to explore the differences in the processes adopted by the groups of tutors. Implications for the contextualised model of SRL and the framework of signifiers are considered, as well as for learning analytics and models of SRL more broadly.

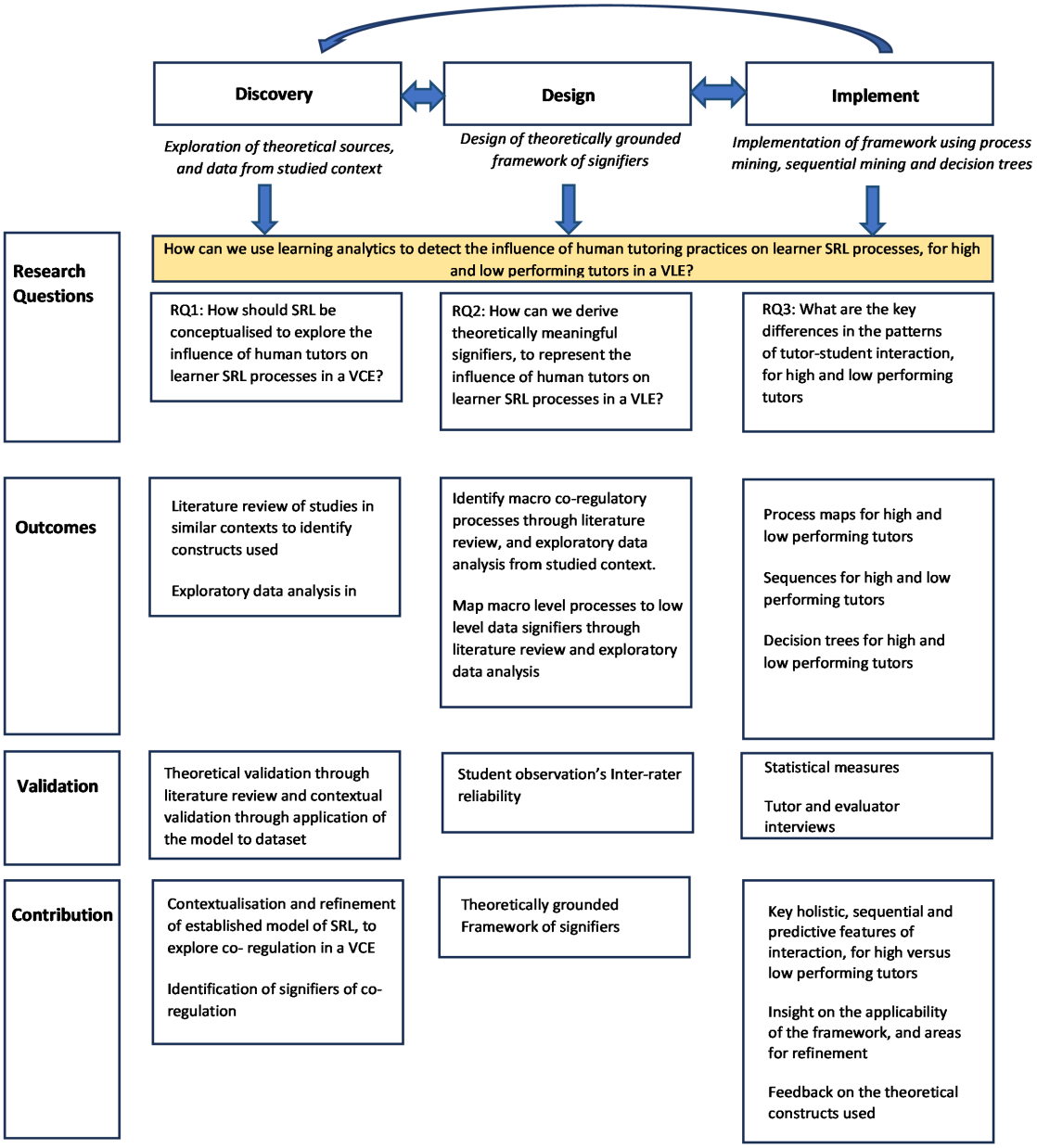


Figure 7. Overarching DBR cycle mapped to the research questions, outcomes, validation and contribution

3.1.4 DBR cycle mapped to phases of work

In this section, I will describe the phases of work planned for each stage of the DBR cycle, as set out in Figure 8.

.

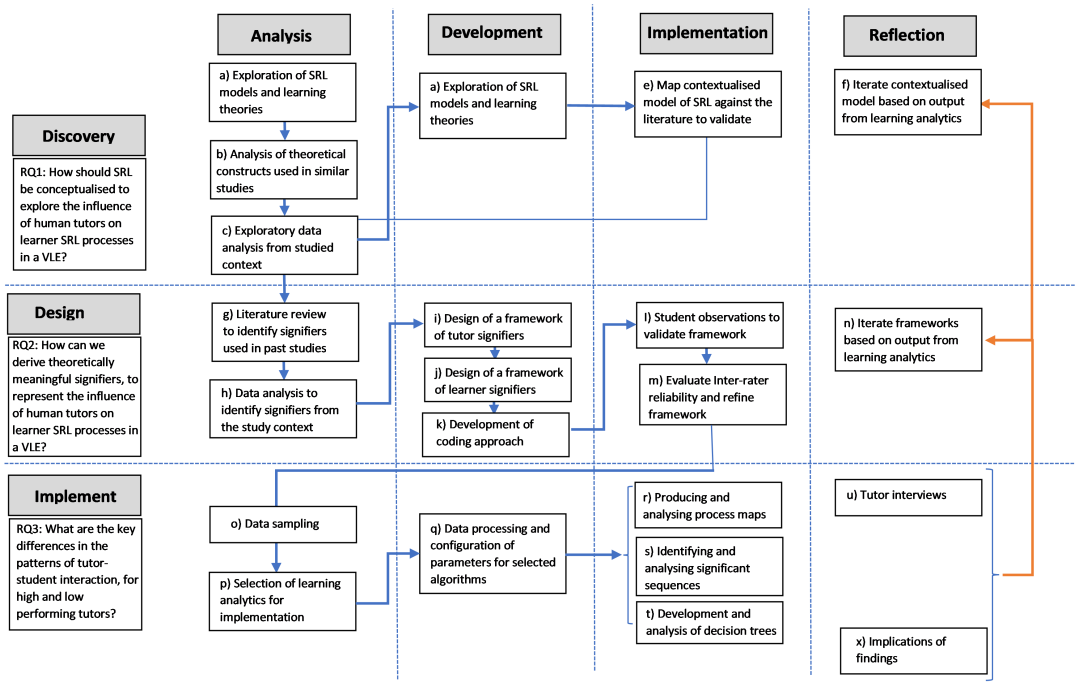


Figure 8. The DBR cycle mapped to phases of work

Figure 8 maps the 4 phases of work ('Analysis', 'Development', 'Implementation' and 'Reflection') to each of the stages of work ('Discovery', 'Design', and 'Implement'). The blue arrows represent the workflow within each stage, while the orange arrows represent the iterative workflow between stages.

In the first stage **'Discovery'**, I address RQ1: how should SRL be conceptualised to explore the influence of human tutors in learner SRL processes in a VLE? The analysis stage required an exploration of the relevant literature from various aspects. Firstly, the literature was reviewed to identify studies conducted in similar contexts to evaluate the theoretical constructs and consider whether these could act as a theoretical basis for my research. Secondly, established models of SRL were identified from the literature and assessed for their suitability as a basis for my research, against formulated criteria. Thirdly, a model of SRL was adopted and interpreted for the study context. The interpretation was validated by

mapping the new sub-categories of the model to the literature, to establish the relationships declared.

In the second stage **'Design'**, I addressed RQ2: how can we derive theoretically meaningful signifiers to represent the influence of human tutors on learner SRL processes in a VCE? This stage of work was based on the contextualised model of SRL developed in stage 1. Firstly, I considered various approaches towards building a framework of signifiers, and assessed the suitability of each approach for my research. Secondly, I considered the data sources available for analysis, and the usefulness of the various sources. As the focus of this stage of work is the optimal representation of how human tutors co-regulate learners, data sources were considered regardless of the ability to automatically extract the relevant features of the data at present. Automated extraction of the relevant features of data is seen as being outside the scope of this research study, and manual coding is allowed for. Following the consideration of approaches and data sources, I designed the framework of signifiers through a mixed methods approach. Firstly, signifiers were identified from the literature and mapped to components of the theoretical model. Secondly, the framework was iterated through the review of the data generated from a sample of 50 randomly selected TSL sessions. This process was implemented to develop a framework of signifiers for learner operations, as well as tutor practices.

The mixed methods approach consists of three main stages:

- Identification of a theory driven framework ('identify')
- Adaptation of the theoretical framework to fit the research purposes ('adapt')
- Merging the fine-grained actions from the data to the adapted theoretical framework ('integrate').

This approach enables the development of frameworks which are both theory-driven (and thus generalizable to a larger number of learning contexts), and data driven (and thus

practicable to apply to the specific learning context). Further, the signifiers noted in the development of a framework through a mixed methods approach are observable actions rather than broad definitions, which enable these frameworks to be operationalised through learning analytics (Cukurova et al., 2016). Frameworks which are solely driven by theory are often broad and high level in nature to allow them to apply to a diverse range of learning contexts. A mixed methods approach is therefore particularly appropriate in an online learning context, in which fine grained trace data is collected, meaning that purely theoretical frameworks need to be refined to more granular forms.

Recent studies have implemented versions of the mixed methods approach, as granular frameworks for online learning environments are developed. For example, Saint et al. (2020) developed the Trace-SRL framework, which is a methodology for implementing process mining for SRL and includes the development of a theoretically driven framework for fine grained online trace data. Within this methodology:

- Firstly**, a theory and model for SRL is adopted** by the researcher
- Secondly, the **macro level processes of SRL** are identified from the **literature**
- Thirdly, the **micro level processes** are mapped onto these macro level processes, through a review of the **empirical data**
- Finally, the **online trace data is extracted and coded** according to the micro level processes

I have adapted the approach for the specific context of my research, and outlined the stages of the approach in table 1 below. Steps (2c)-(3) are repeated to develop signifiers for both tutor practices and learner operations.

Stage	**Description**	**Chapter and section number**
Identification of theory driven framework	As discussed in chapter 2, the Winne and Hadwin (1998) model of SRL is	4.1

		adopted as the basis for my research	
Adapt the framework to fit research purposes	2a) Interpret the model of SRL for my study context	The contextualised model of SRL developed in chapter 4.1 will be used as the basis for the framework of signifiers	4.2
	2b) Restructure the contextualised model	Re-organise the model to enable optimal development of a framework of signifiers	4.2.2
	2c) Review the data generated from the VCE	Review the characteristics and grain size of the data generated from the VCE, prior to identifying potential signifiers	4.2.3
	2d) Map the literature to the framework	Identify potential signifiers of co-regulation from the literature	4.2.4
	2e) Map potential signifiers to the framework	Adapt the signifiers identified from the literature to the study context	4.2.4
Integrate empirical data into the framework	50 sessions are analysed to refine and build the list of signifiers		4.2.5

Following the development of the framework of signifiers, I developed an approach towards coding the signifiers in real time. Each of the signifiers was defined in granular terms, and relevant examples were provided to aid the coding process. A numerical code was provided for each of the signifiers. With regards to the coding process, given that manual coding was required this thesis followed the approach that was set out in Cukurova et al. (2022) ; The authors of this study developed a simple tool which enabled them to manually code events observed in real time, providing data on the event observed, timestamp, and session ID. I developed a similar coding tool, and used this to capture data from sessions that were played back for observation. A snapshot of the data collected is provided in Appendix I.

The reliability of the framework of signifiers was evaluated through the coding of 10 randomly selected sessions, with 2 trained coders, and an inter-rater reliability of ordinal k alpha 0.724. The framework was iterated on the basis of the results of the inter-rater reliability test, and the inter-rater reliability was evaluated again on the basis of coding another 15 randomly selected sessions. An inter-rater reliability of ordinal k alpha 0.842 was achieved, which was judged sufficient to proceed with coding the main dataset.

In order to validate the framework, 30 sessions were observed in person. As tutors for the sessions are based in Sri Lanka, sessions could only be observed from the learner perspective. The recordings of the sessions were later coded using the framework, and compared to the recorded observations. The framework was refined in light of the findings from the observations.

In the third stage **'Implement',** I addressed RQ3: what are the key differences in the patterns of tutor-learner interaction for high and low-performing tutors? The first step was a consideration of the criteria for high and low performance. These criteria could refer to performance judged by human evaluators, learners, or through traditional attainment measures such as progress scores. I decided to explore performance from the perspective of human evaluators, as this would allow me to identify the patterns of interaction demonstrated by groups of tutors ranked as high or low performing by human evaluators, and enable discussion of the opportunity to use learning analytics to support formative tutor training. The human evaluators were highly experienced tutors who were regarded by the TSL team to exemplify good tutoring practice, and had been selected by TSL to form part of their quality assurance team. Outside of these credentials, there was no other independent evidence available on the expertise of the human evaluators. In addition to using human evaluator rankings as criteria for performance, a smaller group of tutors was sampled based on performance as measured by progress scores. Progress scores are the difference between the

learner score on a diagnostic test, and post tutoring session quiz, for a specific learning outcome. This sample was taken to enable discussion of the differences in processes adopted by high versus low-performing tutors, when the metric for performance is varied. Further details of the approach taken towards data sampling, and a breakdown of the data sample collected, are provided in 3.2 of this thesis.

Following data sampling, specific forms of learning analytics were selected for implementation. In 2.4 of this thesis, I reviewed literature that illustrated a gap in using process mining and sequential mining to identify holistic and pattern-based differences in how tutors co-regulate learners. In particular, these forms of analytics have rarely been used in combination with decision trees. I thus propose to use these three forms of analytics to plug the gap in existing research. Prior to the implementation of each of these analytics, I reviewed the potential algorithms which could be implemented and assessed their suitability for my study. I also processed data into the relevant formats and configured parameters, prior to implementing each of algorithms for my data sets. The quality of the resulting process maps, sequences and decision trees was assessed using statistical measures where possible. In addition, a set of 30 interviews were conducted with tutors who were randomly selected from the high-performing and low-performing groups, to complement the insight from the learning analytics.

Finally, the output from the learning analytics was evaluated in light of the existing theoretical and empirical research, as well as the complementary self-report data sources (tutor interviews). The implications for the theoretical constructs and frameworks used were considered, and these were iterated as appropriate, prior to the analytics being executed once again. In addition, the implications for learning analytics in general, as well as broader learning theories and operational considerations, were also explored.

3.1.5 Ethics

I ensured that my research is compliant with the Data Protection Act (2018; see Appendix G). Table 10 below maps data analysed to ethical considerations, such as obtaining informed consent for the use of data. As noted, Third Space Learning stores all data related to the use of the TSL platform and VCE. It provides access to view the sessions through its online application, via the cloud, and I do not save or store the VCE platform and session data myself. Third Space Learning has an agreement in place with the customer schools, which gives informed consent for the VCE platform and session data to be viewed and analysed for research purposes. I have viewed sessions and transcribed data to create a body of text, which is securely stored on my encrypted laptop. All data has been anonymised, and any references to student names have been removed from the transcripts. I have also been careful to ensure that there is no reference to any personal information in the session transcripts, as these consist of natural, unstructured dialogue.

In addition to analyzing the online sessions, I have conducted observations of 30 sessions in person, to validate the inferences drawn about student-tutor dialogue and actions in the VCE. As tutors are based in Sri Lanka, the observations have been designed to focus on learners. The observations have been conducted face to face using pre-designed templates to collect data during the sessions. All observations have been conducted with the informed consent of both TSL, and the relevant schools, parents and students. Further, the data gathered has been anonymized prior to being stored securely on my encrypted laptop, and any references to personal information has been removed.

Finally, I have also collected qualitative and quantitative data from a set of 30 semi-structured tutor interviews. I obtained informed consent from tutors prior to conducting these interviews, and anonymised interview transcripts.

Table 10. Mapping data types to ethical considerations.

Types of data analysed	**Ethical considerations**		
	Data storage	**Informed consent**	**Anonymisation**

VCE session recordings (including tutor-learner audio, and video showing interactive whiteboard, VCE editing and reward tools, chat function, pointer)	Stored by TSL TSL provided view only access to myself	Schools have given their informed consent to TSL for the use of VCE session data for research purposes	I anonymised all observation and transcription data
VCE platform data (student post session surveys, tutor post session surveys)	Stored by TSL TSL provided view only access to myself	Schools have given their informed consent to TSL for the use of VCE platform data for research purposes	I anonymised all VCE platform data collected
Student pre and post attainment scores	TSL provided an anonymised data set, which has been securely stored on my laptop	Schools have given their informed consent to TSL for the use of VCE platform data for research purposes	I ensured all data was anonymised prior to analysis
Student observations	I observed students face to face and collected data using a pre-designed template. The observation data is securely stored on my laptop	Schools, parents and students were all required to provide informed consent prior to the observations being conducted. Any students who did not return the signed consent sheet were excluded from observations	I anonymised all observation data
Tutor rankings by human evaluators	TSL provided the tutor rankings, which have been securely saved on my laptop	Tutors have given their informed consent to TSL for the use of their data for research purposes	I anonymised tutor data prior to the use of this data

3.2: Methodology (data sampling)

Introduction

In this chapter section, I will set out the characteristics of the data sampled for my thesis. This chapter will follow the below structure:

- I will start by setting out the different sets of data used, aligned to stages of work (3.2.1)
- I will then review the structure and characteristics of each data set

3.2.1 Data-sets used, mapped to stages of work

Table 11 sets out the data-sets used, mapped to the research questions and stages of work. Each dataset is unique, with no overlap between datasets. As illustrated, most of the data analysed is generated from the Third Space Learning (TSL) VCE platform. Prior to selecting data from the VCE, I determined the scope of the data collection, and filtered sessions to exclude data that was out of scope:

- The VCE provides SAT and diagnostic programmes for learners; SAT sessions prepare learners for standardised tests, while the diagnostic programmes focus on curriculum areas where learners require more support (as identified by diagnostic tests and teachers). As the two programmes offer uniquely structured sessions, I chose to include only diagnostic programme sessions in my analysis, to enable fair comparison. The diagnostic programme was chosen as the sessions offer richer opportunities for learners to engage with the content, and demonstrate their regulatory skills, while the SAT sessions are more focused on exam readiness.
- I excluded sessions conducted with learners who had English as a second language (ESL), or additional learning needs. Tutors may adapt their approach while coaching ESL learners or learners with additional needs, and the inclusion of these sessions may distort the overall results from the analysis.
- I also excluded any sessions which were reported to have technical issues with the audio or video quality, or where the session was not completed (e.g. it was interrupted by the classroom teacher).

Table 11. Data sets mapped to research questions.

Research question	Stage of work	Data-set label	Source of data-set
RQ1	Exploratory data analysis to develop contextualised model of SRL	A	Data collected from 15 sessions played back from the VCE (random selection)
RQ2	Development of a framework of signifiers to represent tutor co-regulation of learners in a VCE	B	Data collected from 50 sessions played back from the VCE (random selection)
		C	Observations of learners, collected from 30 sessions (random selection of learners from pool of participating schools and learners)
		D	Data collected from 25 sessions to evaluate the inter-rater reliability of the framework of signifiers
RQ3	Implementation of learning analytics (including process mining, sequence mining and decision trees)	E	Data collected from 59 sessions conducted by a mixture of high and low-performing tutors, where the criteria for tutor performance is human evaluator rankings
		F	Data collected from 41 sessions conducted by a mixture of high and low-performing tutors, where the criteria for tutor performance is learning outcomes (i.e. learner average progress scores)
	Analysis of the process maps, sequences and decision trees	G	Data collected from interviews with 30 tutors, including a mixture of high and low-performing tutors (based on human evaluator rankings)

Research question 1

3.2.2 Data set A

Data-set A refers to the recordings of 15 sessions which were selected at random from the VCE and played back from the platform.

The session recordings include both audio and video data:

- The audio data includes the natural unstructured dialogue between a human tutor and learner
- The video data includes:
    - An interactive whiteboard which presents slides, including the maths content and problems that the learners will work through
    - Learners and tutors can write, erase, draw, and highlight on the whiteboard. They can also move their cursor over the whiteboard
    - Tutors are also able to offer the learners effort points, emojis, stickers and pictures to incentive and reward them.
    - Tutors and learners can also choose to use a chat function which is displayed to the side of the whiteboard.
    - There is no video of tutors or learners themselves.

I have also provided the breakdown of data sample *A* in table 12.

Table 12. Breakdown of data sample A (used to conduct exploratory data analysis, and build the contextualised model of SRL.

Number of session recordings	**Number of tutors**	**Number of learners**	**Length of audio (minutes)**
15	14	15	748

Research question 2

3.2.3 Data set B

Data-set B refers to the 50 sessions played back from the VCE, which were used to build and refine the framework of signifiers. After applying filters to exclude data that was out of scope, the sessions were selected at random.

Table 13 below provides the breakdown of the data sample *"B"*.

Table 13. Breakdown of data sample B (used to build the framework of signifiers).

Number of session recordings	**Number of tutors**	**Number of learners**	**Length of audio (minutes)**

50	46	48	2,450

3.2.4 Data set C

Data-set C refers to the 30 sessions observed to validate the framework of signifiers, from the learner's perspective. The 30 sessions were observed across 6 schools, with the learners observed being selected at random from the pool of learners who were engaged in sessions, and had completed forms consenting to participate in the research. Each session was observed for approximately ten minutes.

Table 14 below provides the breakdown of the sessions observed:

Table 14. Sessions observed to validate the framework of signifiers (from the learner perspective).

Number of schools	**Number of learners**	**Number of sessions**	**Length of observations (minutes)**
6	30	30	300

3.2.5 Data set D

Data-set D refers to the sessions analysed to evaluate the inter-rater reliability of the framework of signifiers. Events were manually coded from these sessions by both raters, prior to the inter-rater reliability being evaluated.

Table 15, Data-set D (sessions analysed to evaluate inter-rater reliability).

Number of session recordings	**Number of tutors**	**Number of learners**	**Length of audio (minutes)**
25	23	25	875

Research question 3

3.2.6 Data set E

In order to develop the data-set for the implementation of learning analytics, I collated a sample of VCE sessions, conducted by high-performing and low-performing tutors. Bannert et al. (2014) found that sampling learners by performance enabled them to highlight structural differences in their learning behaviour. I therefore sampled tutors by performance, using as criteria human evaluator scores for tutor ability to promote student-centred teaching. The definition for student centred teaching is set out in the rubric below, and is similar to self-regulated learning (although not identical). The criteria for scoring SCT are provided below in Table 16.

Table 16. Criteria for human evaluators to assess tutor skill at promoting SCT (student-centred teaching), mapped to co-regulation.

Student-centred teaching rubric	Overlaps with co-regulatory tutoring practices
How effectively (quality, appropriateness, necessity) did the tutor pre-assess the student to identify student needs?	Effective diagnosis enables the tutor to pitch the learning at the appropriate level, enabling a degree of learner independence within the learning process
How accurately did the tutor understand the student needs?	
How well did the tutor adapt to the student needs? - Breaking down concepts - Scaffolding e.g. using learning aids - Using different methods - Adjusting pace - Addressing gaps in knowledge - Follow up questions	Adapting to the learner needs can support deeper cognitive processing by learners, which forms part of effective self-regulation
How effectively (quality, appropriateness, necessity) did the tutor post assess the student to check for understanding?	Post-assessment enables the tutor to pitch the learning at the appropriate level, enabling a degree of learner independence within the learning process
How well did the tutor promote conceptual understanding over computational methods?	Promotion of conceptual understanding over computational methods relates to deeper cognitive processing which forms part of effective self-regulation
How well did the tutor promote pupil independence and autonomy?	Promotion of learner independence is a fundamental aspect of co-regulation
How well did the tutor use specific praise/motivation to positively impact the student mindset? (Including use of effort points)	Learner affect, well-being and mindset are recognised as key contextual factors which influence how effectively a learner can self-regulate
How well did the tutor demonstrate listening skills in the session? (Is the tutor aware of everything that happens during the session?)	Effective listening skills are relevant for appropriate diagnosis, adaptation, and personalisation of learning, all of which

	can influence learner engagement and likelihood of effective self-regulation.
How well did the tutor discuss real life situations in the lesson?	Learner motivation is recognised as a key contextual factor which influences how effectively a learner can self-regulate. Discussing real-life situations can enhance the perceived value of the learning and thus motivation.
How much value was the tutor able to add to the student's learning (new learning, new strategies, confidence, fluency)?	Modelling learning strategies can help a learner to effectively self-regulate
How well did the tutor use variation in tone to keep the student engaged?	Learner motivation is recognised as a key contextual factor which influences how effectively a learner can self-regulate. Discussing real-life situations can enhance the perceived value of the learning and thus motivation.
How well was the tutor able to use interests and previous knowledge about the student to impact interaction?	
How well was the tutor able to extend (challenge, making links) the student's learning? (If applicable)	Extending learning where appropriate can boost learner motivation and engagement, and encourage self-regulation
How well did the tutor manage any SEN needs of the student? (If applicable)	N/A
How well did the tutor demonstrate different engagement techniques? (If applicable)	The range of engagement techniques used by the tutor can influence depths of cognitive processing, and effective self-regulation
How well was the tutor able to 'turn around' engagement and interest of the student? (If applicable)	Learner motivation is recognised as a key contextual factor which influences how effectively a learner can self-regulate.
How well did the tutor personalise learning based on the student interests? (If applicable)	Personalising learning to student interests can boost student motivation and effective self-regulation

As illustrated in Table 16, there is a significant overlap between student-centred teaching (SCT) and effective co-regulatory tutoring practices. Based on the Winne and Hadwin (1998) model of self-regulation, effective self-regulation includes cognitive, metacognitive and affective elements. Co-regulatory instructional cues are included within the architecture of the model as external context which influences the cognitive, metacognitive and affective dimensions of self-regulated learning. Tutoring practices which shape learner engagement and cognitive processing, such as SCT practices, can therefore be

regarded as linked to co-regulation. While there is significant overlap between SCT practices and effective co-regulation, it should be noted that the SCT rubric in Table 16 is mainly focused on cognitive dimensions of learning, and to a lesser extent, affective dimensions. There is limited focus on the how instructional cues can co-regulate the metacognitive dimensions of SRL (e.g. through encouraging goal-setting or reflection).

Figure 9 presents the distribution of scores awarded by human evaluators to tutors for promoting SCT. These scores were awarded at the end of academic term 1 (2019). They were awarded by human evaluators, following a review of 1 or 2 sessions per term (with a maximum of 20 minutes being reviewed per session). The scores are awarded on a fixed scale, with evaluators being able to score tutors with regard to SCT as achieving either 0%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% or 100%.

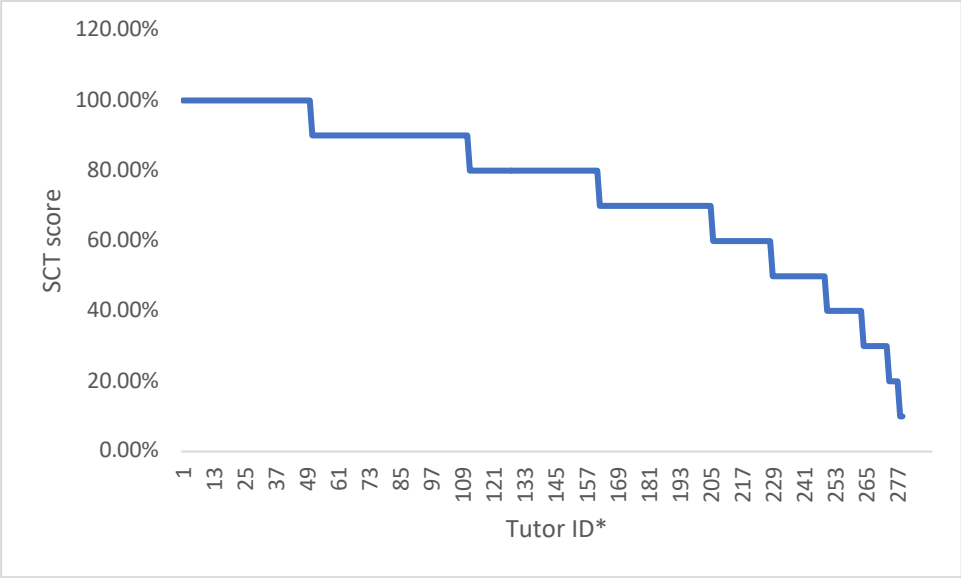


Figure 9. Tutor scores for promoting student-centred learning, awarded by human evaluators *tutor ID's have been anonymised

Out of a total sample of 277 tutors, 50 tutors achieved the highest SCT score of 100%. 15 tutors were randomly selected from this pool of 50 tutors to form the 'high-ranked tutors' data sample. 16 tutors achieved an SCT score of 30% or lower. These tutors were used to build the 'low-ranked tutors' data sample.

59 tutoring sessions (in total) were randomly selected for the selected tutors, and the relevant data was collected and recorded for these sessions. Each tutoring session consists of introductory/warm up, objectives, problem solving and reflection slides, with a flexible order between parts of the session. To ensure that the sessions would be comparable, we only collected data for the problem-solving sections of the session.

The breakdown of the data sample E is provided in table 17 below.

Table 17. Data-set E (used for implementation of learning analytics) *based on human evaluator rankings for SCT.

Tutor category*	**Number of session recordings**	**Number of tutors**	**Number of learners**	**Length of audio (minutes)**
High-ranked tutors	29	15	29	1022
Low-ranked tutors	30	16	30	1054

The data was analysed to ensure that there was no skew in prior levels of learner attainment between the high-ranked and low-ranked tutor data sample. There are two main indicators of prior levels of learner attainment:

- Prior levels of learner attainment may be indicated by the learner scores on the pre-programme diagnostic quiz for the relevant learning outcomes. Learners are requested to answer 2 questions categorised as being at different levels of complexity for each learning outcome, in the pre-programme diagnostic quiz. However, answering the diagnostic questions is not mandatory. In my sample, 16 (out of 29) of the learners in the high-ranked tutor dataset provided answers to the diagnostic questions for the relevant learning outcomes. Similarly, only 17 (out of 30) of the learners in the low-ranked tutor dataset provided answers to the diagnostic questions for the relevant learning outcomes.

- Prior levels of learner attainment may also be indicated by the learner's school teacher, who indicate the level at which learners are working in maths, prior to the programme starting. However, this indicator is for maths in general, and is not learning outcome specific

To start with, I have analysed the available data on the learner diagnostic scores. In order to compare learner performance, I have awarded points based on the number of questions answered correctly, with more points being awarded for more complex questions:

- One point is awarded for correctly answering the question categorised as being at the 'working at' level (i.e., relatively less complex questions)
- Two points are awarded for correctly answering the question categorised as being at the 'exceeding' level (i.e. relatively more complex questions)

Based on the above approach, the learner pre-programme diagnostic scores are calculated. The chart below shows the learner pre-programme diagnostic scores for the high-ranked tutor data sample, and the low-ranked tutor data sample. As demonstrated, there does not seem to be a skew towards higher levels of prior attainment for the high-ranked tutor data sample, with most learners attaining diagnostic scores of either 0 or 1, when the diagnostic questions for the relevant learning outcomes were answered.

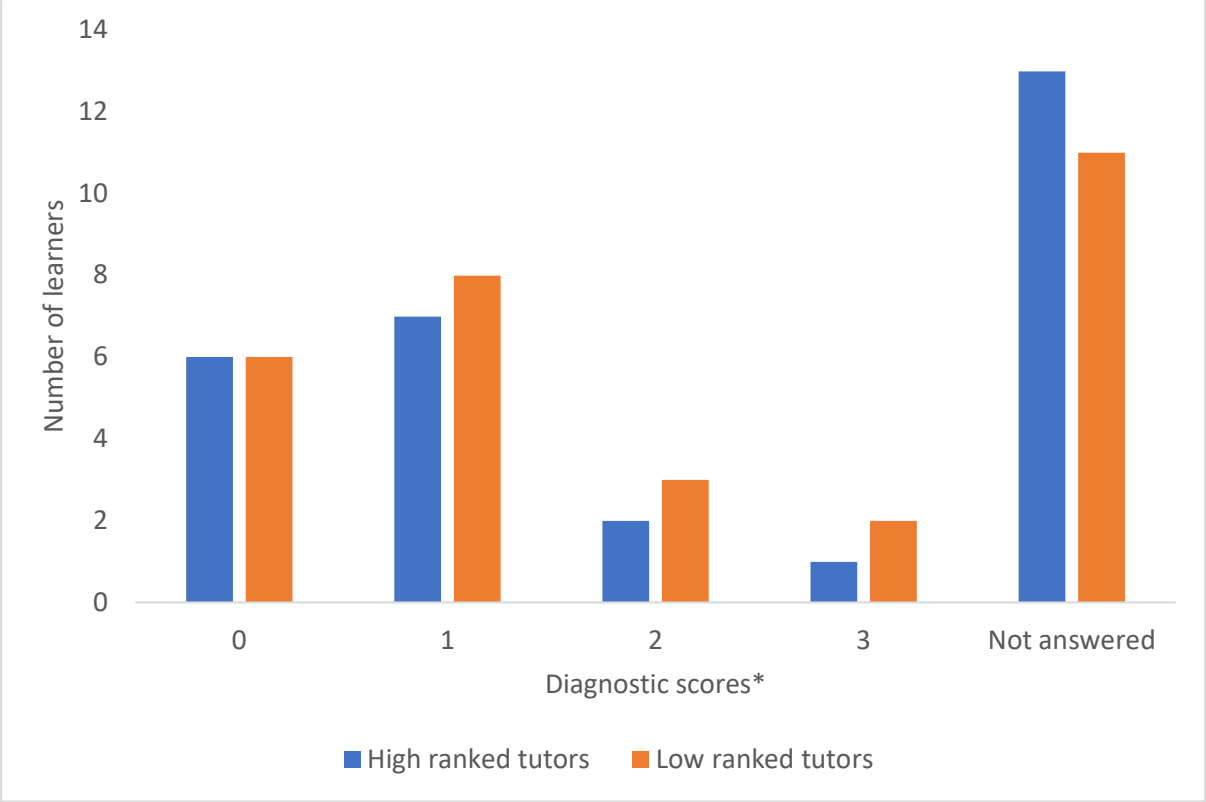


Figure 10. Learner diagnostic scores for high-ranked and low-ranked tutor dataset

Following the analysis of learner diagnostic scores, I compared the teacher-indicated levels of learner attainment, for the high versus low-ranked tutor datasets. School teachers can indicate learners as working at one of the following levels in maths, prior to starting the programme:

- Working Towards Expectations
- Meeting Expectations
- Exceeding Expectations

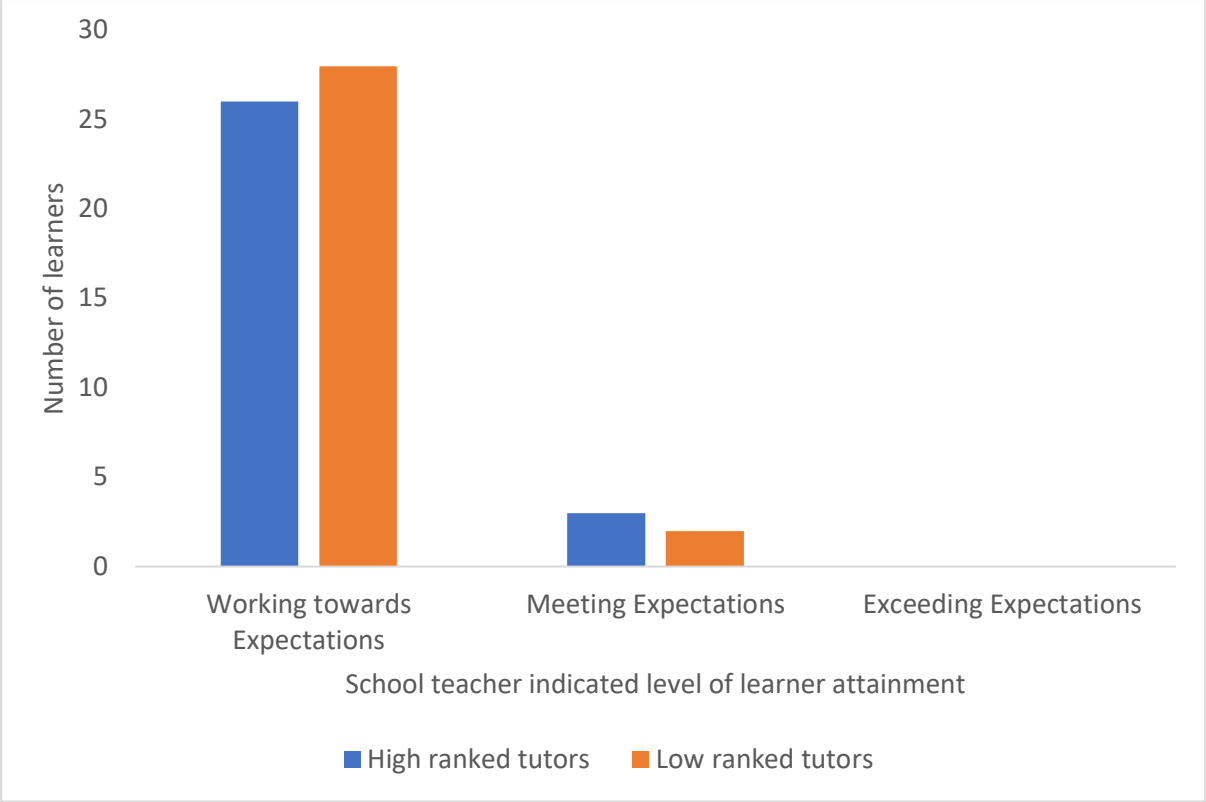


Figure 11. Learner levels of prior attainment, as indicated by school teachers

As demonstrated by Figure 11, most learners in both the high-ranked, and low-ranked tutor datasets, are indicated as working at the 'working towards' level, prior to starting the programme. There is no demonstrable skew towards higher levels of prior learner attainment, in the high-ranked tutor dataset.

3.2.7 Data set F

While SCT is a similar construct to SRL, it is not identical to self-regulation. In addition, the human evaluator rankings are generated using a relatively small sample of data, and may not be reflective of tutor skills in general. The human evaluator rankings are also subject to human bias and inconsistencies. I therefore experimented with the use of other criteria to assess tutor performance in promoting learner SRL: learning outcomes.

Learning outcome scores have been assessed on the basis of average learner progress scores. A '*progress score'* has been defined as the difference between the learner score on the diagnostic questions for a specific learning outcome ('*score a'*), and the learner score on the

post session quiz for that same learning outcome (*'score b'*). I have calculated *score a* and *score b* as follows:

- *Score a* has been calculated by summing together the points awarded for the pre-programme diagnostic quiz. The learner answers two questions for each learning outcome in the diagnostic quiz, with each question being categorised at a different level of complexity. I have awarded points for questions answered correctly in the pre-programme diagnostic quiz, with more points awarded for answering questions which are categorised as being more complex:
    - One point is awarded for correctly answering the question categorised as being at the 'working at' level (i.e., relatively less complex questions)
    - Two points are awarded for correctly answering the question categorised as being at the 'exceeding' level (i.e. relatively more complex questions)
- Similarly, *score b* has been calculated by summing together the points awarded for the post session quiz. The learner answers two questions at the end of a session, for a learning outcome, with each question being categorised at a different level of complexity. I have awarded points for questions answered correctly in the post session quiz, with more points awarded for answering questions which are categorised as being more complex:
    - One point is awarded for correctly answering the question categorised as being at the 'working at' level (i.e., relatively less complex questions)
    - Two points are awarded for correctly answering the question categorised as being at the 'exceeding' level (i.e. relatively more complex questions)
- The learner progress score for a specific learning outcome is calculated through subtracting *score a* from *score b (score b – score a)*. As the learner progress score is

for specific learning outcomes, there may be multiple learner progress scores for the same learner.

- The *average learner progress score* for a tutor is calculated by taking the average of all the learner progress scores, for learning outcomes which have been taught by that tutor

The sessions database was filtered to only include sessions, for which a learner progress score could be calculated (i.e. the learner had answered diagnostic questions and post session quiz questions). To enable a fair comparison between tutor performance, learning outcomes with a student pre-programme diagnostic score that is greater than 1 were excluded from the analysis. In addition, the database was further filtered to only include tutors who have at least 10 data points across which the average learner progress score has been calculated.

After following the approach described in the paragraph above, the distribution of average learner progress scores set out in Figure 12 was obtained. 10 tutors with average learner progress scores that are greater than 1.50 were selected for the high-performing tutor dataset, and 10 tutors with average learner progress scores of lower than 0.55 were selected for the low-performing tutor dataset.

For each tutor, at least 2 sessions were selected at random for analysis from the filtered session database, with 21 sessions being recorded for high-performing tutors, and 20 sessions being recorded for low-performing tutors. The relevant data was collected and recorded for these sessions. Each tutoring session consists of introductory/warm up, objectives, problem solving and reflection slides, with a flexible order between parts of the session. To ensure that the sessions would be comparable, we only collected data for the problem-solving sections of the session.

Table 18 provides the breakdown of data-set F.

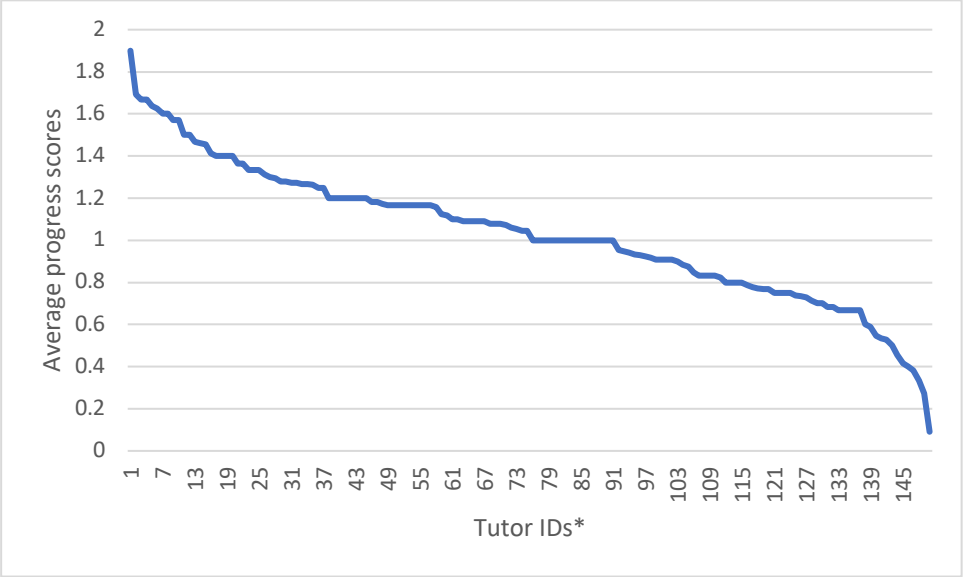


Figure 12. Distribution of average learner progress scores for tutors *tutor ID's have been anonymised

Table 18. Data-set F (used for implementation of learning analytics) *based on average learner progress scores.

Tutor category*	**Number of session recordings**	**Number of tutors**	**Number of learners**	**Length of audio (minutes)**
High-performing tutors	21	10	21	693
Low-performing tutors	20	10	20	640

The learner data was reviewed to ensure there was no skew in prior levels of learner attainment between the high APS tutor and low APS tutor sample. As described earlier in this section, the session database was filtered to only include sessions with diagnostic scores of either 0 or 1. Comparing the diagnostic scores for high versus low APS dataset does not illustrate a skew towards a specific score for either dataset, as per Figure 13.

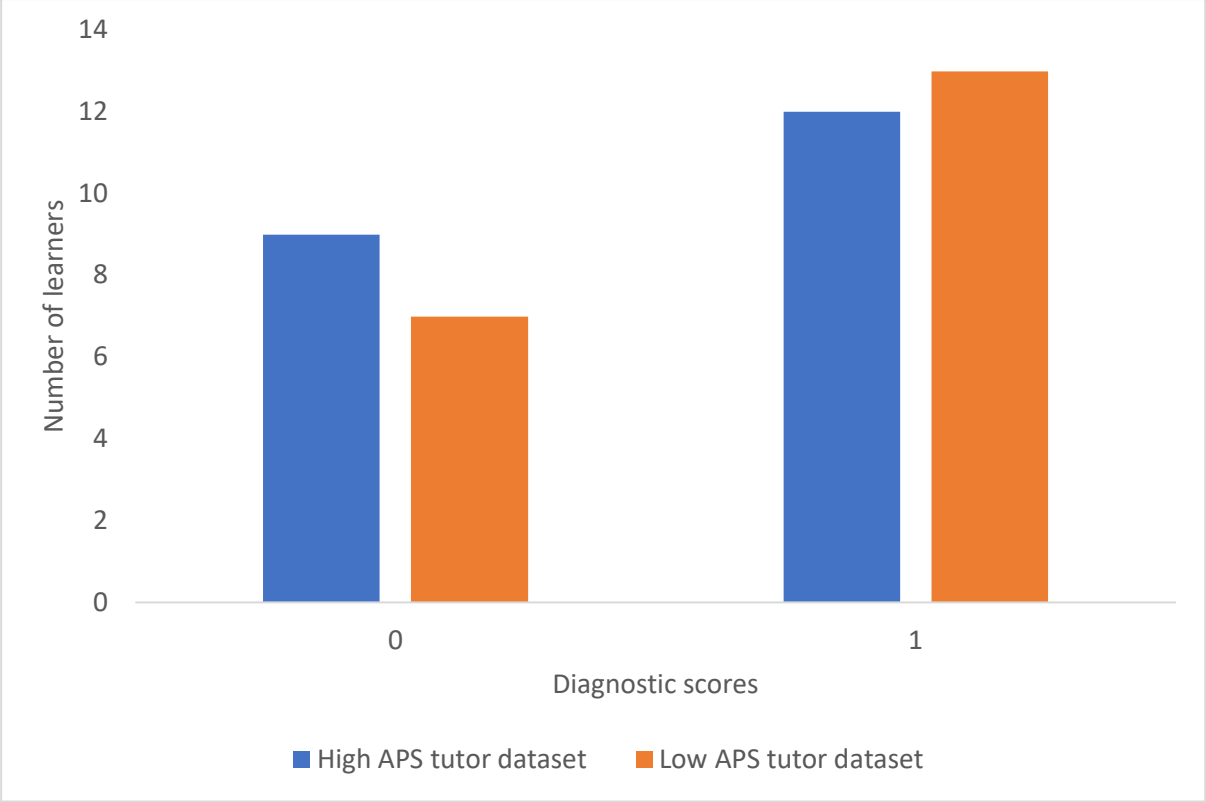


Figure 13. Learner diagnostic scores for high APS and low APS tutor dataset

In addition, there is also no demonstrable skew in prior levels of learner attainment, as indicated by school teachers prior to learners commencing the programme. Figure 14 illustrates that the vast majority of learners in both datasets are at the 'Working Towards Expectations'.

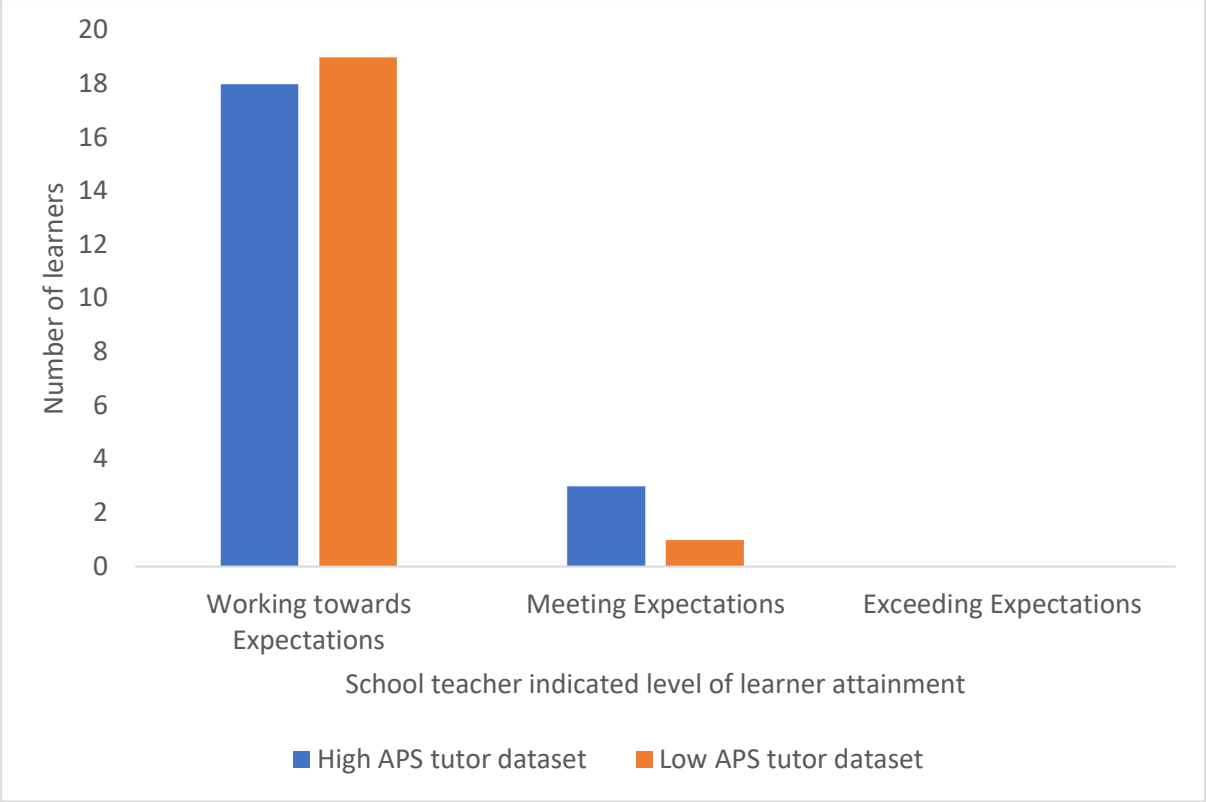


Figure 14. Learner levels of prior attainment, as indicated by school teachers

3.2.8 Data set G

In addition to analysis of the data generated from online VCE sessions, interviews were conducted with high-ranked and low-ranked tutors. The aim of these interviews was to obtain insight into the perceptions and beliefs held by tutors regarding self-regulated learning, and triangulate this against the tutor behaviours observed during online VCE sessions. This would help (1) validate the results obtained from analysing the VCE sessions, and (2) enable a deeper, more rounded understanding of the differences between high-ranked versus low-ranked tutors with regard to SRL.

The tutor interviews were conducted between January-May 2022, using the following approach:

- The most current version of the tutor rankings was used to compile a list of 30 tutors who scored in the top 5 percent and bottom 5 percent for student-centred teaching. 15 tutors were randomly sampled from the group which achieved in the top 5 percent of

scores for SCT, and 15 tutors were randomly sampled from the group which achieved in the bottom 5 percent of scores for SCT.

- Tutors were contacted to obtain their consent to being interviewed and the resulting data being used for research purposes.
- Interviews were set up and conducted over Zoom video-conferencing. Interviews were transcribed, with each interview lasting for approximately 20 minutes

Table 19. Overview of the tutor interview programme

Tutor ranking	**Number of tutors**	**Interview format**	**Total length of interviews (minutes)**
High	15	Zoom video-conferencing	310
Low	15		322

The interviews were conducted on a semi-structured basis, with the help of a discussion guide. The discussion guide was drafted by myself and reviewed by experts within the field of self-regulated learning. The discussion guide was organised to generate tutor insights on particular areas of interest, which would then be analysed through a thematic analysis approach. The relevant themes from the discussion guide are summarised in the table below, alongside examples of the questions presented to tutors, see Appendix L.

Table 20. Themes underlying the tutor interviews, and the analysis conducted of the interview output.

Theme	**Examples of questions presented to tutors**
Contextual questions	Level of experience in teaching (online and offline)
Epistemological beliefs and perceptions of SRL	Beliefs about the nature of the learning process (e.g. role of tutor versus role of learner) Understanding of the terms ‘student-centred learning’ and ‘self-regulation’ Perceived importance of developing SCT and learner SRL, through tutoring practices
Tutoring practices with regard to SRL	Perceived tutoring practices adopted to develop SCT and SRL Perceptions on how tutoring practices are adapted to the specific needs of the learner

Degree of comfort with online tutoring, with regard to SRL	Preference towards teaching online or face to face Ease of promoting SCT and SRL online versus face to face
Role of the VCE in tutoring sessions, and to promote learner SRL	Extent to which the online learning platform supports promotion of SCT and SRL Role of the online learning platform, within sessions

Thematic analysis of interview output

Thematic analysis was implemented, using a deductive approach informed by the contextualised model of SRL and theory-based framework of signifiers.

Thematic analysis followed the below process:

1. **Familiarization with the data.** Following the completion of the interview programme, interviews were recorded and transcribed. I reviewed the transcripts to familiarise myself with the output from the interviews
2. **Coding the data**. Informed by the contextualised model of SRL and theory-based framework of SRL, a deductive approach was taken towards thematic analysis. The types of analysis which would be conducted with interview output were considered, and are set out in Table 21. In light of the potential forms of analysis, a set of codes was generated (e.g. level of tutor experience, tutor growth mindset, tutor interpretation of SRL). The interview output was highlighted to match the codes, and the set of codes was further iterated as the interview output was analysed to ensure that they were comprehensive
3. **Building and refining themes** that emerged from the interviews. The codes were reviewed and categorised into various themes. The interview output was re-examined to ensure that the categories of themes were comprehensive, and that they effectively captured the nuanced output from the interviews

4. **Writing up** the output from the interviews. The thematic output was analysed. The implications of the interviews were considered to (1) validate the results obtained through the implementation of learning analytics, and (2) explore the co-regulatory differences between high-ranked versus low-ranked tutors.

Table 21. Interview analysis.

Theme	**Examples of questions presented to tutors**	**Analysis conducted**
Contextual questions	Level of experience in teaching (online and offline)	Differences in the teacher responses to interview questions, depending on the degree of teacher experience
Epistemological beliefs and perceptions of SRL	Beliefs about the nature of the learning process (e.g. role of tutor versus role of learner) Understanding of the terms 'student-centred learning' and 'self-regulation' Perceived importance of developing SCT and learner SRL, through tutoring practices	Comparison of tutor perceptions with regard to the nature of learning. To what extent do tutors emphasise the active involvement of learners within sessions? To what extent is a fixed versus growth mindset demonstrated? How does this vary between high and low-ranked tutors? Differences between high-ranked versus low-ranked tutors, in the sophistication of their understanding of SCT and SRL Differences between high-ranked and low-ranked tutors in the perceived importance of co-regulating learners
Tutoring practices with regard to SRL	Perceived tutoring practices adopted to develop SCT and SRL Perceptions on how tutoring practices are adapted to the specific needs of the learner	Comparison of the perceived teaching practises adopted by tutors (from interview data) with the observed tutoring practices (using session data). To what extent do the tutor perceptions align with the observed behaviours? How does this vary between high and low-ranked tutors?
Degree of comfort with online tutoring, with regard to SRL	Preference towards teaching online or face to face Ease of promoting SCT and SRL online versus face to face	Differences between high versus low-ranked tutors, in their level of comfort in promoting SRL online. Are high-ranked tutors more aware of the potential difficulties, and have they developed mitigating strategies?

Role of the VCE in tutoring sessions, and to promote learner SRL	Extent to which the online learning platform supports promotion of SCT and SRL Role of the online learning platform, within sessions	Comparison of how the use of the online learning platform varies between high and low-ranked tutors, in promoting learner SRL. Do high-ranked tutors demonstrate a greater degree of familiarity with platform tools? Comparison of the role of the online learning platform, in VCE sessions conducted by high-ranked versus low-ranked tutors. To what extent are sessions led by the structure of the online learning platform, versus being led by the tutor in response to learner needs (using the platform as is appropriate)?

Chapter 4 Developing signifiers for SRL

4.1 Contextualised model of SRL

Introduction

My research aims to address my overarching research question from a theoretically grounded perspective. My research explores the differences between how high-performing tutors co-regulate learners versus low-performing tutors, in a VCE. Given the lack of research exploring tutor co-regulation of learners in a Virtual Classroom Environment (VCE), there is a need to develop a theoretical approach, which can be used as the foundation for analysing co-regulatory processes. In particular, it is essential to consider how SRL should be conceptualised to explore the influence of human tutors on learner SRL processes in a VCE, as per RQ1. My contribution towards addressing this research question is set out in this chapter.

Section 2.1 of this thesis proposed a process-oriented conceptualisation of SRL to explore the dynamic influence of human tutors on learner SRL processes in a Virtual Classroom Environment (VCE), aligned with current research into SRL. I will build on this broad conceptualisation of SRL to propose a contextualised model of SRL. I follow the structure outlined below:

- I will begin by analysing the theoretical frameworks used in past studies, for their suitability to the studied context, and justify the need to develop a contextualised model and theoretical framework for my research
- I will then review established models of SRL to judge their suitability for the studied context, and select a model for contextualisation
- I will refine the chosen model for the studied context through exploratory data analysis and interpretation of the Winne and Hadwin (1998) for my study context

- I will validate the contextualised model by mapping relevant literature to the elements of the model
- The final iteration of the contextualised model of SRL is presented in section 4.1.3

4.1.1 Theory-based frameworks implemented in previous studies

In Section 2.2 of this thesis, I explored the literature to identify studies of co-regulation and identify theory-based frameworks that have been designed and implemented to study co-regulation.

Section 2.2 presented papers judged as relevant, setting out the research context, types of data collected and theoretical approach adopted. As demonstrated, the majority of studies conducted on co-regulation have been in the context of an offline, traditional classroom environment. With regard to scaffolding by online tutors, work has explored the influence of pedagogical agents in scaffolding learner SRL, but limited work has been done in the context of human-to-human tutoring.

The majority of the studies reviewed either didn't apply a theoretical framework for analysis, (preferring to explore findings in light of the relevant literature) or developed theoretical frameworks for the context of a traditional classroom environment. In the latter case, frameworks have been tailored to capture macro events and processes, rather than fine grained data about teacher-learner interactions.

Researchers are beginning to recognise the need for theoretical models and frameworks which enable the concept of co-regulation to be operationalised and explored in practice. To this end, there have been some recent studies which have proposed more granular frameworks of co-regulation. For example, Andrade and Brookhart (2020) developed a theoretical framework to assess co-regulation through Classroom Assessment (CA), and conducted an exercise mapping the literature to each element of the framework.

They began by broadly defining CA to include assessments conducted in class, as well as assessment through student-teacher dialogue. The Pintrich and Zusho (2002) model of SRL was adopted and expanded to consider the potential ways in which CA may co-regulate learning. Andrade and Brookhart (2020) chose to adopt the Pintrich and Zusho (2002) model as the basis for their work, as it takes a comprehensive approach (phases of SRL), and also considers SRL by area (cognition, motivation, behaviour, context). This would enable the influence of CA on SRL to be considered from a range of perspectives. Previous studies were then reviewed and mapped to each part of the framework to validate the theoretical approach.

While the framework proposed by Andrade and Brookhart (2020) is not entirely relevant for my context, there are elements of the framework which can be used to inform my work. The framework was developed for investigating co-regulation through CA, rather than examining the influence of broader tutoring practices. However, there are specific teaching practices (such as informal assessment or feedback), which act as forms of CA and are thus covered by the framework. In addition, the framework is validated solely through mapping relevant literature to the elements of the framework. While no empirical data is used to refine the framework, the mapping of the literature review to the framework is useful for identifying indicators of certain co-regulatory teaching practices.

As explained in 2.1 of this thesis, a process-oriented conceptualisation of SRL has been adopted, to explore the evolution of tutor influence on learner SRL during a learning activity. In order to operationalise this conceptualisation of SRL, a granular and holistic framework is needed. I am not aware of any existing frameworks that are suited to this context, and propose a contextualised model and framework in this chapter. My approach towards building this framework is set out below.

4.1.2 Approach for building contextualised model of SRL

In my study, I will take the following steps to build contextualised model of SRL

a. Identify and adopt a suitable model of SRL
b. Interpret the model for my studied context
    i. Conduct exploratory data analysis to develop an interpretation of the model
    ii. Map relevant literature to the contextualised model to validate my contextualisation

(a) Identify and adopt a suitable model of SRL

In section 2.1 of this thesis, I reviewed established models of SRL. In this chapter, I will consider the suitability of each of these models as a basis for my research.

I have adopted a model which is sufficiently granular to capture fine grained data related to tutor-learner interactions in a VCE. Given that limited research has been conducted into my studied context to date, it is unclear which factors are likely to be influential in the dynamics of tutor-learner interaction. I would therefore like the model to be holistic in nature, to ensure that the relevant elements of the tutor-learner interactions are captured. Finally, as described in chapter 2.1 of this thesis, my conceptualisation of SRL is process-oriented, and I am seeking to explore the evolution of tutor-learner interactions over the duration of the learning activity. To this end, I am seeking a model which reflects the dynamic, and cyclical nature of SRL.

The review of selected established models of SRL (set out in 2.1 of this thesis) highlights that established models agree on fundamental elements of self-regulated learning (e.g. active approach to learning, intentional use of learning strategies, cyclical nature of SRL, influence of context); however, there are key differences between the models in terms of the overall structure of the frameworks, the role of task and learner conditions within the frameworks, the fluidity of the frameworks, and the aspects of self-regulated learning that are emphasised. For example, while all established models recognise the influence of task and learner conditions on SRL, they differ in whether conditions are integrated into the model as

a component of SRL (Winne & Hadwin, 1998), or are recognised as distinct areas within which SRL operates (Pintrich & Zusho, 2002). These differences have consequences for the suitability of the model for exploring SRL from different perspectives. Given that I am seeking to explore the influence of tutors on learner SRL, a model which integrates external conditions (such as tutor practices) into the model as a component of SRL may enable a more rigorous exploration of how tutors can influence learner SRL.

Another key difference between the Winne and Hadwin (1998) model, versus the Zimmerman and Pintrich models, is the nature of the role that monitoring and evaluation play. While monitoring and evaluation is seen as a key element of SRL in all three models (Azevedo et al., 2011; Governor & Wetzel, 2013; Greene & Azevedo, 2007a; Zimmerman, 2000), the Winne and Hadwin model is distinct in that it considers the impact of monitoring and evaluation at a component level, rather than examining impact by phase of SRL (Azevedo et al., 2011). This is a more granular approach, which enables analysis of how changes on one component of SRL (e.g. task conditions) can lead to changes in other components (e.g. type of operations used by learner). Further, the model is more fluid than other models. While all models are weakly sequenced, the Winne and Hadwin model provides a framework for analysing how a learner can move back and forth within a phase of SRL, as well as between phases (Greene & Azevedo, 2007a, 2007b).

The Winne and Hadwin model has been noted as being suited to the analysis of online data (Winne, 2010; Winne et al., 2013), as it synthesises all the various components of SRL from the literature into a heuristic framework that can be applied to educational settings rather than laboratory settings that have been specifically designed to explore SRL. By examining the information processes that occur within each phase, and how these are driven by the various components of SRL, the Winne and Hadwin model provides a rigorous framework for detecting SRL in a manner which reflects its fluidity.

Given the distinguishing features of the Winne and Hadwin model that are described above, I propose to adopt the Winne and Hadwin model of self-regulated learning as the basis for my work. However, my interpretation and adaptation of this model will be influenced by my review of the Zimmerman (2000) and Pintrich (2000) models. I incorporated the SRL variables identified in the Zimmerman and Pintrich models, and the research that has been conducted on these models, into the framework of signifiers. For example, I will incorporate the variables identified by Zimmerman as forming part of the self-reflection phase including self-evaluation and self-reactions (Zimmerman & Kitsantas, 1996, 1997), into a framework of signifiers that aligns with the Winne and Hadwin model.

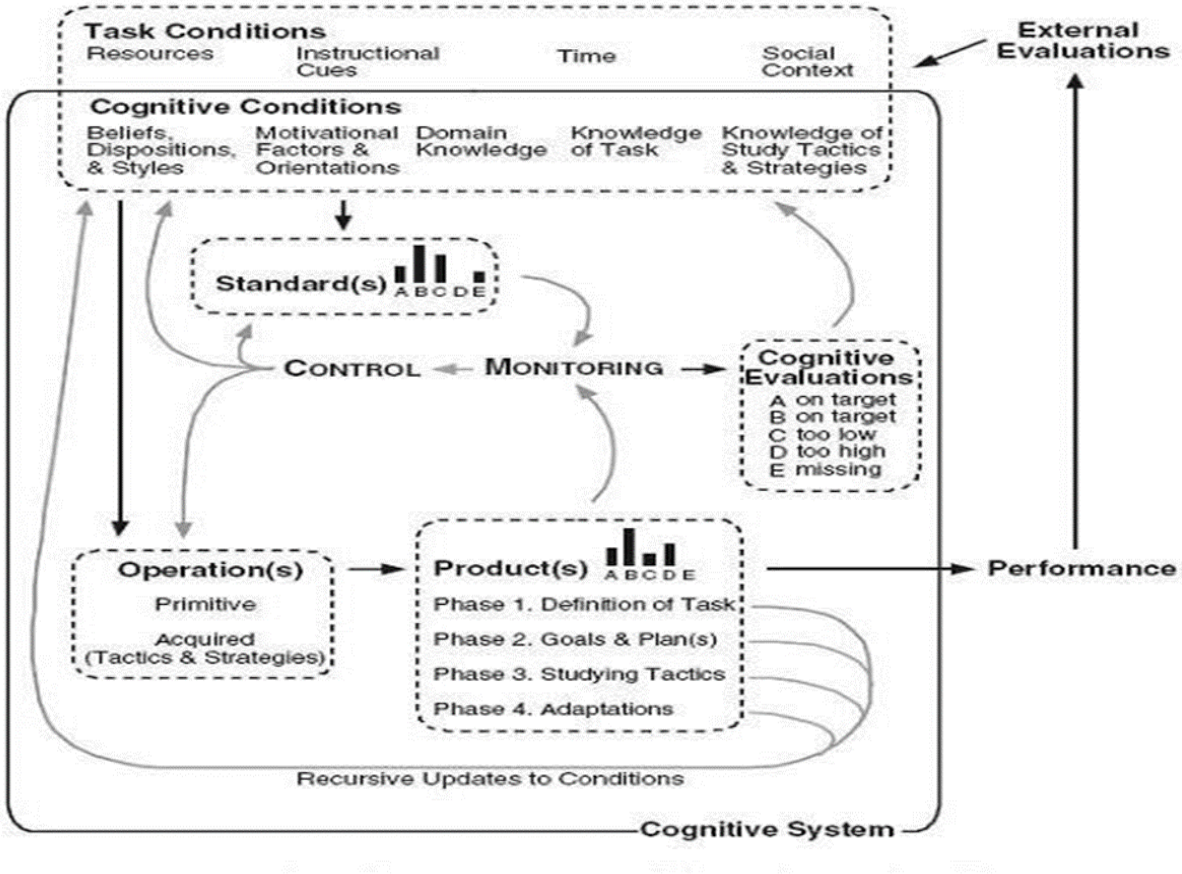


Figure 15. Winne and Hadwin's model of SRL. From "Studying as Self-Regulated Learning," by P. H. Winne and A. F. Hadwin, 1998, in D. J. Hacker, J. Dunlosky, and A. C. Graesser (Eds.), *Metacognition in Educational Theory and Practice*, Routledge (https://doi.org/10.4324/9781410602350-19).

(b)(i) Interpret the model for my studied context

The next step is to interpret the chosen model for my studied context. As my research focuses on understanding how tutors co-regulate learners in a VCE, rather than examining learner SRL in isolation, I begin by analysing the components of the model which align to my research questions. The relevant components will represent how tutor practices can influence learner SRL processes within the broader cognitive architecture of the Winne and Hadwin model.

In the Winne and Hadwin (1998) model, tutor practices are recognised as an example of the Conditions component of the framework, with 'instructional cues' being an external condition that interacts with other components of SRL, such as the internal cognitive conditions of the learner. For example, the nature of tutor praise, and attributions made to effort versus ability, can shape the learners beliefs about the process of learning, and encourage a fixed versus growth mindset (Dweck, 2012). Similarly, learner SRL processes align to Operations component in the framework. Operations are the cognitive and metacognitive processes that learners engage in to regulate their learning. For example, Operations may include the enactment of learning tactics or strategies, or the setting of learning goals and plans. *In order to address my research questions within the context of the Winne and Hadwin model, I must consider how a specific example of an external Condition (tutor practices) can influence learner Operations (SRL processes), while allowing for the inter-related nature of all components within the model.*

While the Winne and Hadwin model recognises the integral relationship between tutor practices and learner SRL operations, the model is silent on the types of tutor practices that can influence learner Operations. In acknowledgment of the inter-related nature of all components of the model, I propose categories of tutor practices which align to each of the components of the model, rather than merely focusing on tutor practices that have the potential to directly influence learner Operations. I have provided examples of potential tutor

practices that map to various components of the Winne and Hadwin model, in Table 22 below. These examples have been developed through:

1. Hypothesising potential tutoring practices which could apply in the context of a VCE, with regard to each component of SRL in the Winne and Hadwin (1998) model of SRL
2. Building and refining the list of tutoring practices by conducting an exploratory data analysis of a sample of sessions from the VCE. 15 sessions were randomly selected from the VCE platform and played back, to support the development of the list of tutoring practices. This exploratory data analysis was useful in highlighting that the tutor style of engagement can influence the level of control that the learner assumes over their learning during the VCE session, and therefore the extent to which the learner will self-regulate.

Table 22. Potential tutoring practices which can co-regulate learners in a VCE.

Components of the Winne and Hadwin framework		**Potential tutor practices that may influence component**
Task Conditions (resources, instructional cues, time, social context)	Resources	N/A (resources centrally set in the VCE)
	Instructional cues	N/A
	Time	N/A
	Social context	N/A
Cognitive Conditions	Beliefs, dispositions, styles	Tutor practices which influenced learning beliefs and mindset
	Motivational factors and orientations	Tutor practices which influenced learner motivation and goal orientation
	Domain knowledge	Tutor practices which can be regarded as forms of informal assessment
	Knowledge of task	Tutor practices which monitored the learner knowledge of the task
	Knowledge of study tactics and strategies	Tutor modelled and prompted learning strategies for the learner
Operations	Primitive	N/A
	Acquired (tactics and strategies)	Tutor practices which influenced shallow versus deep cognitive engagement by the learner
Products	Phase 1- definition of task	Tutor prompted or modelled task definition
	Phase 2 – goals and plans	Tutor prompted or modelled goal-setting and planning

	Phase 3 – studying tactics	Tutor prompted or modelled study tactics
	Phase 4 – adaptations	Tutor prompted or modelled reflection to make fundamental changes to components of SRL
Evaluation	Cognitive evaluations	Tutor prompted the learner monitoring their cognition
Standards	Standards	Tutor prompted the learner to shape the dimensions against which learners monitor themselves

With regard to learner Operations, Winne proposed five categories of Operations in his SMART model (Winne, 2010; Winne et al., 2013). Each operation is a cognitive process through which information is manipulated, and the information input to the process can either be cognitive or metacognitive. The outputs (or Products) of the operation can similarly be both cognitive and metacognitive, e.g. retrieval of information, or a judgement of learning. The SMART Operations are described in Table 23 below (Winne, 2010; Winne et al., 2013), along with examples of traces that were generated in the n-Study learning .platform. The exploratory data analysis described in the sections above was used to consider whether the proposed list of operations adequately captured learner SRL processes, and areas for refinement. I found that the list of operations was comprehensive, as each operation referred to a clearly defined and basic unit of activity. Learner action often appeared to be various combinations of the operations defined.

I therefore propose to use this set of Operations for my studied context, mapping traces from the context to each Operation. The design of the SMART operations sets out clear definitions of processes, which are both theoretically grounded, and capable of being operationalised and tracked using trace data. Patterns of tutor-learner interactions, based on these definitions of learner SRL processes, will enable an exploration of how various tutoring practices can encourage deeper versus shallower forms of engagement, within the context of the Winne and Hadwin model of SRL.

Table 23. SMART operations of SRL as defined by Winne.

Operation	**Description**	**Example of traces (from n-study platform)**
Searching	Directing attention to particular information	Opening successive bookmarks
Monitoring	Comparing information to a set of standards	Highlighting text (e.g., determining that information is important) Selecting a previously made note for review (e.g., determining that further work is needed and forming a judgment of learning)
Assembling	Relating units of information to each other	Tagging Assigning two bookmarks to a relevant folder
Rehearsing	Maintaining or reinstating information in working memory	Reviewing a note Copying/pasting
Translating	Transforming the representation of information	Paraphrasing Describing an equation or diagram in words

Note. Adapted from "Improving Measurements of Self-Regulated Learning," P. H. Winne, 2010, *Educational Psychologist*, *45*(4) (https://doi.org/10.1080/00461520.2010.517150).

In order to bring to life my interpretation of the Winne and Hadwin (1998) model, I present here a narrative example illustrating self-regulation and co-regulation in a VCE. A student (Anwar) worked with his tutor (Salim) on the topic of decimal conversion. Prior to starting a set of questions, Salim asked Anwar how well he understood this topic. This prompt is an example of co-regulation, as Salim is encouraging Anwar to monitor his understanding of the topic. Anwar responded that he had a reasonably good understanding of the topic and thought he could attempt the questions. Anwar worked through the questions, writing, erasing and rewriting his answers a few times. He asked Salim for his help in understanding what one of the questions was asking for. This is an example of self-regulation, as Anwar was monitoring his own understanding without being prompted. In response, Salim asked Anwar a specific, narrow ended question directing his attention towards a certain aspect of the question. This is another example of co-regulation, as Salim chose to adopt a guided form of engagement rather than simply answering the question,

therefore prompting deeper cognitive processing by Anwar. Finally, after Anwar completed the questions, Salim asked Anwar a set of specific, narrow ended questions gradually leading to more open-ended questions. He praised Anwar for his reasoning, demonstrating another aspect of co-regulation (influencing learner Cognitive Conditions).

(b,ii) Validate the interpretation of model

While the SMART model of Operations is an established construct, the categories of tutor practices that can influence learner SRL (as set out in Table 22) needs validation. I propose to validate the categories of tutor practices through mapping relevant literature to each category. The literature review is intended to demonstrate evidence of a relationship between the tutor practice and aspects of learner SRL and is therefore select rather than exhaustive.

Table 24. Potential co-regulatory practices in a VCE mapped to literature.

Components of the Winne and Hadwin framework		**Potential tutor practices that may influence component**	**Relevant literature**
Task Conditions (resources, instructional cues, time, social context)	Resources	N/A (resources centrally set in the VCE)	N/A (resources centrally set in the VCE)
	Instructional cues	N/A	N/A
	Time	N/A	N/A
	Social context	N/A	N/A
Cognitive Conditions	Beliefs, dispositions, styles	Tutor practices which influenced learning beliefs and mindset	(Cimpian et al., 2007)
	Motivational factors and orientations	Tutor practices which influenced learner motivation and goal orientation	(Dweck, 2012)
	Domain knowledge	Tutor practices which can be regarded as forms of informal assessment	(Heritage & Heritage, 2013)
	Knowledge of task	Tutor practices which monitored the learner knowledge of the task	(Andrade & Brookhart, 2020)
	Knowledge of study tactics and strategies	Tutor modelled and prompted learning strategies for the learner	(Dignath & Büttner, 2018; Ewijk et al., 2013; Trevors et al., 2014)

Operations	Primitive	N/A	N/A
	Acquired (tactics and strategies)	Tutor practices which influenced shallow versus deep cognitive engagement by the learner Tutor practices which built inter-subjectivity	(Chi & Wylie, 2014; Hrastinski et al., 2021) (Araya & Aljovin, 2018)
Products	Phase 1- definition of task	Tutor prompted or modelled task definition	(Bell & Pape, 2014; Graesser & McNamara, 2010; Hadwin & Winne, 2001; Winne & Hadwin, 2013)
	Phase 2 – goals and plans	Tutor prompted or modelled goal-setting and planning	
	Phase 3 – studying tactics	Tutor prompted or modelled study tactics	(Trevors et al., 2014)
	Phase 4 – adaptations	Tutor prompted or modelled reflection to make fundamental changes to components of SRL	(Dignath & Büttner, 2018)
Evaluation	Cognitive evaluations	Tutor prompted the learner monitoring their cognition Tutor provided narrative feedback to the learner	(Andrade & Brookhart, 2020)
Standards	Standards	Tutor prompted the learner to shape the dimensions against which learners monitor themselves	(Brookhart & Chen, 2015; Hattie, 2008a)

Limited research has been conducted investigating human tutoring practices that can promote or inhibit learner SRL. The majority of the research conducted to date has focused on traditional classroom environments, or intelligent tutoring systems. The relationship between tutoring practices and SRL remains under-explored. However, there are multiple studies which explore the relationship between tutoring or teaching, the learning environment and aspects of SRL. I utilised these studies to develop hypotheses about the types of tutoring practices that may impact learner SRL operations in the context of a VCE.

Aspects of SRL include learner internal and external conditions, cognitive and metacognitive operations, products of learning, and internal evaluations against internally

held standards. The focus of my research is on how tutoring practices influence learner cognitive and metacognitive operations. However, given the inter-connected nature of the various elements of SRL, tutoring practices which influence other aspects of SRL may indirectly impact learner operations.

Tutor practices influencing cognitive Conditions

I have conducted a detailed literature review to understand how tutoring practices can impact learner internal conditions. For example, there is a vast body of research establishing the relationship between praise and learner motivation. Learner motivation is a key aspect of SRL in all the main theoretical models of SRL (reviewed in section 2.1 of this thesis). In the Winne and Hadwin (1998) model of SRL, motivation is an internal Condition, which interacts with social, environmental and behavioural factors to influence self-regulation. Tutoring practices which impact learner motivation (e.g., through praise), are therefore very likely to impact learner SRL. Praise which reinforces a growth mindset is particularly effective in encouraging learner motivation (Cimpian et al., 2007). Praise for intelligence tends to put students in a fixed mindset and they also tend to opt for easier tasks while praise for effort puts students in a growth mindset and they look for challenging tasks. For example, the study by Cimpian et al. (2007) suggested that positioning an activity as one in which the child is being evaluated ('You're a good drawer') resulted in most children opting to do something else when allowed to. On the other hand, framing an activity as one in which the process or quality of the work is judged ('You did a good job drawing') encouraged children to choose more drawing. Given this body of research, programmes designed to enhance self-regulation tend to foster the rights kinds of feedback and mindset, by focusing on implications of different learning strategies rather than ideas of fixed ability (Dweck, 2012).

In addition to influencing learner motivation and well-being, tutoring practices can have a mediating effect on the impact of learner knowledge on their ability to regulate their learning. For example, tutors can informally assess the degree of domain knowledge and knowledge of task possessed by the learner, and individualise instruction based on the assessment, pitching the learning within the Zone of Proximal Development and enabling learners to self-regulate more effectively (Allal & Pelgrims Ducrey, 2000; Luckin & Luckin, n.d.). In addition, tutors can prompt learners to monitor what they know, and regulate themselves accordingly. The current body of relevant research focuses on the use of informal assessment in the classroom by teachers to determine students' present level of understanding (Andrade & Brookhart, 2020) and to use this information to guide decisions about next steps, including individualizing instruction (Heritage & Heritage, 2013). The majority of research on teacher practice suggests that many teachers are not yet using informal assessments to co-regulate learners by individualising instruction (Schneider & Andrade, 2013; Stahnke et al., 2016). However, the existing studies set in traditional classroom environments provide sufficient basis for proposing that tutors may co-regulate learners through informal assessments in a VCE. This poses an interesting avenue for research as the one-to-one nature of tutoring may lend itself to the individualisation of instruction, further than a traditional classroom environment. In addition to mediating the effect of domain knowledge and task knowledge on SRL, tutors can co-regulate learners through modelling or prompting the use of classroom strategies. This was evidenced by studies set in the context of a traditional classroom environment (Dignath & Büttner, 2018; Ewijk et al., 2013).

Tutor practices influencing Operations

With regard to the tutoring practices that can impact learner Operations, I started by examining the literature on the different styles of learner engagement, prior to considering how tutor behaviours could promote or inhibit these. Chi and Wylie (2014) consider learner

behaviour at a small grain-level while a learner is engaged within a task, rather than larger grains such as attendance or whether a learner does the required work prior to a task etc. They developed a framework which enables simple measurement of levels of engagement, as well as guidelines on promoting levels of engagement e.g., through task design or modification. The framework categorises engagement as Interactive > Constructive > Active > Passive. Each level is associated with different knowledge change (or learning processes), and different depths of understanding.

- **Passive engagement** ('receiving') is an orientation towards learning material and passive receipt of information from an instructor without doing anything else e.g. listening without doing anything else but oriented toward instruction, or reading an entire text passage silently/aloud without doing anything else
- **Active engagement** ('manipulating') when physical action or manipulation is involved e.g., looking for objects in an environment from a description of it, and so forth. Examples of active engagement include:
    - Repeating or rehearsing; copying solution steps; taking verbatim notes
    - Underlining or highlighting; Summarizing by copy-and delete
- **Constructive engagement** ('generating') is when the learner generates output beyond what is provided in the learning materials. E.g. draws a diagram, self explains or takes notes in own words (thus drawing inferences based on material available, and generating a new output). When self-explanation or note taking is exactly the same as instructed via learning materials, then mode of engagement is active rather than constructive. Instructional scaffolding can promote constructive engagement e.g., through content free prompts (e.g., 'what else'), or prompts with some of the content missing (e.g., "OK, and it goes from the atrium to the [pause]?"). Tutoring can be particularly effective at this

- **Interactive engagement** (‘dialoguing’) refers to dialog between two entities when both are participating constructively and there is a sufficient amount of turn taking. It is noted that it is important to take turns as enables students to make on-going revisions to their mental models and incorporate small grained pieces of knowledge.

The Chi & Wylie framework substantiates the hypothesis that tutors can influence learner operations by encouraging different forms of engagement. For example, tutors who tend to lecture during a session without giving the learner adequate opportunities to participate may be encouraging a passive style of engagement. On the other hand, tutors who regularly engage in open ended dialogue with learners, are more likely to encourage constructive and interactive engagement from learners.

In addition to this framework, there are a range of studies exploring the influence of different types of questions on learner engagement. For example, Hrastinski et al. (2021) conducted a study to investigate the different types of questions used by tutors in an online individual maths tutoring environment, and the impact of the various question types on conversation intensity, approach to tutoring, perceived satisfaction and perceived learning. They found that question types can be indirect (i.e., when there is partial interrogative embedded in the statement e.g., ‘the formula includes this and x’), or direct. Direct questions can be open or closed. Some questions started with ‘what’ implying a more closed teacher centred approach, while others that started with ‘how’ or ‘why’ were classified as more open-ended questions. It was also found that approach to tutoring (based on self-reporting questionnaire) was more linked to lecturing or acquisition style, when question styles used a ‘right’ at the end e.g., ‘so the way to solve the problem is xxx. Right?’ Similarly – ‘Do you recognise..’ and ‘What is..’ were teacher centred closed questions linked to lecturing style . In a similar study, Araya and Aljovin (2018) analysed 23,653 written responses submitted by 984 elementary school students to open-ended questions in science and mathematics. It was

found that certain key words in the questions, such as “explain”, had a significant effect on the length of responses and the type of words they used. These studies demonstrate that question types can be signifiers of tutors prompting different styles of engagement from learners.

In addition to the research exploring the influence of question types on learner engagement, there is a vast body of research exploring how teachers can build inter-subjectivity with their learners. Inter-subjectivity has been defined as the process of teacher and learner engagement through which they negotiate a shared understanding (Steffe & Thompson, 2000). Bober and Dennen (2001) found that online instruction can be an effective way of building inter-subjectivity, as it can facilitate improved interaction between tutors and learners. Past research shows that inter-subjectivity can be built through both verbal and non-verbal interactions. For example, the gaze of a teacher has been noted as impacting the state of the learner and the nature of the working environment. Similarly, the pace and tone of the conversation, the amount of turn taking, or the amount of praise and reward given by the teacher, can all impact the inter-subjectivity between teacher and learner. Di Gesú (2021) conducted a brief study on the ability of teachers to build inter-subjectivity during the remote schooling required by the Covid-20 pandemic. They found that the inter-subjectivity was somewhat impacted by the unfamiliar requirements of participating in a Zoom classroom, such as needing to mute and unmute the mic. However, teachers were able to mitigate this by continuing to allow learners space to participate in discussions, and by engaging in discussions around well-being during the time of Covid. Darabi and Jin (2013) found that clear instructions for learners combined with interventions such as explaining, modelling, questioning, and prompts supported inter-subjectivity. Lim et al. (2017) examined the development of inter-subjectivity in a student discussion forum, and found that instructors could help to build this through promoting dialogue at the ‘Understanding’ level of Blooms

taxonomy. This was in contrast to past research which had found that discussion at a higher level was necessary to build inter-subjectivity. Lim et al. (2017) found that a stronger focus on the dialogue may encourage learners to push a discussion to a higher level on Bloom's as they developed their ability to synthesize different aspects of a topic. The research indicates that patterns of tutor-learner engagement are relevant to whether inter-subjectivity is developed. There is a need for further research which examines the specific types and positioning of tutoring practices that influence learner engagement, and can support inter-subjectivity.

Finally, in addition to studies set in the context of traditional classroom environments, research conducted in the context of intelligent tutoring systems (ITS) supports the claim that tutor practices can co-regulate learner operations. For example, Graesser and McNamara (2010) explored the indirect influence of pedagogical agents in ITS on learner SRL operations. Through discourse moves intelligent tutoring systems, such as Auto-tutor, encourage deeper forms of comprehension through guiding the discourse using pumps, hints and prompts, feedback, and eventually using assertions if required. Graesser and McNamara (2010) argue that pedagogical agents are more suited to tracking learner knowledge and adapting discourse moves, than human tutors, who may find it challenging to enact practices which are contrary to their own conversational styles. However, Graesser and McNamara (2010) also note that limited research has been done in this field using expert human tutors (Lepper et al., 1993; Lepper & Woolverton, 2002), and this seems to be an area ripe for further research.

Tutor practices influencing Products

Tutors can directly influence SRL through prompting the generation of products of SRL, or through explicitly modelling the process for generating products of SRL. Multiple

studies have been conducted on explicit prompting of planning, enactment of strategies and monitoring in the field of ITS', HLE's, and classroom environments.

In the context of ITS, pedagogical agents can directly train students on metacognitive and SRL skills, tracking and measuring progress. Examples of pedagogical agents with a direct approach towards training SRL skills are provided in the table below:

Table 25. ITS pedagogical agents which promote generation of SRL products.

Name of ITS	**SRL product**	**Description**
MetaTutor	Planning Enactment of strategies Monitoring	MetaTutor trains students on 13 strategies that are theoretically important for SRL. There is a main agent Gavin that controls the overall learning environment, and three satellite agents which control the phases of SRL, planning, monitoring and applying learning strategies. Each phase can be further broken down under the guidance of the agent. For example, monitoring can be broken down into judgements of learning, feeling of knowing, content evaluation, monitoring the effectiveness of a learning strategy, and monitoring progress towards a goal. Examples of learning strategies are re-reading, taking notes, searching in a goal-oriented manner, drawing diagrams or tables, making inferences, coordinating information from various sources. Each of these skills is measured using data on the students' actions, decisions, ratings and verbal input
iStart (Interactive Strategy Trainer for Active Reading and Thinking)	Enactment of learning strategies	iStart aims to help students become better readers by prompting them self-explain the text (McNamara et al., 2007). It focuses on 5 reading strategies; monitoring comprehension, paraphrasing

		explicit text, making bridging inferences between one sentence and another, making predictions, and elaborating text. Student skills are measured by their ability to correctly identify the reading strategy being used during the process of modelling the strategies by the agents. In the 'Practice' module, students type self-explanations into the system, for which feedback is provided by the agent, and requests made for the student to modify inadequate answers. Measures of the skill are the quality of the self-explanations, relative frequency of positive feedback from the agents, and relative frequency of requests to modify self-explanations
SEEK (Source, Evidence, Explanation and Knowledge)	Monitoring	SEEK (Source, Evidence, Explanation and Knowledge) web tutor aims to improve critical thinking skills for college students, using a 'hint' button which has suggestions for how to guide google search, a 'pop up ratings and justifications', which prompts students to assess the reliability of the information, and finally a pop-up journal, which asks 5 questions about the reliability of the site the learner visited

In addition to studies conducted in the context of ITS, researchers have explored how explicit prompts can scaffold learning in an HLE. For example, Winne and Hadwin (2013) explored the benefit of a software tool nStudy which generates fine grained data to support learners as they engage in SRL. The features of nStudy include:

- Notes, Terms, Glossaries that learners are prompted to label by type, thus encouraging deeper information processing

- Tags, through which text that is being reviewed can be labelled with a type of tag
- Chat tools, with embedded prompts to scaffold learners as they chat. Learners can select a type of role from a drop-down list next to the chat (e.g., planner), and based on that, prompts will appear to help to scaffold the discussion from a co-regulatory perspective. E.g., there may be prompts on how shall we approach this task? What should the next step be?, 'I think we may be a little off track'. Etc

Similarly, Hadwin and Winne, 2001 developed an adaptive electronic textbook, called CoNoteS. This platform includes implicit and explicit scaffolds for SRL:

- Implicit scaffolds include various cues within the platform to set goals (e.g., through a list of suggested objectives), and enact learning strategies (e.g., through prompts to create and label different types of notes)
- Explicit scaffolds include pre-labelled notes and indexes. However, the authors note that these explicit scaffolds may have an unintended detrimental effect, if students feel obliged to complete them as they are pre-labelled, thus reducing their sense of control over the task and possibly impeding quality of work and resulting in cognitive overload.

HLE platforms such as nStudy and CoNoteS use features of instructional design, embedded as cues within an electronic platform, to drive SRL. Similar types of cue may be used by human tutors to scaffold learner SRL both directly as well as indirectly. It is possible that human tutors are advantaged in explicitly scaffolding SRL products, as they are able to judge learner affect and cognitive overload, prior to cueing various types of SRL behaviour.

In addition to the studies conducted in the context of ITS and HLE, some studies have been conducted on the explicit scaffolding of SRL in traditional classroom environments. For

example, Bell and Pape (2014) developed a theoretical framework coupled with a strategy observation tool to explore how teachers scaffolded SRL. The research was conducted over the course of the academic year with 7th grade mathematics students. In this paper, standard maths problem aligned to the curriculum were used. The learning design was aligned to the theoretical cycle of SRL, with students being given time to discuss, make sense of the problem and plan their work (forethought), enact their plans, discuss the effectiveness of the strategies used as a group (self-reflection), and make changes to their own strategies if desired (observation, imitation, performance control). Reflective classroom discourse was used to scaffold development of SRL through focus on the strategies used rather than the correct numerical answer. Explicit discussion of the strategies allowed students to understand that they could analyse each other's learning behaviours. Within the classroom discourse, it was clear that students struggled to precisely explain study strategies, instead describing abstract study behaviours such as study hard or do lots of problems. Through guided discourse, they could be made to make the actual strategies they used (e.g., building on prior knowledge) explicit. This is an example of social interaction or co-regulation being important for the development of SRL skills.

Similarly, Dignath and Büttner (2018) argue that teachers need to show a combination of direct and implicit instruction to have a lasting and significant effect on learner SRL skills. Coded observation data was used to explore how teachers instructed learners on SRL in practice, while the interview data was used to understand their motivations and beliefs underlying their practices. The results showed that direct instruction of SRL was very rare, and interviews of secondary school teachers showed that teachers over-estimated the extent to which they instructed learners on SRL strategies. Teachers also found it difficult to explain to the researchers what was meant by cognitive, motivational and metacognitive strategies,

indicating that teacher training is required to enable them to both directly and implicitly instruct learners on SRL.

Tutor practices influencing Evaluations

Tutors can influence the cognitive evaluations that learners make through the nature of the feedback that they provide, and through prompting monitoring.

Andrade and Brookhart (2020) point to the substantive reviews conducted on the role of feedback and note that feedback has been identified as one of the most important factors influencing the success of self-regulatory programmes in primary schools. They summarise the body of research to state that feedback is most effective when it is detailed and narrative rather than evaluative, supportive, delivered at the right time, and to learners who are receptive and have a reasonably strong sense of self-efficacy. In contrast to formative feedback, summative assessment can be detrimental to learner motivation and performance (Lipnevich & Smith, 2009).

In the context of a VCE, tutors can communicate dimensions of success to learners through metacognitive prompts or modelling. It is likely that the learner perception of success criteria and the subsequent standards developed, will be influenced by a combination of the tutoring practices and relevant resources provided by the VCE (for example, a list of session objectives).

Tutor practices influencing Standards

The proposition that tutors can influence the standards against which learners self-regulate is supported by range of studies set in a traditional classroom environment. For example, researchers have explored the influence of success criteria on self-regulation. Success criteria describe the qualities of excellent work, and can be communicated verbally by teachers, through assessment rubrics (Brookhart & Chen, 2015), or via worked examples

which imply success criteria (Hattie, 2008b). They shape the learning goals that students set for themselves in specific terms.

Students' perceptions of the teachers' criteria for success will vary by student, resulting in differing standards across the classroom. Butler and Schnellert (2015) explored how students' perceptions of their teachers' criteria will influence the manner in which they regulate their learning. For example, students who perceive a reading assignment as being a memorization task will set standards related to memorisation for themselves, and use memorization and recall strategies to meet these standards. On the other hand, students who perceive the same assignment as requiring comprehension will set themselves standards of being able to explain and apply the content of the text.

Teacher expectations of students will influence their communication of success criteria, and in turn, the standards that students set for themselves. Rosenthal and Jacobson's (1968) study of Pygmalion in the Classroom found that students randomly labelled 'bloomers' did increase their attainment over the course of the academic year. While this study has been not always able to be replicated (Hattie, 2008b), it pointed out the now acknowledged influence of teacher expectations on the standards that students set for themselves. Further work established variation in teachers' expectations, based on race, gender, socioeconomic status and prior attainment (e.g., van den Bergh et al., 2010).

4.1.3 Contextualised model of SRL

The design of the Winne and Hadwin model of SRL, contextualised to explore how tutors co-regulate learners in a VCE, is set out below:

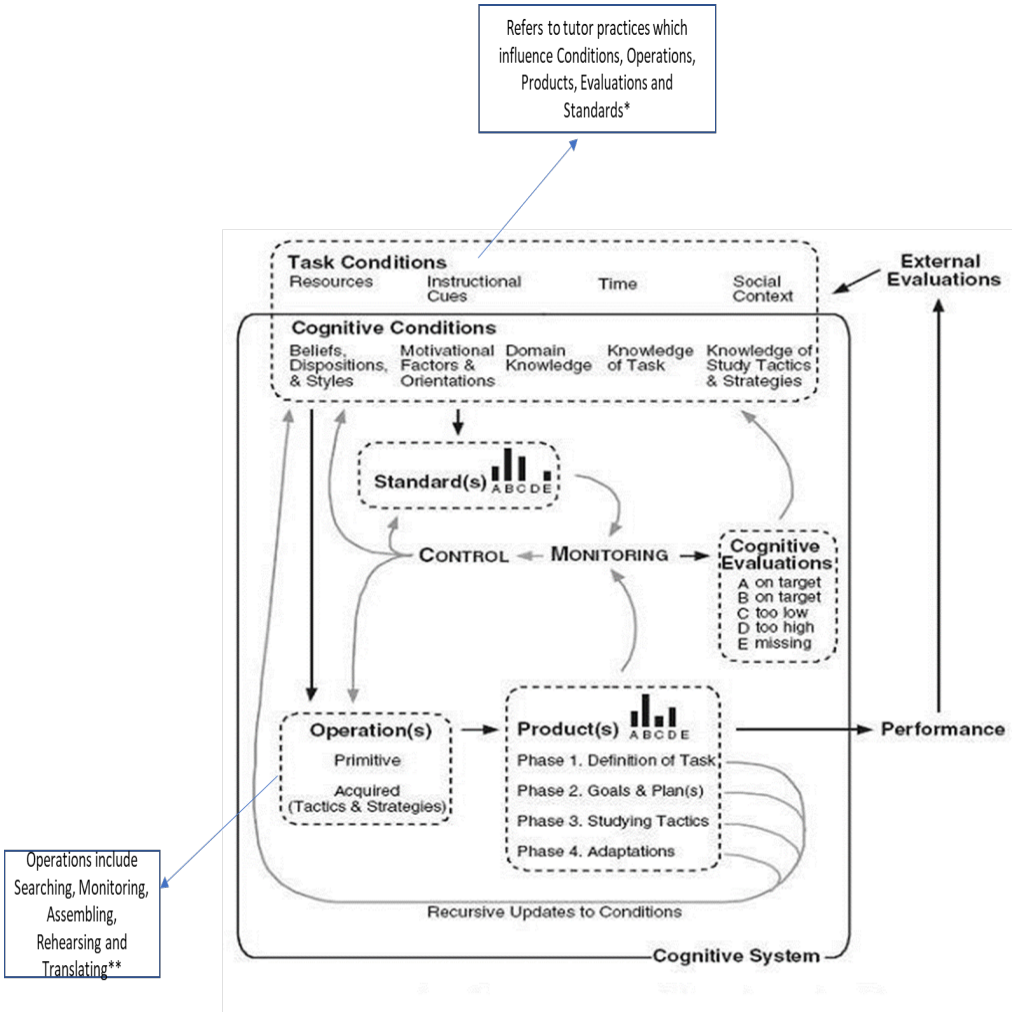


Figure 16. Contextualised Winne and Hadwin model of SRL, for exploring how tutors co-regulate learner operations in a VCE. Adapted from "Studying as Self-Regulated Learning," by P. H. Winne and A. F. Hadwin, 1998, in D. J. Hacker, J. Dunlosky, and A. C. Graesser (Eds.), *Metacognition in Educational Theory and Practice* Routledge (https://doi.org/10.4324/9781410602350-19).

*The full list of tutor practices which can theoretically co-regulate learners in a VCE is set out in Table 24

** The definition of the SMART operations is set out in table 23.

As described in the sections above, the contextualisation of the Winne and Hadwin model is intended to enable the influence of tutoring practices on learner operations to be tracked in a holistic manner, which accounts for the indirect influence of other related factors in the model, such as the learners cognitive state, and evaluations of their own learning.

Winne (2020) refers to -IF-THEN' analysis, which considers the probability of a learner event/state following an action by the tutor. For example, if the tutor praises the learner for their effort, what are the common behaviours exhibited by learners following this practice?

Prior to being able to use the contextualised model to track the influence of tutoring practices on learner behaviours, it is necessary to build a framework of signifiers, using a literature review and empirical data. This work is set out in Section 4.2.

Summary

This Section has set out a key result: a conceptualisation of SRL which can be used to explore how tutors co-regulate learners in a VCE.

As described above, there is a need for theoretically based frameworks which enable the exploration of co-regulation in a VCE. The majority of the current literature either hasn't applied a theoretical framework, or has developed frameworks tailored for a traditional classroom environment. These frameworks tend to use large grained forms of data. There is a need to develop a contextualised model of SRL, which can provide the foundation for theoretical frameworks and enable an exploration of co-regulation, utilising the fine-grained data generated from online contexts.

I have proposed a contextualised model of SRL, which is based on the Winne and Hadwin (1998) model of SRL. The architecture of the Winne and Hadwin model makes it more suited to contextualisation versus other established models of SRL. For example, it integrates external conditions (such as tutor practices) into the model as a component of SRL, enabling a more rigorous exploration of how tutors can influence learner SRL. Further, the Winne and Hadwin model is distinct in that it considers the impact of monitoring and evaluation at a component level, rather than examining impact by phase of SRL (Azevedo et al., 2011). This is a more granular approach, which enables analysis of how changes on one component of SRL (e.g., task conditions) can lead to changes in other components (e.g., type

of operations used by learner). The Winne and Hadwin model is granular, and allows the tracking of learner information processes within a phase, as well as between phases. Finally, the Winne and Hadwin model provides a holistic representation of SRL, as it as it synthesises all the various components of SRL from the literature into a heuristic framework. This enables the model to provide a strong foundation for exploring co-regulation in an online context, given the limited research conducted to date.

In order to contextualise the Winne and Hadwin (1998) model of SRL, I have hypothesised about the potential tutoring practices which can influence learner SRL in a VCE, aligned to each of the components of the model. The proposed contextualised model has been validated through a select literature review evidencing the relationship between the tutoring practices and model components. All components of the model have been considered in this contextualisation, rather than simply focusing on the components of SRL which have a direct relationship with learner SRL processes. This is important, as components of SRL can indirectly influence each other, and a holistic representation of co-regulation must therefore consider tutoring practices aligned to the entire model. Within the context of the model, learner SRL processes are aligned to the learner 'Operations' component of the model, as defined by Winne (2010).

A pictorial representation of the contextualised model is set out in 4.1.3, and the detailed co-regulatory practices are set out in tables 23 and 24. In chapter 4.2, I will explore how the contextualised model of SRL can be used as the foundation for a framework of signifiers of co-regulation in a VCE.

4.2 Theoretically based Framework of Signifiers

Introduction

In this Section, I develop a theoretically grounded framework of signifiers, which can be used to represent how tutors co-regulate learners in a VCE. The framework of signifiers is based on the contextualised model of SRL, proposed in Section 4.1.

The Section follows the structure below:

- I recap the approach adopted to build a framework of signifiers
- Following the approach set out, signifiers are identified for tutor practices in the context of a VCE
- Following the approach set out, signifiers are identified for learner SRL operations in the context of a VCE
- An observation study is conducted to validate the signifiers of learner operations

Approach adopted to build framework of signifiers

As set out in Section 3.1, I follow the mixed methods approach towards building the framework of signifiers (Cukurova et al., 2016). The mixed methods approach consists of three main stages, which are described in detail in the sections below:

- Identification of a theory driven framework ('identify')
- Adaptation of the theoretical framework to fit the research purposes ('adapt')
- Merging the fine-grained actions from the data to the adapted theoretical framework ('integrate').

4.2.1 Restructure the contextualised model of SRL

I will use the contextualised model presented in Section 4.1 to build a framework of signifiers. In this framework, I will group similar tutoring practices together rather than continuing to align the structure of the framework to the COPES model (Winne & Hadwin, 1998). This will allow the framework to be represented in a simpler manner, and enable more

efficient coding of data. Co-regulatory tutoring practices have been mapped to the category of practice, rather than the component of the COPES model.

Models of SRL generally recognise that SRL is driven by factors relating to motivation/affect, metacognition and cognition (Pintrich, 2000; Zimmerman, 2000). Three areas of focus were therefore identified, which are used as the basis for categorising tutoring practices:

- Tutoring practices which influence learner **motivation and affect**
- Tutoring practices which influence learner **metacognition**
- Tutoring practices which influence learner **cognition**

Table 26 maps the contextualised model of SRL developed in chapter 4a of this thesis to each area of focus. Table 27 presents the framework of signifiers organised by the type of tutoring practice (rather than aligned to the structure of the COPES model). It should be noted that tutoring practices are allocated to an area of focus based on a judgement of what the primary focus of the tutoring practice is. It is possible for a tutoring practice to influence multiple areas of focus; for example, praise can influence learner motivation, and have a subsequent positive influence on learner engagement levels. In addition, as described above, tutoring practices are derived from mapping the COPES model of SRL to the literature, rather than scanning the literature for tutoring practices which influence learner motivation/affect, metacognition and cognition. It is therefore possible that there are additional tutoring practices which would influence learner motivation, metacognition and cognition, which are not identified below. This is particularly true for the motivation category, which is not a detailed component in the COPES model.

Table 26: COPES model mapped to co-regulatory tutoring practices by type.

Components of the Winne and Hadwin framework		**Potential tutor practices that may influence component**	**Type of tutoring practices**
Task Conditions	Resources	N/A (resources centrally set in the VCE)	

(resources, instructional cues, time, social context)	Instructional cues	N/A	
	Time	N/A	
	Social context	N/A	
Cognitive Conditions	Beliefs, dispositions, styles	Tutor practices which influenced learning beliefs and mindset	Motivational
	Motivational factors and orientations	Tutor practices which influenced learner motivation and goal orientation	Motivational
	Domain knowledge	Tutor practices which can be regarded as forms of informal assessment	Cognitive
	Knowledge of task	Tutor practices which monitored the learner knowledge of the task	Metacognitive
	Knowledge of study tactics and strategies	Tutor modelled and prompted learning strategies for the learner	Metacognitive
Operations	Primitive	N/A	
	Acquired (tactics and strategies)	Tutor practices which influenced shallow versus deep cognitive engagement by the learner Tutor practices which built inter-subjectivity	Cognitive
Products	Phase 1- definition of task	Tutor prompted or modelled task definition	Metacognitive
	Phase 2 – goals and plans	Tutor prompted or modelled goal-setting and planning	Metacognitive
	Phase 3 – studying tactics	Tutor prompted or modelled study tactics	Metacognitive
	Phase 4 – adaptations	Tutor prompted or modelled reflection to make fundamental changes to components of SRL	Metacognitive
Evaluation	Cognitive evaluations	Tutor prompted the learner to monitor their cognition Tutor provided narrative feedback to the learner	Cognitive Metacognitive
Standards	Standards	Tutor prompted the learner to shape the dimensions against which learners monitor themselves	Metacognitive

Table 26. Co-regulatory tutoring practices categorised by type of practice.

Category of tutoring practice	**Description of category**	**Tutoring practices**
Motivational	Promote sustained learner motivation	Tutor practices which influenced learning beliefs and mindset
		Tutor practices which influenced learner motivation and goal orientation
Metacognitive	Promote learner metacognition through explicit and implicit means	Tutor prompted or modelled task definition
		Tutor prompted or modelled goal-setting and planning
		Tutor prompted or modelled study tactics
		Tutor prompted or modelled reflection to make fundamental changes to components of SRL
		Tutor prompted the learner to monitor their cognition
		Tutor prompted the learner to shape the dimensions against which learners monitor themselves
Cognitive	Promote different forms of learner engagement to build cognitive understanding	Tutor informally assessed learner to elicit current level of information and individualise instruction
		Tutor practices which influenced shallow versus deep cognitive engagement by the learner Tutor practices which built inter-subjectivity
		Tutor provided narrative feedback to the learner

4.2.2 Review the data generated from the VCE

There are two main sources of data from the VCE:

- Audio data, which includes the natural, unstructured dialogue between tutor and learner
- Video data, which includes tutor and learner actions on the whiteboard (e.g., writing/erasing), as well as the use of VCE tools (e.g., effort points/stickers awarded by the tutor)

As discussed in Section 2.3, previous studies have demonstrated a range of different approaches towards collecting and coding data, depending on the specific study context, and

research aims. For example, researchers may want to understand more about the learning tactics internalised by learners, in which case they may use think aloud learning protocols to generate data on these. Alternatively, fine grained online trace data generated by a learner can be mined to map out SRL as an observed series of events.

As discussed in Section 2.3, the use of natural unstructured dialogue between learner and tutor to signify co-regulation is under-researched. While the interactions between learners and pedagogical agents have been subject to some research, the use of human-to-human dialogue would enable analysis that can later support these interactions. Further, natural language dialogue can be analysed from the perspective of understanding tutoring practices that support and inhibit learner SRL, rather than primarily focusing on the development of learner SRL skills (as is the case with most studies on ITS).

I have explored how natural unstructured dialogue between tutor and learner can be used as the primary data source for representing how tutors co-regulate learners in a VCE. While the use of this type of data to explore co-regulation is under-researched, researchers have considered in the past how verbal data can be collected and analysed, and the output from these studies has been used to inform my approach towards using natural unstructured dialogue to signify co-regulation. Chi (1997) defined verbal analysis as the coding of verbal data which is then quantified (e.g. tabulated, counted, relations investigated). The verbal analysis method proposed by Chi (1997) highlighted a range of considerations in the use of verbal data. I have outlined these below, along with implications for my research.

Chi (1997) highlighted that segmentation of data can be at different levels of granularity. For example, it can be at a proposition level, sentence level, exchange of dialogue level. Researchers need to consider whether the type of results obtained from the level of segmentation will be able to meaningfully answer the research question. With regard to my research, as this research aims to understand how tutors can influence learner SRL,

segmenting the data by conversational turn would enable meaningful analysis. I will therefore use conversational turns as signifiers for tutoring practices and learner operations.

In addition, Chi (1997) pointed out that finding criteria for codes can be tricky and sometimes the key contribution of the research. For example, researchers found the use of connecting words by expert children was evidence of a richer connected knowledge base, versus the fragmented knowledge base of novice children who did not use connecting words, and simply stated what they saw. There can be ambiguity around the utterances and what they should be coded as. Chi stated that it is important to be clear around how much context should be taken into account when coding a segment, and whether the preceding and following protocols should be analysed too. The amount of context that is taken into account should remain consistent between codes. As my research aims to explore the influence of tutoring practices on learner SRL processes, context is important and the signifiers will be implemented in a manner which takes account of the context, and considers the preceding and following dialogue exchanges. As coding will be implemented based on the meaning of an utterance, I will provide examples of the micro features of utterances to aid the coding process but the framework of signifiers will not be led by these micro features.

4.2.3 Map the literature to the framework to identify potential signifiers

In this section, key findings from the literature are analysed to formulate signifiers of how tutors may co-regulate learner SRL operations in a VCE.

Table 27. Literature review mapped to framework to identify potential signifiers.

Category of tutoring practice	**Tutoring practices**	**Relevant literature**	**Key findings from literature**	**Potential signifiers of co-regulation in a VCE**
Motivational	Tutor practices which influenced learning beliefs and mindset	(Cimpian et al., 2007)	Verbal praise, focused on effort/process rather than ability can promote sustained motivation	Tutor offers verbal praise, focused on effort/process rather than ability

				Tutor gives non-verbal rewards for effort
	Tutor practices which influenced motivation and goal orientation	(Dweck, 2012)	Discussions on the implications of using different learning strategies can encourage the use of learning strategies	Tutor prompting learner to monitor strategy use Tutor explicitly modelling the use of different strategies
Metacognitive	Tutor prompted or modelled task definition	(Bell & Pape, 2014) (Graesser & McNamara, 2010; Hadwin & Winne, 2001; Winne & Hadwin, 2013)	Teachers and pedagogical agents can explicitly cue learners to plan tasks and set goals, or model task planning and goal setting	Tutor explicitly cues learner to plan the task and set goals Tutor models how to plan the task and set goals
	Tutor prompted or modelled goal-setting and planning			
	Tutor prompted or modelled study tactics	(Trevors et al., 2014)	Pedagogical agents can explicitly cue learners to enact learning strategies, or model how to enact learning strategies	Tutor explicitly cues learner to enact learning strategies Tutor models how to enact learning strategies
	Tutor prompted or modelled reflection to make fundamental changes to components of SRL	(Dignath & Büttner, 2018)	Teachers can encourage learners to practice guided reflection	Tutor explicitly cues learner to conduct guided reflection
	Tutor prompted the learner to monitor their cognition	(Andrade & Brookhart, 2020; Trevors et al., 2014)	Teachers and pedagogical agents can prompt learners to monitor cognition and affect	Tutor prompts learner to monitor cognition and affect

			MetaTutor uses pedagogical agents to prompt learners to evaluate their learning. These prompts include: Monitoring Progress Towards Goal (prompts relating to quality of work and time spent) Judgements of Learning (prompts on how well the learner has understood the work) Feeling of Knowing – (prompts on how well the learner already knows the work) Content Evaluation (learner evaluates relevance and quality of content supplied)	Tutor prompts learner to evaluate the quality of their work Tutor prompts learner to monitor the on time spent the on question Tutor prompts learner to evaluate their level understanding Tutor prompts learner to evaluate prior knowledge Tutor prompts learner to evaluate relevance and quality of content, and learning aids
	Tutor prompted the learner to shape the dimensions against which learners monitor themselves	(Brookhart & Chen, 2015; Hattie, 2008)	Teachers can communicate success criteria, verbally or through rubrics, and influence the multi-dimensional standards adopted by learners	Tutor verbally communicates success criteria to learner
Cognitive	Tutor informally assessed learner to elicit current level of information and individualise instruction	(Heritage & Heritage, 2013)	Diagnostic questions can enable teachers to individualise instruction and encourage learner engagement	Tutor asking specific, closed ended questions to diagnose student understanding

				and knowledge
	Tutor practices which influenced shallow versus deep cognitive engagement by the learner Tutor practices which built inter-subjectivity	(Chi & Wylie, 2014; Hrastinski et al., 2021) (Araya & Aljovin, 2018)	ICAP framework proposed, which analysed various forms of learner engagement: Interactive engagement: Turn taking, and constructive dialogue Constructive engagement: Generation of new output Active engagement: Physical action taken by learner (e.g., repetition or manipulation Passive engagement: Learner passively receives instruction Open ended questions are perceived to be signifiers of constructive engagement being cued Questions starting with 'how', why', 'explain' are perceived to be signifiers of an open-ended teaching approach. Questions starting with 'what' are perceived to be signifiers of a teacher centered approach	Tutors can cue interactive and constructive engagement, through open-ended questions (e.g., starting with 'how', 'why' or 'explain') Tutors can cue active engagement through specific closed ended questions (e.g., starting with 'what', or 'where') Tutors can cue passive engagement through lecturing to students Tutor utterances aimed at building rapport with learner (e.g., how was your weekend?)

			Teachers can cultivate a sense of learner belonging to improve engagement	
	Tutor provided narrative feedback to the learner	(Andrade & Brookhart, 2020)	Narrative feedback is related to positive learning outcomes	Tutors providing narrative versus summative feedback to learners

The examples of literature provided in the table above are drawn from a range of contexts, including traditional classroom environments, HLEs, and ITSs. Limited work has been done in the study context of one-to-one tutoring in a VCE. The hypothesis is that the cues and prompts which have been used to co-regulate in other study contexts, will also apply in the context of a VCE. The analysis that I have conducted will test this hypothesis, as well as consider the potential impact of the VCE context on the patterns of behaviour observed.

The studies analysed explore the influence of both direct and indirect practices on components of learner SRL. Indirect practices include the design of a learner centred environment, in which a learner is encouraged to take ownership of their learning processes. For example, they can include the nature of the resources used by the teacher (Dignath & Büttner, 2018), or the design of the learning environment provided by an ITS (Graesser & McNamara, 2010). Direct practices refer to the prompting of SRL processes, either explicitly or implicitly. For example, instructors/pedagogical agents can explicitly discuss what a learning strategy is, and appropriate strategies to use in specific contexts (Dweck, 2012), or they can implicitly encourage the use of learning strategies through modelling the use of the strategy (Trevors et al., 2014), or by asking questions which guide learners towards the use of that learning strategy. For the purposes of my study context, findings on direct practices

which influence learner SRL are of greater relevance towards understanding how tutors can co-regulate learners in a VCE, as tutors have little control over the resources/design of the VCE study context.

In addition to directly influencing learner SRL through explanations/prompts/modelling, tutors can cue different forms of engagement from the learner, and potentially encourage them to take greater control over their learning. For example, Chi and Wylie (2014) developed the ICAP framework, which categorises learner engagement depending on the degree of learner activity and independence demonstrated. There is a spectrum, ranging from:

- Passive engagement (where the learner is entirely reliant on the teacher/resources, and shows minimal activity)
- active engagement (where the learner continues to be led by the teacher/resources, but shows physical signs of involvement in learning)
- constructive engagement (where the learner is demonstrating greater control over their learning by generating new output beyond that provided to them by the teacher/resources), and finally
- interactive engagement (where the learner demonstrates sophisticated control over their learning, through dialogue with their teacher and generation of new output through this dialogue).

Tutoring practices which co-regulate learners have been identified by analysing literature on the different modes of learner engagement, and considering how tutors can prompt this engagement. For example, tutors who adopt a lecturing style, asking minimal questions, are encouraging a passive form of engagement. In contrast, tutors can encourage interactive and constructive forms of engagement by asking learners open-ended questions, encouraging them to make constructive contributions and participate in dialogue. Relevant

literature on the engagement cued by various question types has also been analysed. For example, past research has demonstrated that open-ended questions are perceived to be signifiers of constructive engagement being cued (Hrastinski et al., 2021). Questions starting with 'how', why', 'explain' are perceived to be signifiers of an open- ended teaching approach. Questions starting with 'what' are perceived to be signifiers of a teacher centered approach (Araya & Aljovin, 2018).

4.2.4 Integrate empirical data into the framework of signifiers

Following the literature review, I reviewed empirical data to refine the framework of signifiers. 50 tutoring sessions were selected at random and the recordings of the VCE sessions were observed. During my observations, I considered whether the signifiers of the tutoring practices (captured in table 28) were appropriate and if additional signifiers were required. I also considered whether there were additional tutoring practices which were not currently captured by the framework, but could be considered to co-regulate learners.

The findings from the review of the empirical data are as follows:

- Motivational tutoring practices:
    - The review of the empirical data from the VCE demonstrated that tutors can use the platform to offer stickers, pictures and effort points to learners. These are signifiers of the tutor aiming to motivate learners, through incentives and rewards
    - The review of the tutor-learner dialogue indicated that tutors often ask students about their interests and hobbies. These are signifiers of tutors aiming to motivate learners through building a rapport with them
- Metacognitive tutoring practices:

- Tutoring practices aimed to building metacognition mainly focused on prompts for the learner to monitor/reflect, with provision of examples of metacognitive processes being extremely rare
- Tutors can prompt learner monitoring and reflection through verbal cues (e.g. are you sure?), as well as non-verbal cues (e.g. tone and pausing).

- Cognitive tutoring practices:
  - Based on the literature review, tutors were defined in the earlier iteration of the framework as cueing different forms of engagement based on their tutoring style.
    - Open ended engagement: Tutors can cue interactive and constructive engagement, through open-ended questions (e.g., starting with 'how', 'why' or 'explain')
    - Guided engagement: Tutors can cue active engagement through specific closed ended questions (e.g., starting with 'what', or 'where')
    - Tutor-led engagement: Tutors can cue passive engagement through lecturing to students
  - The review of empirical data refined my understanding of how tutors can influence different forms of learner engagement. In particular, the empirical review indicated that 'tutor-led' engagement was a broad category which encompassed situations in which the tutor was directing learners to take a specific action to solve a problem, as well as situations in which the tutor was explaining a concept to the learner to scaffold understanding. In order to more accurately represent the various forms of tutor-learner engagement, the 'tutor-led' engagement category was replaced by 'directive engagement', and 'explanatory engagement'. In addition, a new category of engagement was

identified termed ‘passive engagement’. This refers to tutors observing the learner or repeating what the learner says, without offering any further guidance. I distinguished between guided/open-ended engagement, and passive engagement, as the former were proactive teaching practices aimed at scaffolding engagement, while passive engagement was fully student led.

The iterated list of tutoring practices and signifiers, integrating the findings from the empirical review, is presented below in Table 29.

Table 28. Iterated framework of signifiers (tutor practices).

Tutor practices	**Audio signifiers**	**Video signifiers (whiteboard/log-file events)**
Boosting learner affect and motivation	Praise (e.g., well done)	Drawing emoji Effort points, pictures, stickers
Metacognitive prompts	Prompts to monitor understanding, plan approach, or reflect (e.g., are you sure)	
Directive engagement	Utterances framed as instructions with use of command words. Tutor interrupting the learner Tutor giving the answer to the learner	Tutor fills in question answers
Explanatory engagement	Utterances in which the tutor explains part or all of the concept and approach	
Guided engagement	Utterances in which the tutor asks specific, closed ended questions (e.g., starting with what or which)	Tutor writes on whiteboard
Open ended engagement	Utterances phrased as open-ended questions (e.g., starting with how or why). Appropriate pausing.	
Tutor passive engagement	Tutor observing learners Pauses between learning activities Tutor repetition	

Framework of signifiers of learner SRL operations

In order to develop the framework of signifiers for learners, we followed the steps below:

1. Conducted a literature review to identify signifiers for each of the learner operations
2. Conducted an empirical review to refine the framework of signifiers
3. Conducted an observation study to further validate the framework of signifiers for learners

4.2.5 Learner SRL processes

As reviewed in detail in literature review section, operations refer to the nature of the learning process rather than the type of information being processed; the SMART operations defined by Winne (Winne, 2010; Winne et al., 2013) can act on both cognitive and metacognitive information.

In order to develop signifiers of learner operations in the study context, I have considered the definition of the operations provided by Winne and Hadwin, and adapted these for the study context. I have also considered relevant studies, which research similar learning processes to those defined by Winne and Hadwin, and analyse whether the signifiers utilised could be relevant.

The potential signifiers identified for the study context are set out below:

- Searching: Winne and Hadwin defined the searching operation as the learning process in which the learner is seeking to locate information. In the study context of this thesis, signifiers of this process would include cursor movements, or learner questions seeking to establish where information is located

- Monitoring: This operation is defined as the learner evaluating the extent of their knowledge about a topic, their degree of understanding, and the efficacy of the learning strategies deployed. Studies in ITS aiming to develop learner metacognition (e.g., Trevors et al., 2014) utilise learner signifiers from the learner actions in the online environment. In the context of learners working through textbooks, Azevedo and Hadwin (2005) considered underlining to be evidence that the learner had monitored the relevance and importance of the information underlined. In the study context of a VCE, monitoring is likely to be signified through a combination of verbal signifiers (e.g. learner saying they don't understand), and non-verbal signifiers (e.g. learner writing and then erasing).
- Assembling: This refers to the construction of new output by the learner, which goes beyond information that has been provided to them. It is similar to the Constructive level of engagement in the (Chi & Wylie, 2014) framework. A constructive level of engagement was signified by the production of a solution to a problem. Similarly, assembling in a VCE can be signified by the learner solving a maths problem. In addition, learner utterances in which the learner generates new output beyond that provided, can also be a signifier of the Assembling operation
- Translating: This refers to the transformation of material learnt into new forms. There is some overlap with the interactive level of engagement in the Chi and Wylie (2014) framework. Multiple turns of constructive dialogue signify that the learner is engaging in deeper forms of learning, and transforming material learnt.

Table 29. Framework of signifiers for learner operations (based on literature review only).

Learner Operations	**Relevant studies**	**Potential signifiers in the study context of this thesis**
Searching	(Chi & Wylie, 2014)	Manipulation of content and learning resources (e.g., searching with the cursor in an online environment)

Monitoring	(Azevedo & Hadwin, 2005; Trevors et al., 2014)	Utterances asking for help, or evidencing self-monitoring
Assembling	(Chi & Wylie, 2014)	Utterances or actions constructing information through new contributions
Rehearsing	(Chi & Wylie, 2014)	Utterances repeating information provided
Translating	(Azevedo & Hadwin, 2005; Chi & Wylie, 2014)	Constructive interactions with another person or agent

4.2.6. Empirical review for learner framework of signifiers

Following the literature review, I reviewed empirical data to refine the framework of signifiers for learner operations. 50 unique tutoring sessions were selected at random and the playback of the VCE session was observed. Based on this review of empirical data, the signifiers of learner operations were defined more precisely.

The empirical review clarified the need to distinguish between assembling and translating. Assembling refers to the construction of new output, but does not include the transformation of this into new forms. It is therefore signified by the learner formulating answers to questions, while translating is signified by the learner engaging in deeper forms of dialogue with the tutor (e.g., by asking challenging/clarifying questions).

Table 30. Framework of signifiers for learner operations.

Learner operations	**Audio signifiers in a VCE**	**Video signifiers (Whiteboard and log file events)**
Searching	Utterances intended to locate information	
Monitoring	Utterances overtly monitoring (e.g., 'I'm not sure') Asking for help	Writing and erasing
Assembling	Responding to a question, through on utterance or by writing	Computing answer using a column chart or number line

Rehearsing	Repeating tutor or question Copying out question	
Translating	Asking a challenging or clarifying question (e.g., starting with 'why' or 'how') Explanatory utterances	

4.2.7 Observations of learner operations

In order to validate the framework of signifiers of learner operations, in-person observations of learners using the VCE platform were conducted. The following approach was adopted to conduct the observations:

- TSL (the industrial supplier of the VCE) contacted customer schools in London to introduce the research, and obtain consent from schools, parents and learners for participation in the observation study
- 8 schools responded and agreed to participate in the observation study. The schools were anonymised in the data collected and labelled from A-H.
- I reviewed relevant literature to assist with the development of the observation protocol. I also conducted a visit to one of the participating schools ('school A'), to observe the VCE sessions in progress. 5 learners were observed engaged in VCE sessions, for ten minutes each.
- Following the literature and empirical review, an observation protocol was developed, to identify aspects of learner behaviour which could be used as signifiers of the learner operations.
- I conducted a visit to a participating school ('school B') accompanied by a trained rater. Students conducted their sessions on the VCE in parallel to each other, and the rater and I simultaneously observed 5 students for ten minutes each. Our observations

were compared, and an inter-rater reliability of 88% was established, which was judged to be sufficient to proceed.

- I conducted school visits to the remaining schools ('schools C-H') to conduct the observations. All observations were time-stamped. In total, 30 sessions were observed, and the full breakdown of the data captured through the observation study is provided below.
- Following the school visits, I obtained session recordings of the relevant sessions, and coded the segment of the session observed, based on the framework of signifiers of learner operations presented in Table 32.
- During the observations, students conducted their sessions on the VCE in parallel to each other. I observed individual students at regular intervals of ten minutes each, randomly selecting students for observation.
- I analysed the data, by checking if a learner operation signified by online data was also signified by the session observations, at the given time-stamp. Where the operations signified matched, the signifiers were labelled as being reconciled to each other. Where the operations signified did not match, or where one source of data signified an operation while the other did not, the signifiers were labelled as being unreconciled.

Development of observation protocol

Table 31. Observation protocol

Learner operations	Observational signifiers	Rationale
Searching	Eyes scanning the screen Hands moving the mouse for a sustained period of time	Searching refers to the location of information by the learner. In the study context, this would be signified by the learner scanning the screen and moving the cursor.
Monitoring	Learner pauses after writing or speaking and shows signs of thought	Monitoring refers to the learner evaluating their progress and understanding.

	(e.g. frowns, postural changes increase)	In the study context, observational indicators would include signs that the learner is engaged in thought, based on facial expressions and postures
Assembling	Learner is typing on the keyboard Learner is writing on pen and paper	Typing on the keyboard or writing with pen and paper are signs that the learner is generating new output and engaging in the assembling operation
Rehearsing	Eyes tracking along screen while learner speaks	Eyes tracking along the screen while learner speaks is a sign that the learner is reading aloud content from the screen, which is a form of rehearsal
Translating	Explanatory gestures combined with talking in turns	Signs that the learner is engaged in constructive dialogue with the tutor
Disengaged	Looking around the room Head in arms Stretching Talking with a fellow classmate	The learner may demonstrate signs of disengagement with the tutoring session

The observation protocol was developed by interpreting the SMART operations for the context of learners working with tutors in VCE. This included a literature review to identify potential observational signifier of the operations, and an empirical review of VCE sessions in progress.

While there is no research that I am aware of, which identifies observational signifiers for the SMART operations of learners working in a VCE, there is a significant body of research exploring how facial expressions, eye tracing, and gestures can be used to signify learner engagement (see Dewan et al., 2019) for a review of the state of the art in engagement detection). Most of this research appears to be focused on how learner engagement can be automatically detected using data on facial expressions, eye tracing and gestures. The state of learner engagement is often manually labelled, and the observational data is analysed through machine learning techniques to identify automated signifiers for the state of learner

engagement. There is limited research on observations which can be collected manually and labelled using human judgment to identify the operations that learners are engaged in (Dewan et al., 2019), potentially due to the difficulties of doing this at scale. In the case of my research, the observational study is intended to validate the framework of online signifiers of learner operations, and can therefore be conducted on a smaller scale. The findings from studies automating observations to signify learner engagement are used to inform the development of the observation protocol. These findings from the literature review are complemented through the empirical review, to help formulate the observational signifiers.

Research has shown that signifiers are more accurate when they combine multi-faceted data across modalities (Grafsgaard et al., 2013; Psaltis et al., 2018); for example, eye tracking data can be combined with data on posture and facial expressions, and experimented with physiological data (e.g. heart-rate, respiration rate), or online trace data to detect engagement . In light of these findings, multi-faceted observations are used to signify each of the SMART operations in the observation protocol; for example, searching can be signified through a combination of eyes scanning the screen, and learners using their mouse to move the cursor across the screen.

With regard to posture, research has shown that increased postural movement can be a sign of either negative affect or disengagement (Grafsgaard et al., 2013). For example, the learner may slouch back in their chair or lean forwards. For the purposes of my study context, increased postural movement is interpreted to be a sign of monitoring, when it is combined with the learner pausing the flow of their work, and facial expressions show signs of thought. On the other hand, when the increased postural movement refers to the learner focusing away from the screen (e.g., looking around the room, or head in hands), then it is likely to be a sign of disengagement

With regard to gestures, research into tutor and learner interactions has explored what different types of gestures signify. Researchers have traditionally labelled gestures as being either deictic, beat or iconic (Feyereisen, 2006). Deictic gestures (pointing), and iconic gestures (representing physical phenomena) often meet the function of completing speech, while beat gestures allow the speaker to emphasise certain points. Research has begun to explore sub-categories, through examining explanatory gestures where the learner appears to be reasoning or drawing a relationship between two variables. These sub-categories of gestures are indicators that the learner is engaging in a higher form of engagement, and when accompanied by signs of dialogue (turn taking), signify the translation operation.

In addition to the findings from the literature review, an empirical review of 5 sessions at one school was conducted to help build the observation protocol. The empirical review enabled the identification of behaviours which could indicate the occurrence of the learning operation in the study context. For example, rehearsing refers to the learner repeating content provided by an external source. In the context of a VCE, learner repetition can either refer to the learner repeating the tutor or the learning material displayed on the screen. It can be signified by the learner's eyes tracking along the screen while they speak, indicating that they are reading aloud. In addition, the observation protocol was refined to include ensure the list of signifiers was comprehensive. For example, in addition to writing online, it was noted that students occasionally used pen and paper. Although this was a rare event, it was regarded as another signifier of the students generating output (i.e., assembling). Similarly, additional signs of disengagement were noted, such as stretching, yawning, or talking with friends.

Data captured

Following the school visits to develop the observation protocol, and evaluate inter-rater reliability, 6 school visits were conducted to capture observations based on the observation protocol. The breakdown of the data captured is provided in Table 10:

Table 32. Breakdown of the data captured using the observations protocol.

Number of schools	**Number of learners**	**Number of sessions**	**Number of unique LO's**	**Length of observations**
6	30	30	26	300 minutes

Process for coding the observations

The process for coding observations was agreed following a discussion with the trained rater, and included the following elements:

- Prior to commencing observations, I confirmed with the school that the students being observed had returned forms consenting to participate in the study, and any students who had not returned the forms were excluded from the observations. 5 students were randomly selected from the remaining pool of students for observations at regular intervals of ten minutes each.
- I positioned myself a distance away from the students being observed, to minimise the impact of my presence on the session. I faced the student being observed, to ensure that I could capture observations on their facial expressions, gestures and posture. The observations were timed to occur over regular intervals.
- As discussed in 3.1, signifiers are coded using online data on the basis of conversational turns and event changes. For example, a single code of rehearsing will be applied when the learner reads aloud, and this code applies until there is a conversational turn or change in behaviour (for example, the tutor asks a question, or the learner begins to solve a problem). If the learner returns to reading aloud again, after the conversational turn or change, then a second code of reading aloud is applied. Similarly, with the observational data, a single code is applied based on an

observation of an operation occurring, and this code applies until there is a change in the behaviour observed.

Analysis of results

Table 33. Total number of operations signified, across all sessions observed.

Operation	Signified by observations and online data	Signified by observations but not online data	Signified by online data, but not observations	Percentage reconciled
Searching	35	5	0	87.5%
Monitoring	50	7	3	83.3%
Assembling	120	3	0	97.6%
Rehearsing	100	10	12	81.9%
Translating	30	0	9	76.9%

The operations signified by the observations collected, were compared to the operations signified by the online data collected for the same sessions. The results are presented in Table 12.

It is notable that assembling has a particularly high rate of reconciliation. This is because there is a close alignment between the signifiers for online data and observations. The 3 occurrences signified by observations (but not online data) refer to instances where the learner is working on pen and paper. The industrial supplier of the VCE discourages the use of paper, and requests students to do all working out in the VCE instead. However, there are some rare occurrences where the student will use pen and paper for constructing solutions.

In contrast, the lowest rate of reconciliation occurs for the translating operation (76.9%). This is partly due to the potential misalignment between the signifiers coded from online data versus observations; the signifiers for online data refer to the type of utterances, while the signifiers for observations refer to the type of gestures being used. The relatively high proportion of translating operations which are signified by online data but not by observations may also demonstrate that learners who are engaging in constructive dialogue

do not consistently use explanatory gestures while speaking. Further research is required to explore this topic, but is outside of the scope of this thesis.

Overall, there is a high percentage of reconciliation between the operations signified by observations versus online data. All operations demonstrate a reconciliation of over 75%. This provides evidence that the signifiers for learner operations using online data from the VCE, can detect the occurrence of learner operations with a high rate of accuracy.

Summary

In this Section I have proposed a theoretically grounded framework of signifiers to represent how tutors co-regulate learners in a VCE. I have based the framework on the contextualised model of SRL (proposed in Section 4.1). The model has been re-organised to enable the efficient coding of signifiers of tutor co-regulatory practices.

In order to identify signifiers for tutor co-regulatory practices and learner operations, I have drawn on the literature to identify signifiers for various study contexts. These signifiers have been adapted to my specific study context, prior to being subjected to an empirical review to extend and refine the framework. The empirical review has developed my understanding of tutor co-regulatory practices and learner SRL processes. For example, the empirical review indicated that 'tutor-led' engagement was a broad category which encompassed situations in which the tutor was directing learners to take a specific action to solve a problem, as well as situations in which the tutor was explaining a concept to the learner to scaffold understanding. In order to more accurately represent the various forms of tutor-learner engagement, the 'tutor-led' engagement category was replaced by 'directive engagement', and 'explanatory engagement'. In addition, a new category of engagement was identified, in which the tutor is entirely passive, and simply observes or repeats the learner. I distinguished between guided/open-ended engagement, and passive engagement, as the

former were proactive teaching practices aimed at scaffolding engagement, while passive engagement was fully student led.

In order to validate the framework of signifiers of learner operations, in-person observations of learners using the VCE platform were conducted. A high rate of reconciliation (over 75%) between the observations and coded data indicates that the learner signifiers are valid and can be relied upon.

The final versions of the framework of signifiers for tutor co-regulatory practices and learner operations are presented in Tables 29 and 31. In Chapter 5, I will implement these frameworks using various forms of learning analytics, to explore the differences in how high-performing versus low-performing tutors co-regulate learners in a VCE.

Chapter 5: Analysing SRL and co-regulation

In this chapter, I analyse SRL and co-regulation using a combination of learning analytics. I start by using process mining to describe the latent processes through which tutors co-regulate learners during sessions in a VCE, identifying differences in the structure of the interactions. I move onto present my findings from implementing pattern mining, highlighting statistically significant patterns which occurred in sessions conducted by high performing tutors versus low performing tutors. Finally, I explore how significant patterns can be used as the basis for decision trees which classify tutors into high-performing and low-performing groups.

5.1 Process maps for sessions conducted by high-performing versus low-performing tutors

Introduction

In this Section, I will implement the framework of signifiers (developed in chapter 4.1) and analyse the patterns of labelled signifiers using process mining. As described in chapter 2.4 of this thesis, process mining enables tutor co-regulation of learners to be investigated from a dynamic and process-oriented perspective. This aligns with my conceptualisation of SRL, as a holistic process which is rooted in context and evolves with contextual changes.

Real-life processes, such as tutor co-regulation of learners, can be highly unstructured and are unlikely to conform to an ideal form. Process mining can offer valuable insight by discovering latent processes in a dataset. Through applying process mining techniques, such as the fuzzy miner algorithm, it is possible to visualise hidden processes at configurable levels of detail. These visualisations provide explorative and interactive analytics, which both:

1. Provides insight into the structure and distinguishing features of the processes under consideration
2. Acts as the foundation for the formulation of specific research questions, which can be investigated through further levels of statistical analysis

Table 35 below sets out the approach towards process mining, including a brief overview of each stage, and the relevant chapter section in which each stage of implementation is described. Following the implementation of process mining, I will discuss my findings in the context of the output from tutor interviews. I will conclude by summarising my findings and the further analysis planned.

Table 34. Stages of process mining mapped to chapter sections.

Stage	**Description**	**Chapter section**
Data sampling	Consider what the most appropriate approach towards data sampling is, in light of the research questions and the nature of the data-set	5.1.1
Select process mining algorithm	Explore a range of process mining tools and algorithms, and analyse them for suitability to my data-set and research questions	5.1.2
Data processing	Describe how data is processed into an event log	5.1.3
Production of process maps (data sample E)	Production of process maps from a range of perspectives and differing levels of detail	5.1.4
Production of process maps (data sample F)	Production of process maps from a range of perspectives and differing levels of detail	5.1.5
Analysis of process maps	Analysis of the key features and structural differences between process maps	5.1.6

5.1.1 Data sampling

Data was sampled by tutor performance, to allow an exploration of the differences in the way that learners are co-regulated by high-performing tutors versus low-performing tutors. From the perspective of process mining, sampling data from high-performing and low-

performing tutors is advantageous in that distinct processes are more likely to emerge amongst these groups. Bannert et al. (2014) noted that process mining could more effectively discover clear processes amongst extreme groups, and therefore chose to sample high-performing versus low-performing learners. I similarly chose to apply process mining to a data set of high versus low-performing tutors. However, it is important to note that while tutors have been sampled based on performance measures, there are no performance measures available at the level of a VCE session. The sessions from which data is sampled are therefore selected at random, from all the sessions conducted by high-performing and low-performing tutors (see 3.2 for more details). Some noise in the data is therefore possible. The effect of this noise was mitigated by examining process maps from different perspectives and configured to differing levels of details; this should have enabled a clear description of the underlying process to emerge, despite any noise within the data.

Two sets of criteria have been used as the basis for evaluating tutor performance:

- Human evaluator rankings for tutors, with regard to the student-centred teaching (SCT) pillar of evaluation
- Learning progress outcomes

As described in Chapter 3, TSL uses a rubric to evaluate tutor performance which includes scoring tutors on their ability to promote student centred teaching (SCT). SCT is similar (although not identical) to self-regulation, and human evaluator rankings for SCT may therefore act as a measure of how effectively tutors co-regulate learners. The human evaluator rankings are based on observations of the characteristics of randomly selected sessions. My research examined whether these rankings correspond with distinct granular patterns of tutor-learner interaction. As well as being an interesting area for research, the presence of these patterns could offer opportunities for later research into automated detection of favourable and maladaptive patterns of tutor-learner interaction.

In addition to using human evaluator rankings, average learning progress outcomes (APS) are piloted as an alternative measure of tutor performance. Past research has established a link between SRL and positive learning outcomes (e.g., van de Graff et al., 2022) and it is reasonable to hypothesise that learners with positive learning outcomes are being more effectively co-regulated. Exploring the use of an alternative measure of tutor performance is also useful in that it enables comparison between the different patterns of co-regulation, depending on the performance criteria applied, and offers a rich avenue for research.

The data sample used for process mining is set out in Section 3.2 of this thesis.

5.1.2 Selecting and applying the appropriate algorithm

As discussed in Section 2.4, the second stage of process mining is selecting the most appropriate algorithm for the scope of work. Process Mining consists of a wide range of algorithms, which vary in their ability to deal with complex processes, and noise within data (van der Aalst et al., 2011). Prior to selecting an algorithm, education researchers must consider how complex the behaviour patterns are likely to be, whether it is possible for learners to engage in a choice of patterns, and how much noise there is likely to be in the data.

Table 36 sets out the process mining algorithms that have been used in notable studies on process mining for learning processes, including SRL amongst other topics. As shown, the algorithms used include Inductive Miner, Fuzzy Miner, and pMineR, with the latter two algorithms being used most frequently. A systematic review conducted by Saint et al. (2021) stated that Fuzzy Miner was the most frequently used algorithm, used in 9 SRL studies. The pMineR has been used in 6 studies, while the Inductive Miner algorithm has been used in 2 SRL studies.

Table 35. Process Mining algorithms used in Learning Analytics research.

	Inductive Miner	**Fuzzy Miner**	**pMineR**

Toolkit	Pro M, BupaR	Pro M Disco Celonis	pMineR package
Metrics	Frequency, time	Frequency, time	Transition probability
Key features	Guarantees process Model soundness	Simplified visualisations, highlighting key events and transitions	First Order Markov Models, enabling quick comparisons
Examples of process mining studies	(Bogarín et al., 2018; Cerezo et al., 2020)	(Bannert et al., 2014; Maldonado-Mahauad et al., 2018; Sedrakyan et al., 2016)	(Ahmad Uzir et al., 2020; Matcha et al., 2019; Saint et al., 2020)

My work builds on the work done by Saint et al. (2021). In order to select an algorithm and toolkit for my research, I implemented a sample from my dataset using inductive miner and fuzzy miner. These two algorithms were chosen for comparison due to their differing approaches. Inductive miner distinguishes between choice and parallelism within the process models generated, and guarantees the model performance against conventional criteria (van der Aalst, 2012). In contrast, fuzzy miner does not distinguish between choice and parallelism within the process model generated, and the models produced do not accurately replay all pathways from the event log. However, fuzzy miner prioritises a clear visualisation of the process over the conventional criteria for process models, and seeks to emphasise the most prominent events and transitions. This can be very useful in unstructured datasets, where the research aim is to gain insight into the key features of the process, rather than to compare the process against a pre-defined structure.

In the following subsections of this Section, I select an algorithm for my research by implementing a data sample for low-ranked tutors (data sample E). I sought to compare the resulting models and metrics, and evaluate their research utility, in the context of my research, and dataset. At this stage of my research, I did not have any evidence on the different patterns which may emerge for the low-ranked tutor data sample versus the high-

ranked tutor data sample. The low-ranked tutor data sample was therefore selected at random for an evaluation of the research utility of the different algorithms.

Inductive Miner

Description and process model

The Inductive Miner algorithm works by finding major splits within the log, detecting relations, and defining sub-logs, and then repeating the process for the sub-logs. The models produced can follow a Business Process Model Notation, or a Petri Net. The algorithm can be implemented through Pro M, or Bupa R. The visual Inductive Miner version has been noted for its ability to show a clear process, with the animated replay feature enabling researchers to visualise where time is being spent in the process.

Inductive Miner guarantees a sound process model against all four conventional criteria (replay fitness, precision, simplicity, generalisability). While this can be effective whilst investigating structured processes, it can lead to complex process maps when dealing with unstructured processes such as SRL.

I experimented with the use of Inductive Miner with the dataset for the low-ranked tutors. The algorithm is configured to be robust to noise within the dataset, using the noise threshold parameter. This parameter can be configured to a value between 0.00 and 1.00, where a value of 0.00 guarantees perfect log fitness. Given the unstructured nature of the data-set, the noise threshold is set to a relatively high value of 0.60.

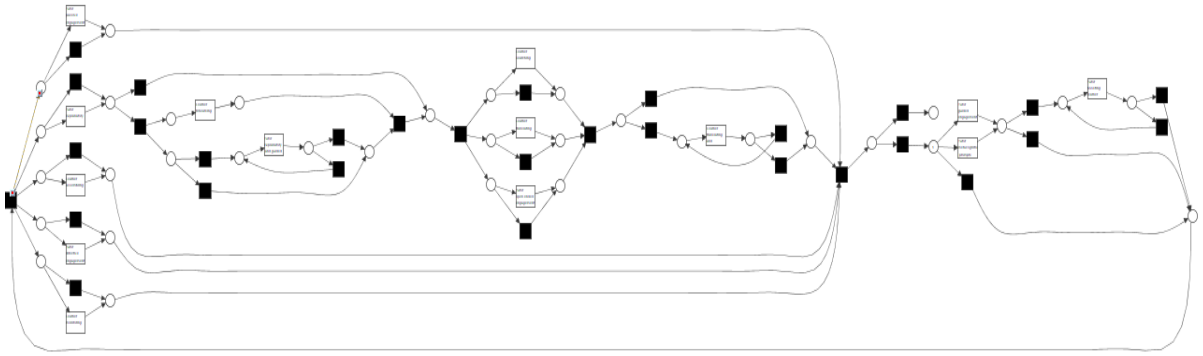


Figure 17. Petri model of low-ranked tutor dataset, mined with Inductive Miner algorithm

A brief definition of the components of the Petri net are provided below. Please refer to Appendix H for a more detailed description of how to read the process model:

- Circles: Represent places (states or conditions).
- Black Squares: Silent transitions (internal routing and synchronization).
- White Squares with Words: Labeled transitions (observable activities).
- Arrows: Arcs that show the flow between places and transitions.
- Tokens: Indicate the current state of the process

Evaluation of research utility

As illustrated, the resulting model is very hard to read or interpret. The model produced performs well against the conventional criteria for process model performance (van der Aalst, 2012), as evidenced by a comparison of the event log data with the petri net. However, the model is hard to read or interpret, as it shows multiple, siloed pathways, with the overall process model being looped on itself. This means that the model accurately replays the event data, but that it is challenging to develop a clear understanding of the process followed.

In addition, the model does not indicate which events or transitions are more important, than others. Parallelism used by the model obfuscates key transitions that may be occurring between events, with the main transitions shown occurring between blocks of events, rather than at the event level. The model prioritises accurate reflection of event data, over an insightful visualisation, and this contradicts my key research aim of discovering the processes underlying effective tutor co-regulation of learners. The inductive miner algorithm therefore appeared to be of limited research utility for my thesis.

Fuzzy Miner

Description, and process model

The fuzzy miner algorithm has been noted for its ability to produce simplified, descriptive visualisations of the processes occurring. While Fuzzy Miner can be useful in identifying what patterns are common to groups of learners, it is important to note that it cannot be statistically confirmed, as the output is descriptive (Bannert et al., 2014). However, the description of the process being followed can highlight the types of behaviour and interactions that are associated with positive learning outcomes. This can advance our understanding of SRL processes in new learning environments. A detailed description of how the algorithm works (including the metrics applied) is provided in the Appendix C.

The fuzzy miner algorithm can be implemented in Pro M, or the commercial versions Disco and Celonis. I chose to implement fuzzy miner in Disco, due to the optimal configuration of metrics, and holistic process graphs offered by the toolkit. Fuzzy miner in Disco has been refined, so that each node and transition is assigned a value from 0-100, using the metrics set out in the Appendix C, where 0 is least important and 100 is most important. Researchers can use interactive sliders to set the level of complexity to be displayed with regard to nodes and transitions, and Disco will automatically select default values for the metrics, and accordingly select the nodes and transitions which are to be displayed by the

graph. This feature of Disco means that less time needs to be spent by researchers on configuring the underlying metrics to display the desired level of detail in the process graph. Whilst configuration of metrics to adapt the level of detail displayed by the fuzzy miner algorithm is an interesting area for further research, it is not the focus of this current thesis. My work focuses on the implementation of process mining algorithms to derive insights on how tutors can co-regulate learners, rather than experimenting with the precise configuration of metrics used by the algorithm. In addition to enabling optimal configuration of the fuzzy miner algorithm, Disco also allows the user to view the model from different perspectives, with a built-in filter, allowing process models to be quickly re-built to include certain events, relations or attributes. Finally, while fuzzy miner in Pro M will cluster events to provide a higher-level view, and label each cluster as cluster 1, cluster 2, etc, the Disco version will produce a graph where each node is labelled as an event class. Disco therefore provides a holistic view of a process despite the abstraction of detail.

I experimented with implementing my low-ranked tutor dataset in Disco. The Disco tool was configured to show the most significant events and transitions. The resulting process model is shown below:

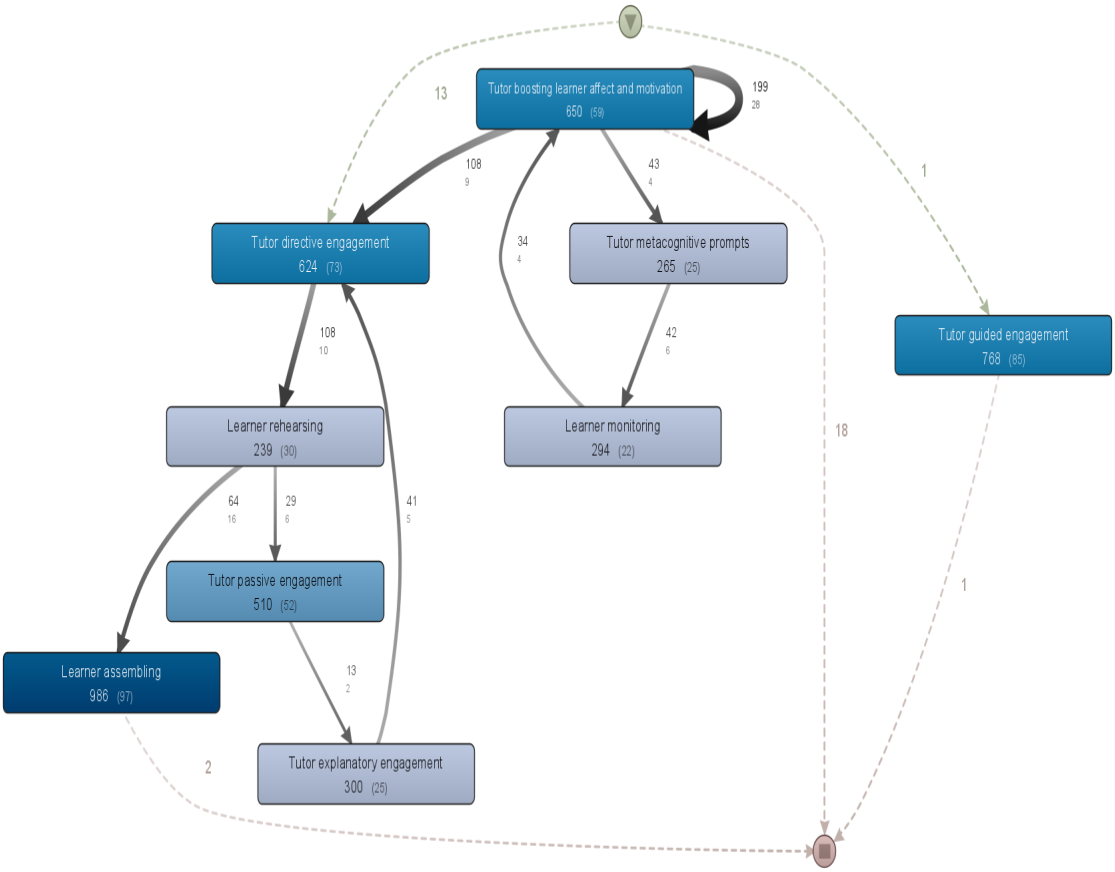


Figure18. Process model of low-ranked tutor dataset, mined with Disco algorithm

The Disco process model visualises the transitions between events and allows the researcher to identify key structural features of the processes observed in tutor-student dyads across the group. Important event classes and transitions are emphasised through the shading of the boxes and the thickness of the lines; darker shades and thicker lines denote more important event classes and transitions. The small triangle at the beginning of the graph signifiers the beginning of the process, while the rectangle at the end signifies the end of the process. Dotted lines are lines which have been artificially created to preserve a link between a node and the remainder of the process graph. The dotted lines represent the occurrence of less significant paths, while the vertical positioning of the boxes helps in understanding the order of activities within the process.

Evaluation of research utility

The process model produced by Disco is not sound, based on the conventional criteria for process models; it does not replay all the patterns found in the event log, and is not sufficiently precise. This is because the model is built using a combination of both log metrics and derivative metrics, based on which it prioritises certain patterns for visualisation over others. The visualisations produced offer insights into the key features of the process through which tutors co-regulate learners in a VCE. For example, for the low-ranked tutor group, it is immediately clear that ‘’Tutor directive engagement’ is a fundamental activity within the low-ranked tutor sessions, as illustrated by the colour of the node, as well as the thickness of the lines connecting it to other nodes in both directions. The ability of Disco to produce simplified visualisations which highlight the key structural features of the process makes it well suited to research in Education, particularly on SRL. SRL processes are complex and unstructured, as they are led by human agency. Descriptive models can thus be more useful to explore the development of SRL, rather than models which are statistically sound but lack clear visualisations. I therefore chose to use Disco to further our understanding of how tutors can promote or inhibit the development of SRL amongst learners in a VCE. The findings from Disco were validated and advanced using statistical analysis, such as pattern mining and decision trees.

5.1.3 Data processing and configuration of parameters

Having chosen an appropriate process mining tool by working on a small subset of the available data, I now describe the next stage of process mining for the whole data set. In the next stage of process mining, the event data is processed into an event log. In order to build an event log for Disco, event data is required which comprises of the following fields:

- Trace: This refers to the identifier for the sequence of cases being analysed. In my research, the trace is the ID for the VCE session from which data has been collected

- Event: This refers to the event occurring. Events can be defined at different levels of granularity, e.g., a type of practice versus a micro level signifier. In my research, the events analysed are defined by the framework of signifiers set out in Section 4.2 of this thesis. I start with analysing processes from the perspective of tutor practice, and learner operations categories, before considering processes from the perspective of the micro signifiers recorded.
- Time-stamp. This can refer to the event began, or when it was completed. If both a start and end time stamp are recorded, process maps can be used to analyse the median time lag between events, and assess performance and bottlenecks within the process. In the case of my research, only a start time-stamp is noted during data collection given the manual nature of the coding process, and this is simply used by the algorithm to establish an ordering of events.

The Appendix D sets out further details of how the data was processed into an event log and imported into Disco.

5.1.4 Producing process maps

As described in the sections above, process mining is an explorative and iterative process. The aim of process mining is to gain insights into latent processes, which can then be validated and refined through other forms of analysis. In order to implement process mining effectively. I produced and analysed process maps from different perspectives, and at different levels of detail as explained below.

I will start by describing process maps for sessions conducted by high-performing versus low-performing tutors, where performance was determined by human evaluators (data sample E), followed by repeating this approach for data sample F, where performance was determined by learning outcomes). I analysed processes from a range of different perspectives including:

- From **different levels of granularity:**
    - Implementing process mining for categories of tutor practice types and learner operation types.
    - Producing process maps which use the individual signifiers of tutor practices and learner operations.
- From **tutor versus learner perspectives:**
    - I filtered process maps to examine tutor and learner behaviours in isolation, for the groups of sessions conducted by high-performing and low-performing tutors. The aim was to focus with greater depth on how the practices adopted by high-performing tutors differ from low-performing tutors, and similarly, how the behaviours adopted by learners differ when sessions were led by high-performing versus low-performing tutors.

Process maps are initially described at high levels of detail, with only the most important events and transitions being displayed. I then move onto examine the process maps at increasing levels of detail, to gain further insight into the key events and transitions denoted in the initial process map.

5.1.5 Process mining, using data sample E (using evaluator rankings as the criterion for tutor performance)

This subsection follows the structure below:

i) Process maps produced, using categories of tutor practices and learner operations
    (1) Analysis of fundamental differences between process maps with reference to more detailed process maps
ii) Process maps produced using signifiers of tutor practices and learner operations
    (1) Analysis of fundamental differences between process maps with reference to more detailed process maps

iii) Process maps produced from tutor versus learner perspectives

(1) Analysis of fundamental differences between process maps with reference to more detailed process maps

iv) Synthesised analysis of the insights from process mining for data sample E.

(i) Process maps produced using categories of tutor practices and learner operations

I implemented process mining for the sessions conducted by high-ranked and low-ranked tutors (data sample E), using the framework of signifiers set out in Section 4.2. The initial process maps were produced using categories of tutor practices and learner operations as event classes. Given the relatively high number of event classes, the process maps were initially configured to show only the most important events and transitions. The resulting process maps are set out in Process map 1 (for high-ranked tutors) and Process map 2 (for low-ranked tutors). For convenience, the online links to process maps are also provided below:

Process map for sessions conducted by high-ranked tutors using event categories and sliders set to 0 percent

Process map for sessions conducted by low-ranked tutors using event categories and sliders set to 0 percent

As described elsewhere in this thesis, the colour of the node represents the importance of the event node to the overall process, with darker nodes being more prominent in the process. The thickness of the line represents the importance of the transition. The Appendix C sets out the metrics which are used to judge the importance of event nodes and transitions, and these include a combination of log based and derivative metrics. The numbers set out in the process map for event nodes and transitions refer to the absolute frequency of the event or transition within the data sample. These metrics are useful for further understanding the prominence of events and transitions within the overall process. However, it is important to

note that algorithm determines the importance of events and transitions using a combination of metrics, extending beyond measures of frequency.

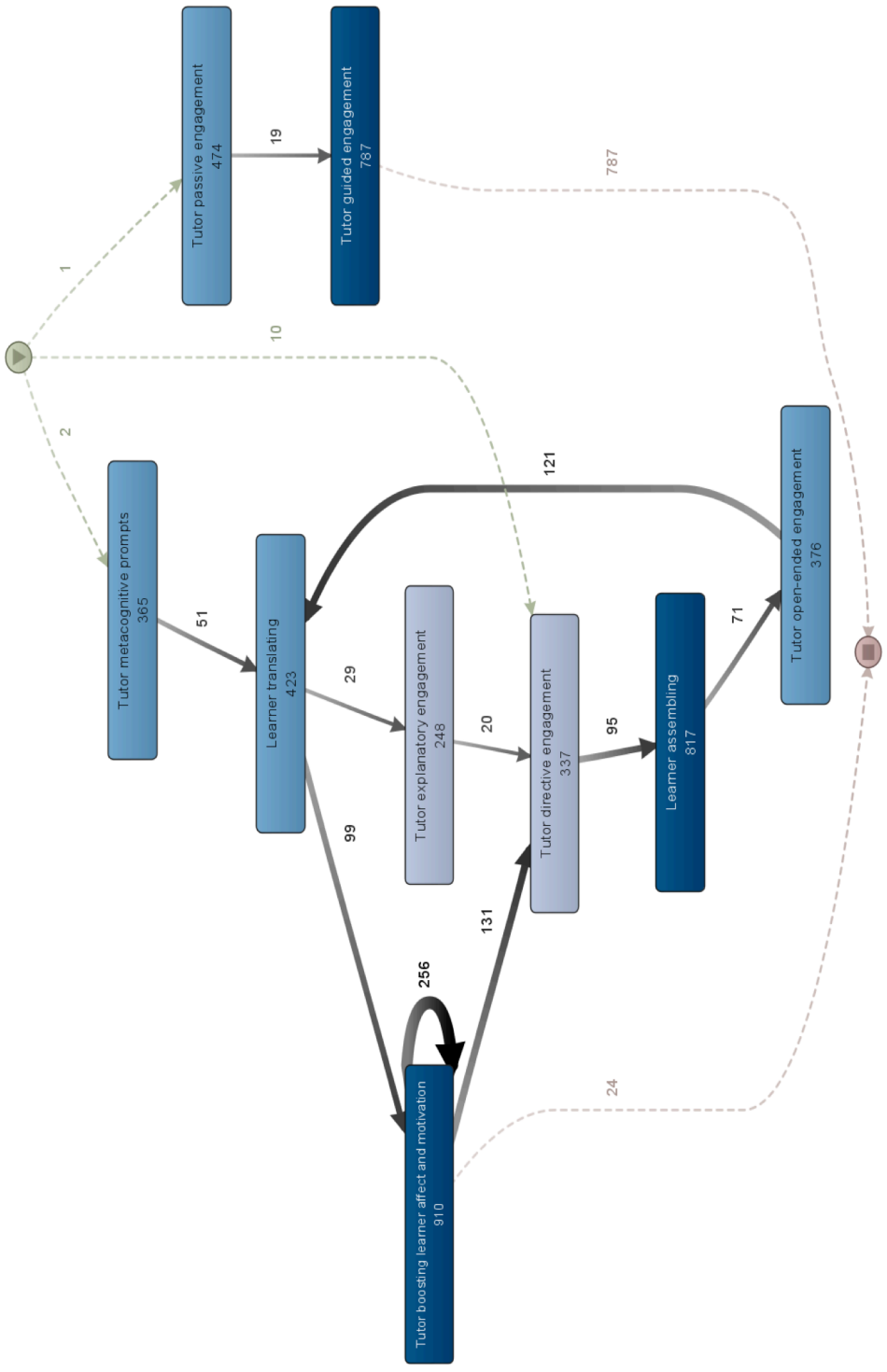


Figure 19. Process map 1, for high-ranked tutor sessions (sliders set to 0%)

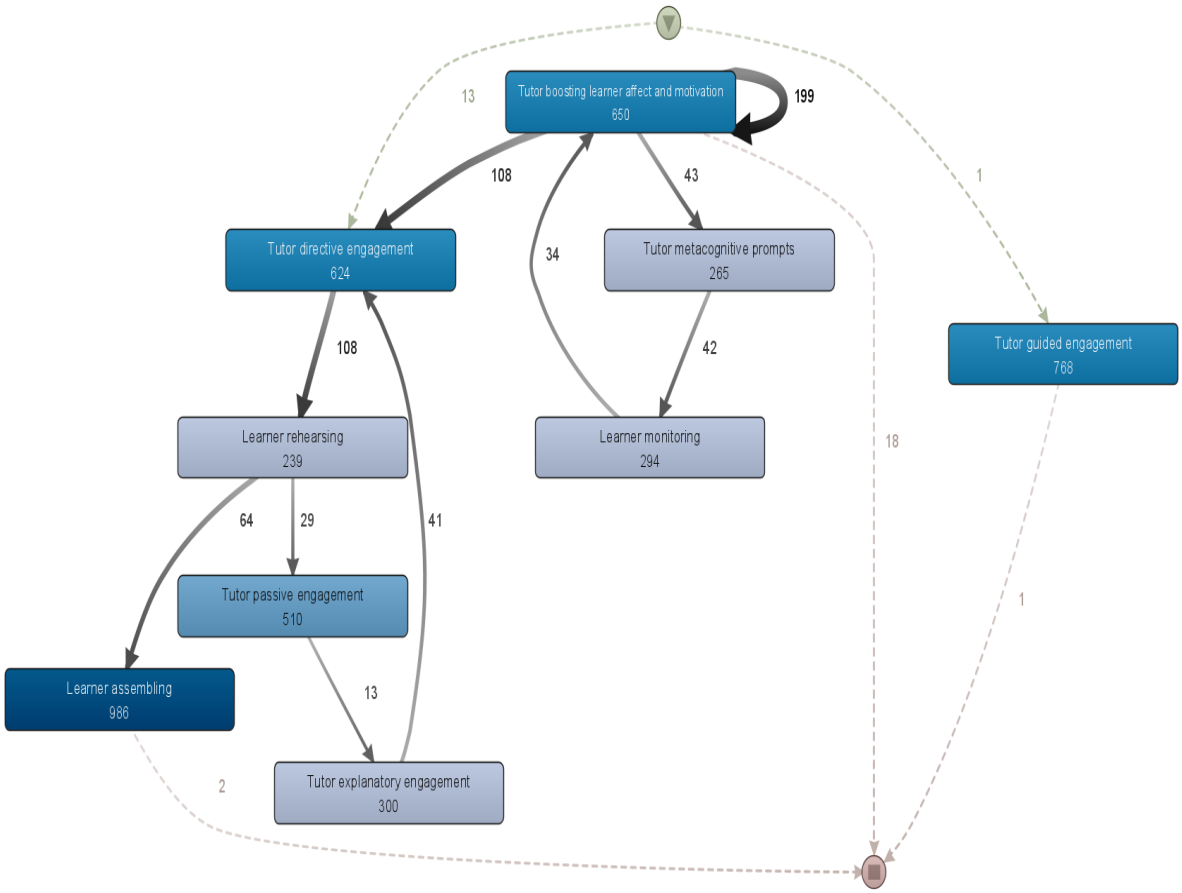


Figure 20. Process map 2 for low-ranked tutors (sliders set to 0%)

Analysis of process maps produced using categories of tutor practices and learner operations

My analysis found that the process maps for the low-ranked tutor group differed from the process maps for the high-ranked tutor group in terms of both the structure of the map, and also with regard to the types of prominent events and transitions. Overall, there is a relatively narrow range of tutor-centred events that are considered important in the low-ranked group, while high-ranked tutors demonstrate a more diverse range of open and interactive practices.

Low-ranked tutors heavily rely on a small number of practices, out of which most are tutor-centred and cede limited autonomy to the learner. Some of the significant observations are listed below.

- **Tutor directive engagement** is highly prominent, both in terms of the significance of the event and key transitions to the remainder of the process. Tutor directive engagement refers to the tutor instructing the learner to take a particular action, or adopt a specific learning tactic, e.g. to read a question aloud or utilise a column chart. It is a tutor-centred practice and offers limited control to the learner.
- **Tutor directive engagement frequently transitions to learner rehearsing**, indicating the tutor scaffolding learner through encouraging shallow learning activities such as reading aloud from the slide.
- There is also a relatively **frequent transition from tutor directive engagement – learner rehearsing- tutor passive engagement – tutor explanatory engagement**. Tutor passive engagement refers to the tutor observing the learner without offering any constructive input. Whilst control is handed to the learner here, the transition to explanatory engagement indicates that the practice tends to be used by low-ranked tutors to diagnose the specific types of support required by the learner and to plug the gaps through explanation, rather than to encourage open-ended engagement from the learner.
- There is a **loop between tutor explanatory engagement-tutor directive engagement**, indicating that **low-ranked tutors repeatedly engage in tutor-centred paths** of interaction with learners.
- **Tutor-learner interaction appears to be closely aligned to the VCE problem-solving activities** rather than extending to conceptual dialogue, as indicated by the

prominence of the learner assembling node versus the absence of the learner translating node from the process map.

- **Open-ended, conceptual dialogue is not a significant feature** of the process, as indicated by the absence of tutor open-ended engagement, or learner translating from the high-level process map.
- In addition to the tutor-centred paths of interaction (starting with tutor directive engagement), there is a **separate path of interaction which encourages learner metacognition**. However, this path is **relatively weak**, and appears to be relatively isolated from the remainder of the process, with no recurring or strong transitions evident.
- **Tutor boosting learner affect and motivation** is a **significant feature** of the process and acts as a starting point for a transition to tutor directive engagement node, and tutor metacognitive practices.

In contrast to the low-ranked tutors, sessions conducted by the **high-ranked group exhibit a broader range of significant events and transitions, including open-ended tutoring practices** which encourage more sophisticated learning operations.

- **Learner assembling** is **highly prominent** within the process map, both in terms of the unary significance of the node, and the key transitions to other nodes. Learner assembling has **strong relation with tutor open ended engagement, which loops back to learner translating.** This transition indicate that **high-ranked tutors use learner activities within the VCE as a foundation to encourage sophisticated learner operations,** such as learner reasoning.
- **Tutoring practices which encourage open ended engagement and metacognition also feature prominently.** Both open-ended engagement and metacognitive practices appear to be well integrated with the overall process.

- There is a transition from **tutor passive engagement to tutor guided engagement.** This indicates that **tutors may use passive engagement to diagnose gaps in learning**, and **scaffold the learner through tutor guided engagement** rather than tutor explanatory engagement.
- **Tutor boosting learner affect and motivation event node is well embedded within the process**. It frequently follows learner translating, indicating that the high-ranked tutors reward learners for engagement in sophisticated learning operations.
- From the learner perspective, there is also a greater range of operations displayed, with **learner assembling and learner translating both featuring prominently** in the process map.

While the initial process maps provide valuable insight into the key differences between the process maps for high ranking versus low-ranking tutors, there is a need to examine the processes in more detail. In particular, to identify a wider range of transitions which occur for key events. The sliders for the process maps in Disco were therefore configured to show only the most important events (i.e., set to 0%), but to include the most important 25% of all transitions (i.e. the paths slider is set to 25%). The resulting process maps are set out in process maps 3 and 4 (in the Appendix F), for sessions conducted by low-ranked and high-ranked tutors respectively, and the corresponding online links are also provided below.

Process map 3 LOW HE tutors, event categories, paths sliders at 25 percent

Process map 4 HIGH HE tutors, event categories, paths sliders at 25 percent

The **process map for sessions conducted by high-ranking tutors is broadly similar, regardless of whether the paths slider in Disco were set to 0% or 25%.** The main difference is that an additional transition for tutor guided engagement is displayed, between tutor engagement- tutor boosting learner affect and motivation. A new path, tutor passive

engagement - tutor guided engagement – tutor boosting learner affect and motivation, is also displayed. These transitions indicate that high-ranked tutors can use the practice of observing learners to micro scaffold learners where necessary, with learners being encouraged and rewarded for their effort. The new transitions displayed also demonstrate that guided engagement is well-integrated with the remainder of the process discovered.

The process map for sessions conducted by **low-ranked tutors** differs, depending on whether the paths slider is set to 0% or 25%, in the following ways:

- **The learner rehearsing event node is more prominent in the detailed map**, as multiple transitions are shown from this event to other parts of the process. Shallow learning operations are therefore a key part of the process for sessions conducted by low-ranked tutors, despite the low unary significance of the event class.
- The **Tutor boosting learner affect and motivation** node appears to be **more connected** to the remainder of the process, in the detailed map, with a number of transitions ending in tutor boosting learner affect and motivation being shown. However, on close examination, the new transitions are:
    - Tutor directive engagement – tutor boosting learner affect and motivation.
    - Learner rehearsing – tutor boosting learner affect and motivation.
    - Tutor boosting learner affect and motivation – learner rehearsing.

  Each of these paths starts or ends with a shallow learning operation (learner rehearsing), or a tutor led practice (tutor directive engagement). Tutors thus appear to be **rewarding shallow engagement within learning, or compliance with tutor instructions**, rather than embedding praise and reward within a more diverse range of events.

- Similarly, an additional path is also discovered, which integrates tutor guided engagement into the remainder of the process: **Tutor-guided engagement – tutor explanatory engagement.**

Tutor guided engagement appears to be used to scaffold learners through micro questions, with tutor led explanations being used to further plug any gaps in understanding. This contrasts with the process discovered for high-ranked tutors, which uses tutor guided engagement as a starting point for open ended forms of engagement, and encouraging learner reasoning.

(ii) Process maps produced, using signifiers of tutor practices and learner operations

I have implemented process mining using the signifiers of tutoring practices, and learner operations, as the event classes. The process maps are set out below, and the corresponding online links to the maps are also provided below. These maps show only the most important events and transitions (i.e., the event and path sliders are set to 0%).

Process map 5 LOW HE tutors signifiers view sliders at 0 percent

Process map 6 HIGH HE tutors signifiers map sliders at 0 percent

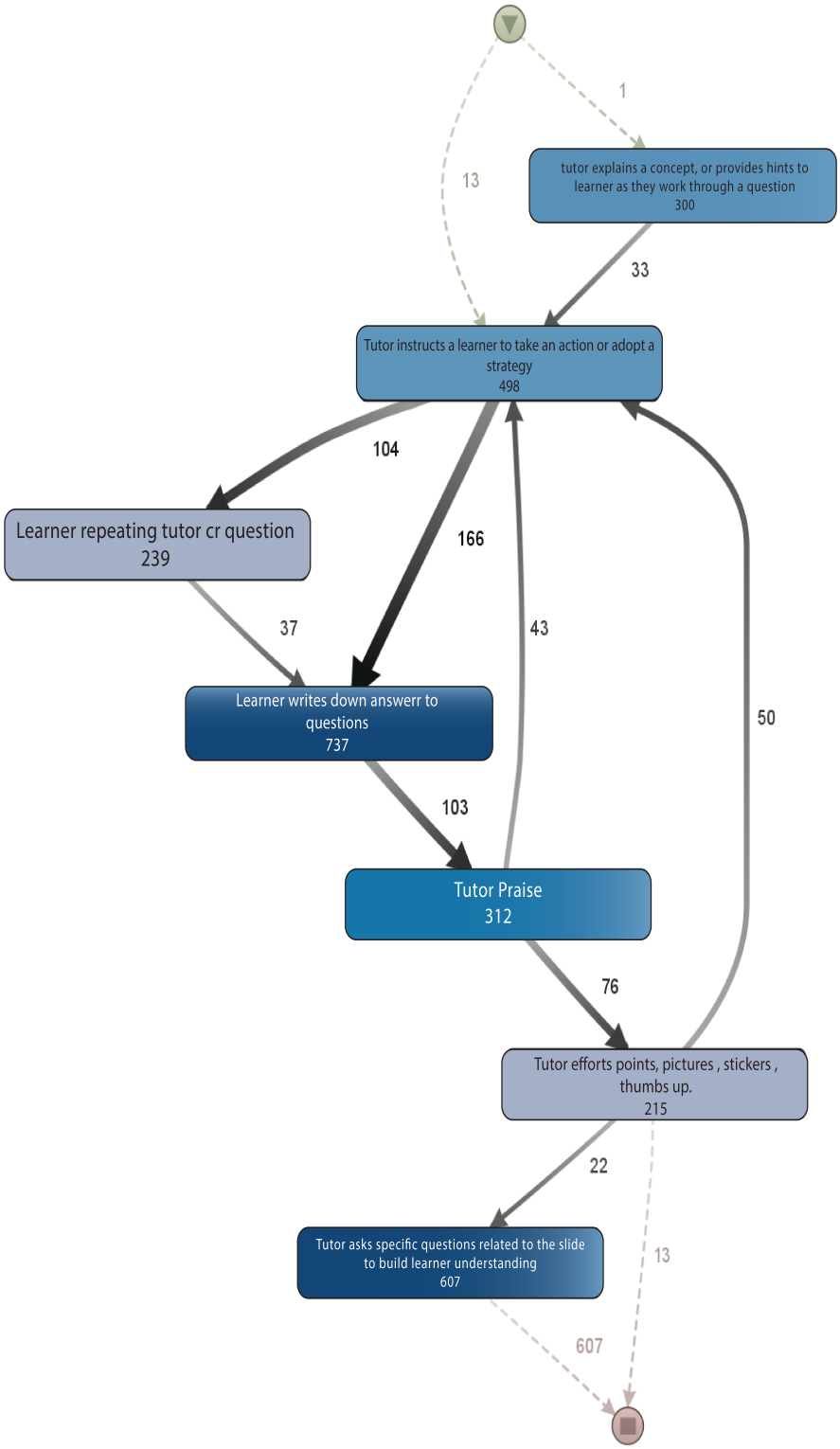


Figure 21. Process map 5 low-ranked tutors (signifiers view, sliders set to 0%)

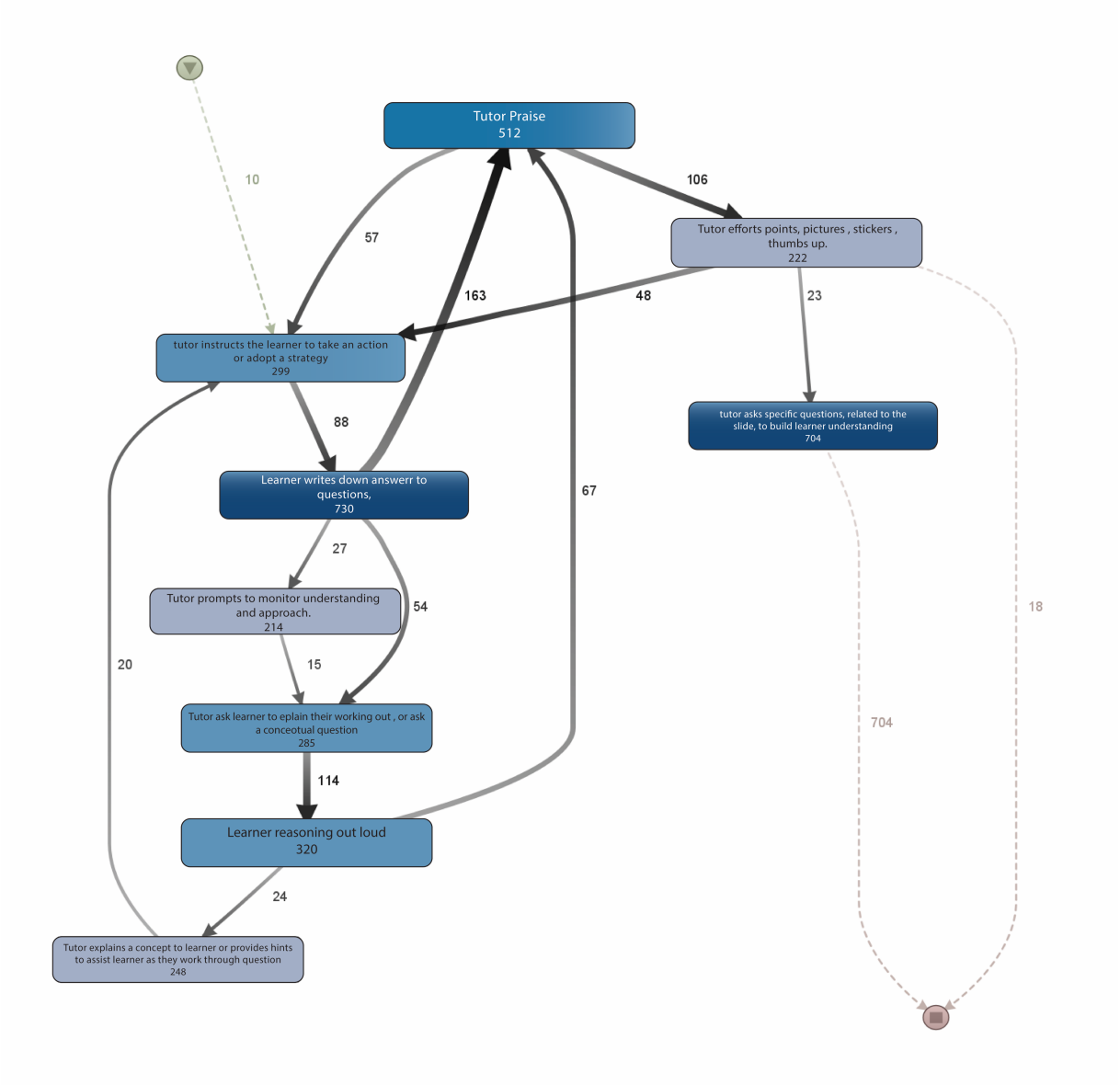


Figure 22. Process map 6 high-ranked tutors (signifiers view, sliders set to 0%)

Analysis of process maps produced, using signifiers of tutor practices and learner operations

The process maps set out in Figure 21 and 22 discover a more granular process, as they analyse fine-grained signifiers rather than macro event categories. For example, the macro event category tutor directive engagement includes the signifiers below:

- Tutor utterance which instructs the learner to take an action or adopt a strategy.
- Tutor interrupts the learner to take control of the session activity or dialogue.
- Tutor provides the answer to the question, either orally or in writing.

A process map which uses event categories as the event class would only represent that tutor directive engagement has occurred. In contrast, process maps which use signifiers as the event class will show the dominant events themselves. While **process maps which use the macro categories are helpful in aligning processes to the theoretical framework, process maps which use signifiers provide a practical representation** of the process. A process map built on event signifiers, can be useful in the context of a VCE, to evaluate how tutors co-regulate learners in practice.

Analysis of the process maps in Figure 21 and 22 illustrates that the fundamental features of the process maps remain the same, regardless of whether the event nodes are based on signifiers rather than event categories. However, analysis of multiple fine-grained signifiers, rather than a more condensed selection of macro event categories, results in certain event nodes and transitions being more or less prominent than expected. For sessions conducted by high-ranked tutors:

- The process map continues to demonstrate a complex process, with a greater emphasis on open ended tutoring practices. For example, there is a strong transition between ‘learner writing answers to questions’ and ‘tutor asks learner to explain their reasoning or asks a conceptual question’.
- Increasing the level of detail in the process map, such that a greater number of transitions are included, produces the process map linked here: Process map 8. This reveals interesting new transitions between signifiers. In particular, a signifier of tutor open-ended engagement (tutor asks learner to explain their reasoning or asks a conceptual question’) is shown to have a number of key

transitions to other events, despite a low unary significance. This illustrates how tutor open-ended engagement practices are fundamental to the structure of the process in sessions conducted by high-ranking tutors. This also illustrates the **value of developing process maps using fine grained signifiers**, to highlight the **key role** that **certain fine-grained events** may play within a process.

- Certain macro event categories, such as tutor metacognitive prompts and tutor passive engagement, are not reflected in the signifiers-based process map. This is most likely because of the low number of types of signifiers which comprise these event categories. A low number of types of signifiers will distort the prominence of the relevant event categories, when a signifiers-based view is adopted.

For sessions conducted by low-ranked tutors:

- The process map continues to highlight a simpler process, dominated by tutor-centred practices, such as 'tutor instructs learner'. This signifier has a high unary significance, as well as a number of key transitions to other events.
- Despite a low unary significance, **shallow learning operations such as learner repeating tutor**, are discovered as having key transitions to other events and being **an integral part of the overall process.**
- In a similar manner to the high-ranked group, certain macro events, such as tutor metacognitive prompts and tutor passive engagement are not reflected in the signifiers-based process map.
- Increasing the level of detail in the process map, such that a greater number of transitions are included, produces the process map linked here: Process map 7. This process map further reinforces the **prominent role of 'learner repeating the tutor or question'** within sessions, which has multiple transitions to other events. It is also clear from this map that **tutor practices are very closely aligned with**

the problem-solving activities of the VCE; the 'tutor instructing-learner writing-tutor asks specific questions' path is by far the most prominent pathway in this map.

(iii) Process maps produced from the tutor and learner perspective

I have also analysed process maps by filtering them from the **tutor perspective** and the **learner perspective**. The aim is to explore tutor and learner behaviour in isolation, in particular analysing the key transitions that build up overall patterns of tutor and learner behaviours.

The process maps are set out as below in Figures 23-26, and have been configured to represent the most important events and the 25% of the most important transitions (i.e., event sliders set to 0% and paths sliders set to 25%). Signifiers of tutor and learner behaviour are used as the event class, to enable a granular representation of the processes followed.

Process map 9_Tutor filter_HIGH HE group

Process map 10_Learner filter_HIGH HE group

Process map 11_Tutor filter_LOW HE group

Process map 12_Learner filter_LOW HE group

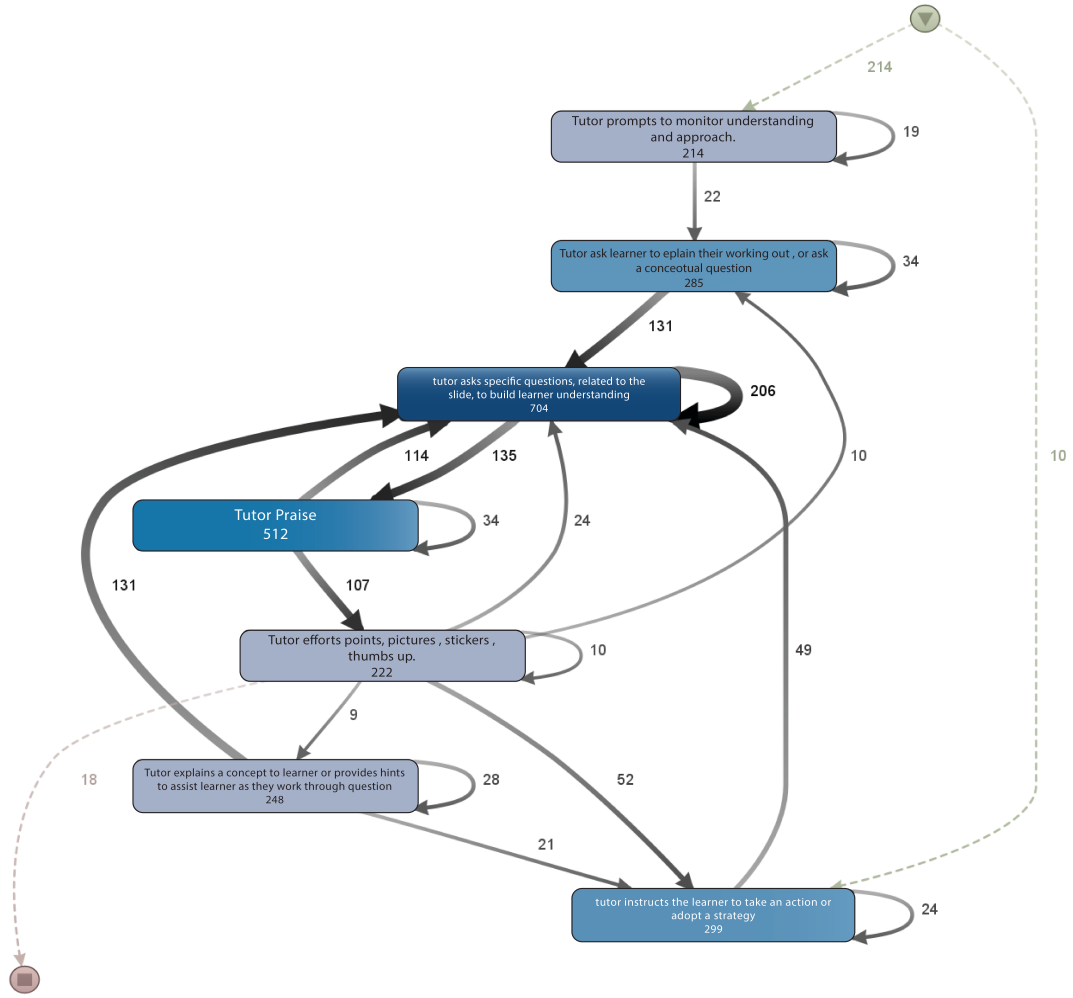


Figure 63. Process map 9 High-ranking tutor sessions (filtered by tutor behaviours)

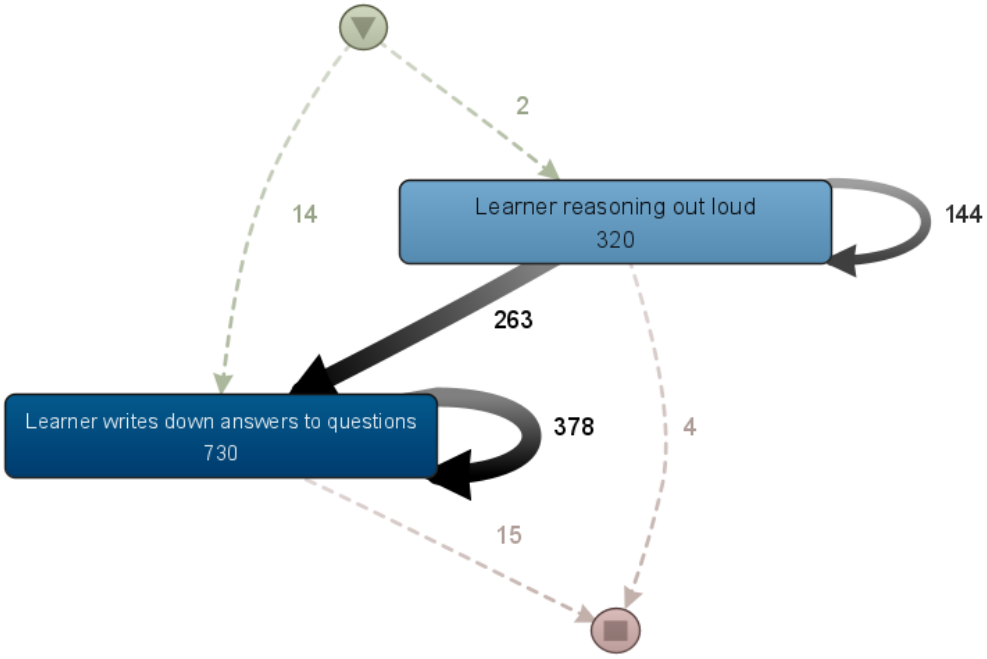


Figure 24. Process map 10 High ranking tutor sessions (filtered by learner behaviours)

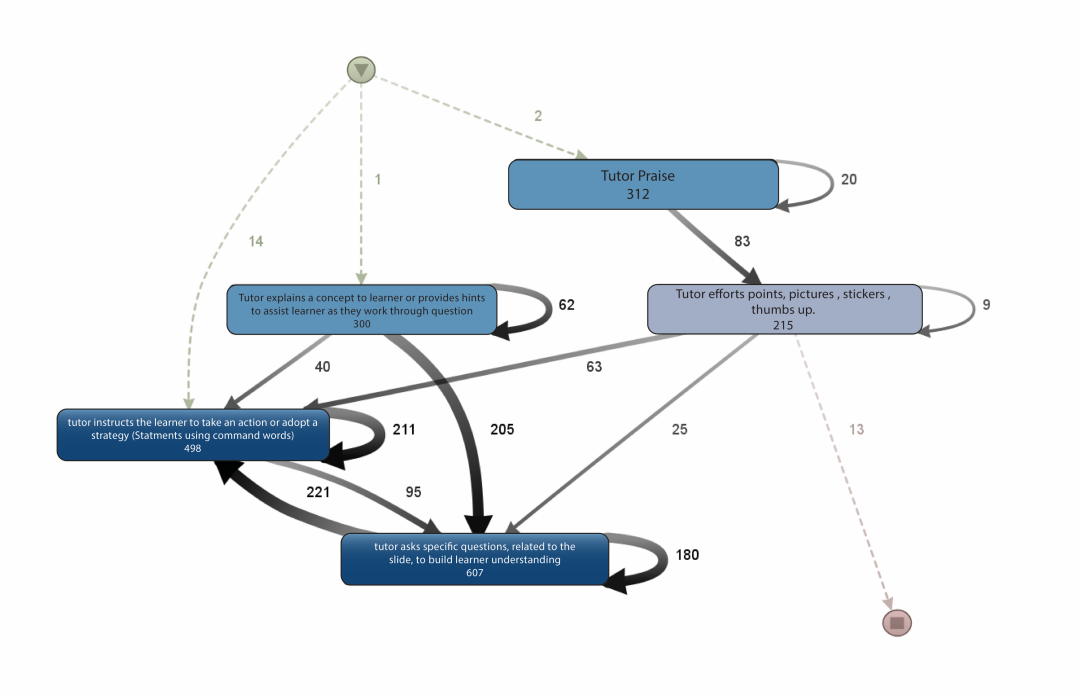


Figure 25. Process map 11 Low-ranking tutors (filtered by tutor behaviours)

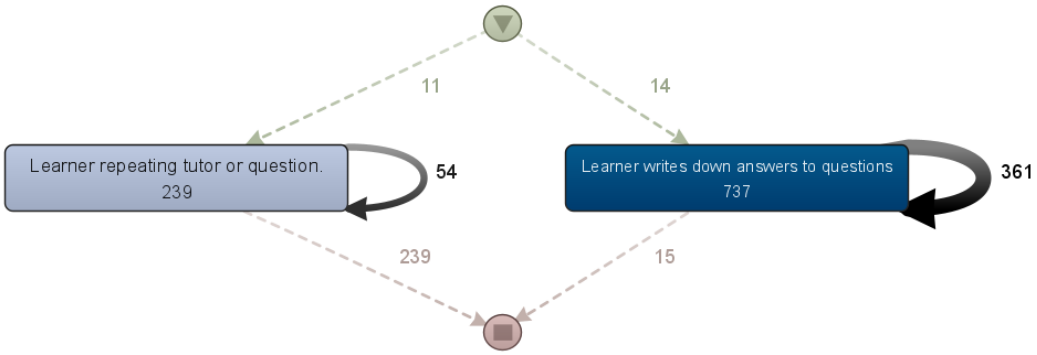


Figure 26, Process map 12 Low-ranking tutors (filtered by learner behaviours)

Analysis of process maps produced from the tutor and learner perspective

Examining the process maps filtered for tutor practices demonstrates a number of key differences in the practices adopted by high-ranked tutors versus low-ranked tutors. Some of the significant insights are summarised below.

- **High-ranked tutors** demonstrate a **diverse range of practices**, and tutors frequently transition from one type of tutoring practice to another. For example, tutors frequently transition from 'tutor asks learner to explain their working out, or asks a conceptual questions' to 'tutor asks specific questions to build learner understanding', as indicated by the thickness of the dark lines connecting the event nodes. Similarly, there are multiple transitions connecting onwards to a range of other practices, such as tutor praise or explanation. While **instructional tutoring practices** also feature in the process map, they do **not appear to be heavily embedded** within the overall process.
- In contrast, **low-ranked tutors** appear to rely on a **narrower range of tutor-centred practices**, such as tutor instructing learner. There are strong transitions to and from 'tutor instructs learner' to 'tutor asks specific question, and there is also a strong

transition from ‘tutor explains concept’ to ‘tutor asks specific questions ‘. While the event ‘**tutor asks specific questions’** is also prominent for high-ranked tutors, it frequently occurs after the tutor asks learner to explain their working out, or asks conceptual questions’, indicating that it is potentially used as a form of scaffolding. In contrast, the ‘tutor asks specific questions’ event node is **connected to other forms of tutor-centred practices**, indicating that low-ranked tutors may potentially prefer tutoring practices which allow them to retain their control and limit learner agency.

Similarly, the process maps filtered for learner practices illustrate that **learners in sessions conducted by high-ranked tutors are more likely to demonstrate higher order operations**, such as ‘learner reasoning’. In contrast, **learners in sessions conducted by low-ranked tutors are more likely to demonstrate shallower learning operations such as ‘learner repeating the tutor or a question’**. I have checked the data sample for a skew in learner attainment by data sample (see Section 3.2), but have not found one. It is therefore possible to argue that the difference in learner operations may be driven by the tutoring practices, rather than vice versa.

(iv) Summary of analysis from process mining (using tutor rankings as criteria for tutor performance)

The use of process mining has discovered distinct latent processes underlying the sessions conducted by high-ranked versus low-ranked tutors. These differences include (1) the range and patterns of tutor-learner interactions and, (2) the varied use of the VCE by the groups of tutors with differently ranked performance.

With regard to the patterns of tutor-learner interactions, the sessions conducted by low-ranked tutors tend to center on a small number of practices, out of which most are tutor-centred and cede limited autonomy to the learner. Low-ranked tutors repeatedly engage in tutor-controlled paths of interaction with learners, with forms of directive engagement being

particularly prominent within the overall process. Open ended, conceptual dialogue does not appear to be a significant feature of tutor-learner interactions. In contrast, high-ranked tutors utilise a more diverse range of practices, with open ended engagement featuring prominently. High-ranked tutors appear to deploy a range of practices to build on from learner solving problems within the VCE, as demonstrated by the multiple transitions between learner assembling and interactive and open-ended forms of engagement.

High-ranked versus low-ranked tutors also seem to use the same tutoring practice to different effect. For example, both high-ranked and low-ranked tutors utilise the practice of observing learners ('tutor passive engagement') to diagnose gaps in learning. However, in the case of low-ranked tutors, this practice most often leads to tutor explanations to plug the gaps in learning. In the case of high-ranked tutors, a more interactive form of engagement (micro questions) is used to plug diagnosed gaps. Similarly, praise and reward are heavily embedded in the maps for both categories of tutor. However, these practices are linked to shallow learning operations in the low-ranked tutor map, and to more sophisticated learning operations, such as reasoning, for the high-ranked tutor map.

Producing the process maps using signifiers of event categories rather than the categories themselves enables a granular and practical representation of the process for each group of tutors. In particular, signifiers of shallow learner operations, such as learner repetition, are shown to be particularly important in the low-ranked tutor map. Conversely, signifiers of sophisticated forms of tutor-learner engagement, such as 'tutor asking conceptual questions', are demonstrated to play a key role in the high-ranked tutor maps.

The differences in the discovered processes for the sessions conducted by high versus low-ranked tutors point to a difference in the way that the VCE platform is used, by the two groups. High-ranked tutors appear to use the platform as a starting point, from which tutor-learner interactions are built, with learning being extended through rich, open-ended forms of

engagement, such as sophisticated dialogue. In contrast, low-ranked tutors appear to be constrained by the platform, and tend to restrict their practices towards guiding the learners through the activities set out in the platform.

5.1.6 Process mining using data sample F (using average learner progress scores as the criterion for tutor performance)

In the following section, I have experimented with using a data sample that applies learner average progress scores, as the criterion for tutor performance. The aim is to discover whether the co-regulatory process differs with learning outcomes. There is a vast body of research which associates self-regulation with positive learning outcomes (see van de Graff, 2022). It is therefore a reasonable hypothesis that the effectiveness with which tutors co-regulate learners will influence learning outcomes. I will also explore the impact of using average progress scores rather than human evaluator rankings as the criterion of tutor performance, on process discovery. I will consider whether the application of one criterion, rather than another, results in clear differences in the processes adopted, and why this may be.

This section will follow a similar structure to the structure adopted above for data sample E (based on human evaluator rankings of tutors).

i) Process maps produced, using categories of tutor practices and learner operations
    (1) Analysis of fundamental differences between process maps with reference to more detailed process maps
ii) Process maps produced using signifiers of tutor practices and learner operations
    (1) Analysis of fundamental differences between process maps with reference to more detailed process maps
iii) Process maps produced from tutor versus learner perspectives
    (1) Analysis of fundamental differences between process maps with reference to more detailed process maps

iv) Synthesised analysis of the insights from process mining for data sample F.

(i) Process mining, using data sample F (using average learner progress scores as the criterion for tutor performance)

To start with, I produced process maps at the highest level of detail, including only the most important events and paths (i.e. event and paths sliders set to 0%). The resulting process maps are presented below: Figure 27 shows the process map 13 for tutors with a high learner average progress score (APS), while Figure 28 shows the process map 14 for tutors with a low APS. The corresponding online links to the process maps are also provided below.

Process map 13_HIGH APS_event categories_sliders at 0 percent

Process map 14_LOW APS_event categories_sliders at 0 percent

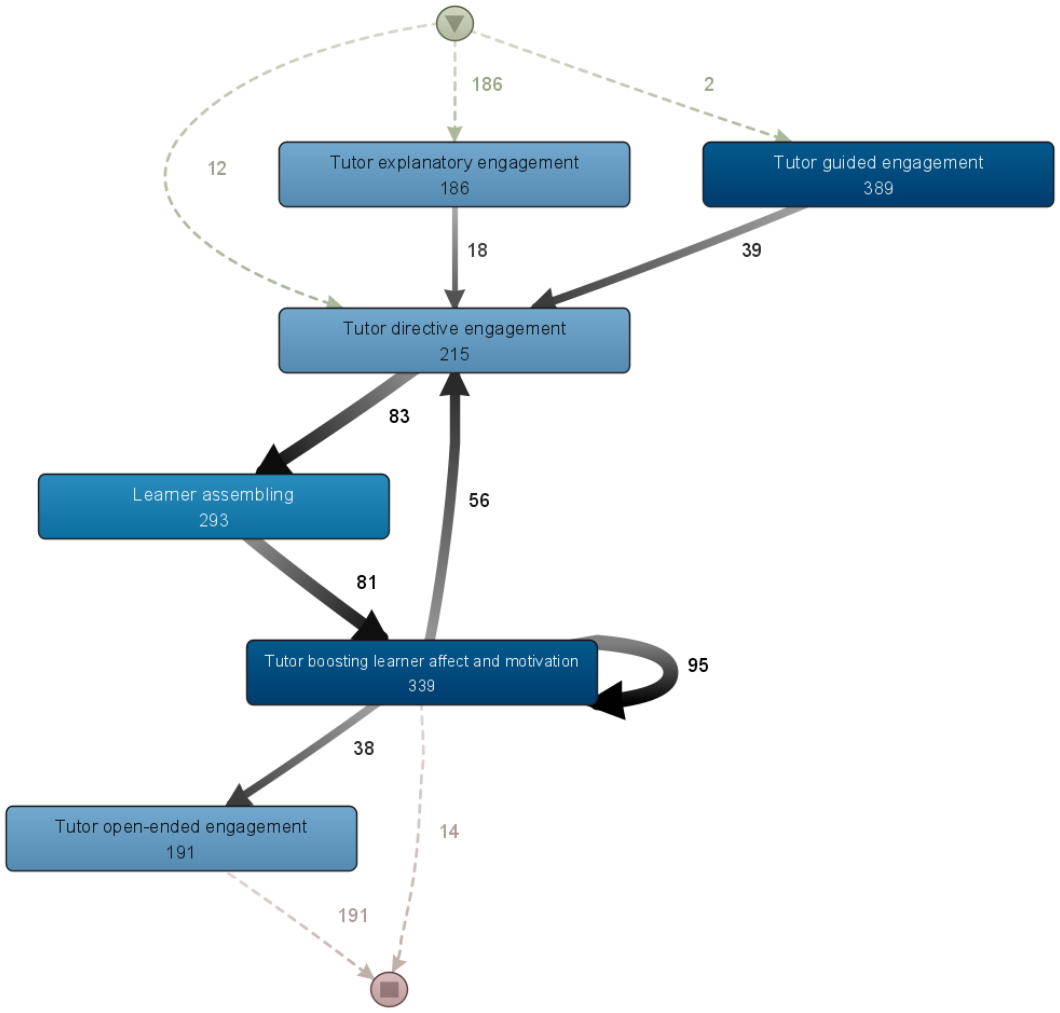


Figure 27. Process map 13 high APS tutors (event categories, sliders at 0%)

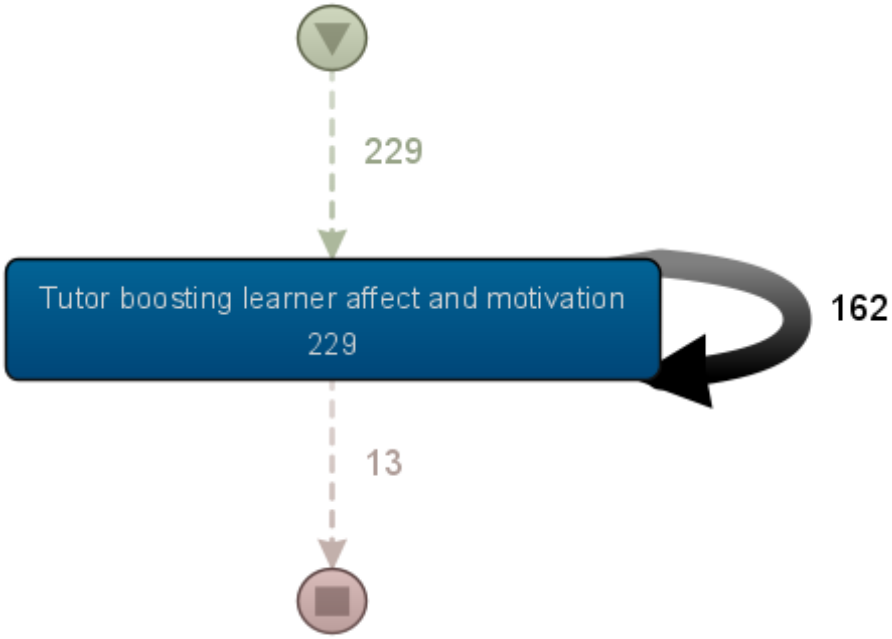


Figure 28. Process map 14 low APS tutors (event categories, sliders at 0%)

Analysis of process maps using data sample F (using average learner progress scores as the criterion for tutor performance)

As shown in Process map 14, the process map for tutors with a low APS only includes one event – 'tutor boosting learner affect and motivation'. This is interesting in itself, as it indicates that tutor practices, such as **praise and reward**, are **disproportionately prominent** for sessions conducted by **tutors with a low APS**. In contrast, the process map, for sessions conducted by **tutors with a high APS**, shows a **broader distribution of tutoring practices**, including a mixture of tutor-centred practices (e.g., tutor directive engagement), and student-centred practices (e.g., tutor open-ended engagement).

In order to enable further analysis of the key events and transitions, the level of detail in the process maps is increased, with event sliders being set to 10% and the paths sliders

being set to 25%. This reveals a difference in the types of tutoring practices that are prominent within the two groups of sessions. As illustrated by Process map 15 and Process map 16, the process maps differ in that there is broader range of significant tutoring practices within the high APS map:

Process map 15_HIGH APS tutors_event categories_event sliders at 10 percent and path sliders at 25 percent

Process map 16_LOW APS_event categories_event sliders at 10 percent and path sliders at 25 percent

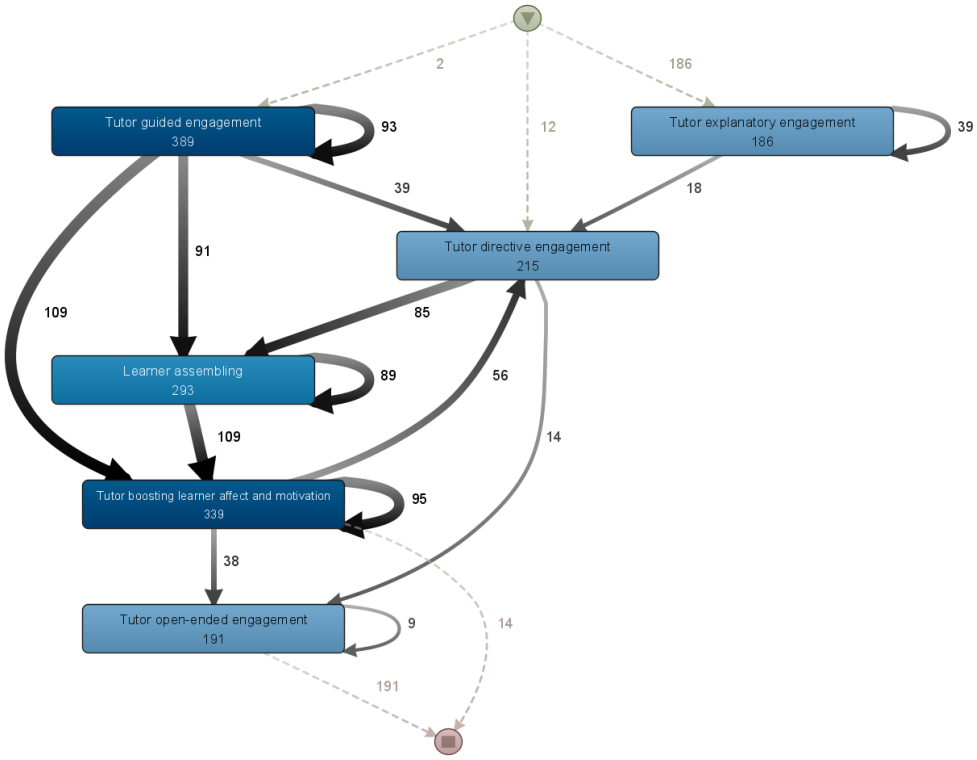


Figure 29. Process map 15 high APS tutors (event category view, event sliders at 10% and path sliders at 25%)

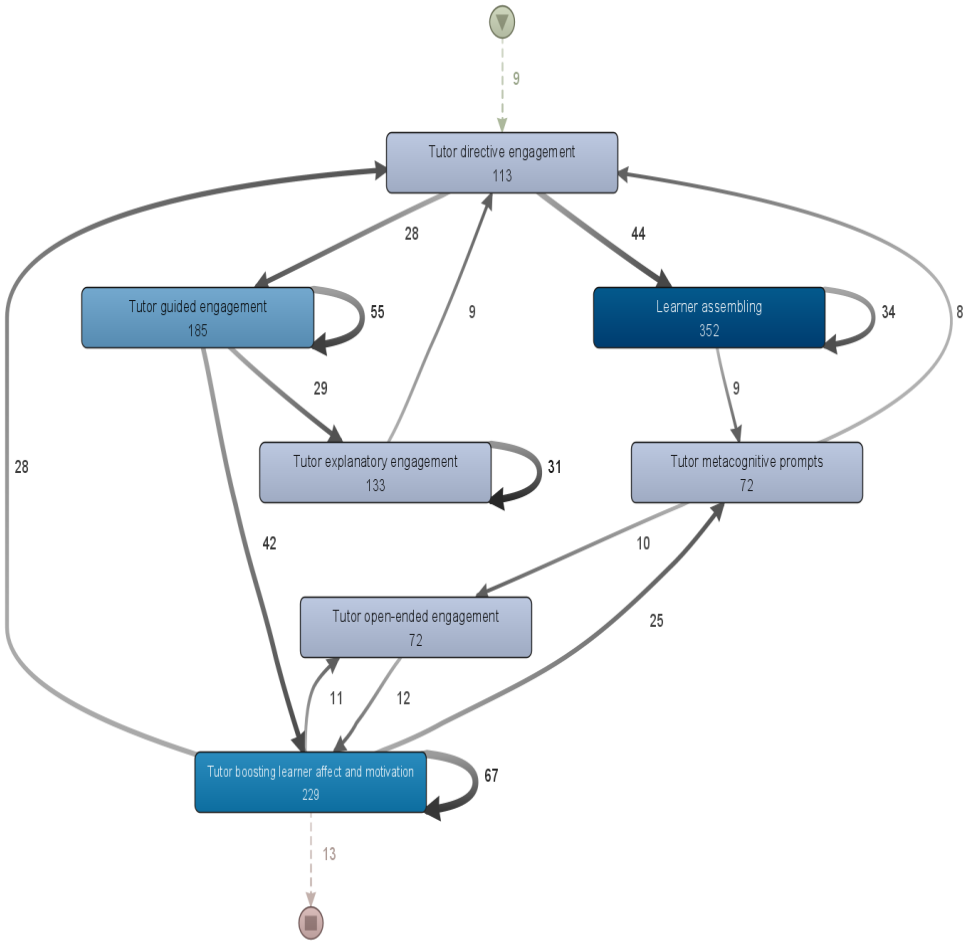


Figure 30. Process map 16 low APS tutors (event category view, event sliders at 10% and path sliders at 25%)

- For the **high APS** map, tutor guided engagement, tutor boosting learner affect and motivation, tutor open ended engagement, tutor directive engagement and tutor explanatory engagement are all significant, as illustrated by the dark colour of the event nodes.
- In contrast, for the **low APS** process map, tutor guided engagement and tutor boosting learner affect and motivation are disproportionately relied upon versus other types of tutoring practices, as illustrated by the difference in the colour of the event nodes.

The above differences between the process maps may reflect the difference in learner engagement in the sessions analysed, with low engagement or domain knowledge prompting tutors in the low APS group to relying on micro strategies such as scaffolding and praise. However, it should be noted that I was not able to observe any evidence that shows learners in the low APS sample demonstrating lower engagement or attainment, prior to commencing the programme The learners in both the high and low APS samples demonstrated similar levels of attainment in the pre-programme diagnostic scores. The alternative hypothesis is that learners show **greater progress in learning outcomes**, when tutors frequently rely on a **wider range of tutoring practices**.

In addition to analysing the differences in the co-regulatory processes by APS scores, I also analysed the extent to which co-regulatory processes were influenced by the criterion for tutor performance chosen. On comparing the process maps generated by each criterion, it appears that process maps generated using **tutor rankings** show a **clearer co-regulatory process**, with open ended tutoring practices featuring more prominently in the high-ranked tutor maps, and directive engagement being more limited than for low-ranked tutors. While tutors with high APS scores also appear to rely on open ended engagement practices more heavily than tutors with low APS scores, the distinction is less clear. This may be as tutors with a **high APS** scores **prioritise learning outcomes** over student centred teaching practices, and therefore rely on more heavily on tutor centred practices when considered necessary.

(ii) Signifiers process maps for data sample F (using average learner progress scores as the criterion for tutor performance).

As shown in Process map 14, the process map for tutors with a low APS only includes one event – 'tutor boosting learner affect and motivation'. In order to enable further analysis of the key events and transitions, the level of detail in the process maps for data sample F

(using event categories) was increased, with event sliders being set to 10% and the paths sliders being set to 25%. Similarly, the signifiers process maps have been produced and are set out in the Appendix F, with the event sliders being set to 10% and the paths sliders being set to 25%.

Process map 17_HIGH APS_signifiers map_event sliders at 10% and path sliders at 25 percent

Process map 18_LOW APS_signifiers map_event sliders at 10 percent and path sliders at 25 percent

Analysis of these process maps reveals that high APS tutors appear to be more proactive than low APS tutors. Further, while both maps are dominated by tutor-controlled practices, open ended tutoring practices appear to be more prominent within the high APS map rather than the low APS map. More specifically,

- As demonstrated by the process maps, the **high APS process maps** appear to **be driven by tutoring practices**, which occupy a more prominent position as the starting point of transitions within the map, than learner operations. Tutor controlled practices, such as 'tutor instructs learner', and 'tutor asks specific questions' are particularly prominent, both in terms of event significance and transitions. However, a form of open-ended engagement, 'tutor asks learner to explain their working out or asks a conceptual question' is also included within the map.
- On the other hand, the **low APS maps** appear to be **driven by learner behaviour**, with the most prominent node being 'learner writes down answers to questions. Similarly to the high APS map, the most important tutoring practices are forms of tutor-controlled practice. However, in contrast to the high APS map, **open ended tutoring practices are not sufficiently important to be included** in the overall process.

(iii)Process maps from tutor versus learner perspective

In the next stage of analysis, I filtered the process maps by tutor actions and learner actions. The resulting process maps are set out in the following figures with online links provided below.

Process map 19_HIGH APS_tutor filter_event slider 10 percent paths slider 25 percent

Process map 20_HIGH APS_learner filter_event slider 10 percent paths slider 25 percent

Process map 21_LOW APS_tutor filter_event slider 10 percent paths slider 25 percent

Process map 22_LOW APS_learner filter_event slider 10 percent paths slider 25 percent

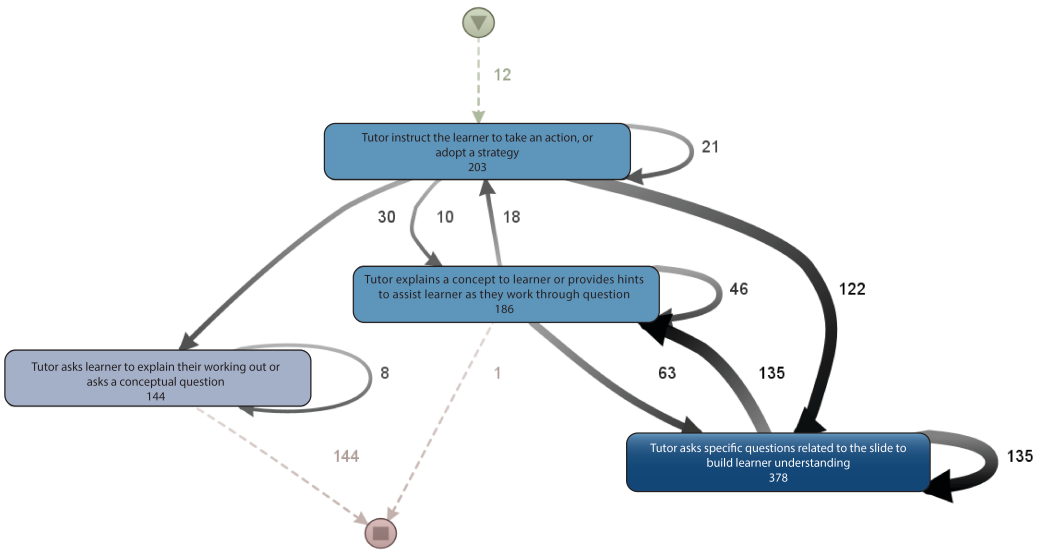


Figure 31. Process map 19 high APS tutors (filtered by tutor behaviours)

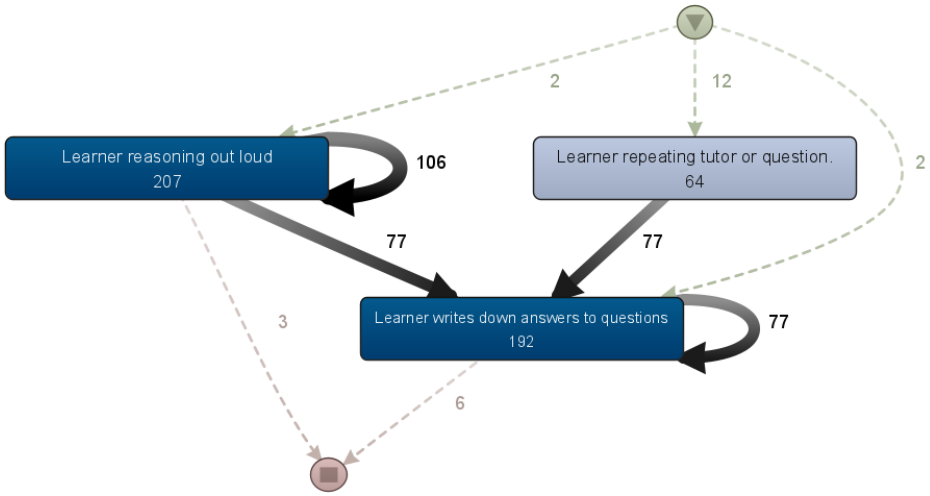


Figure 32. Process map 20, high APS tutors (filtered by learner behaviours)

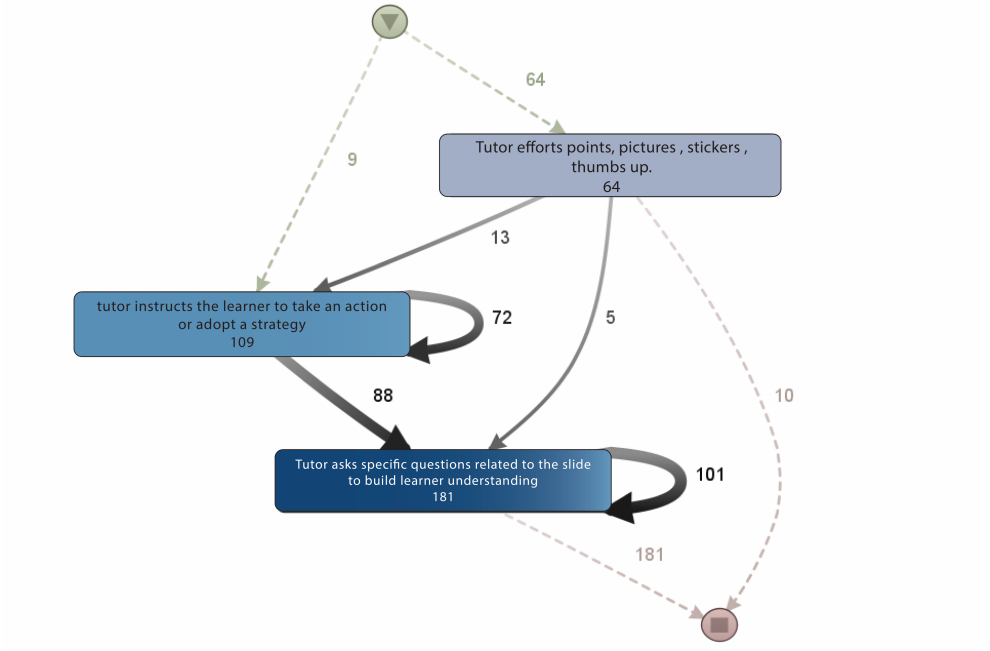


Figure 33. Process map 21, low APS tutors (filtered by tutor behaviours)

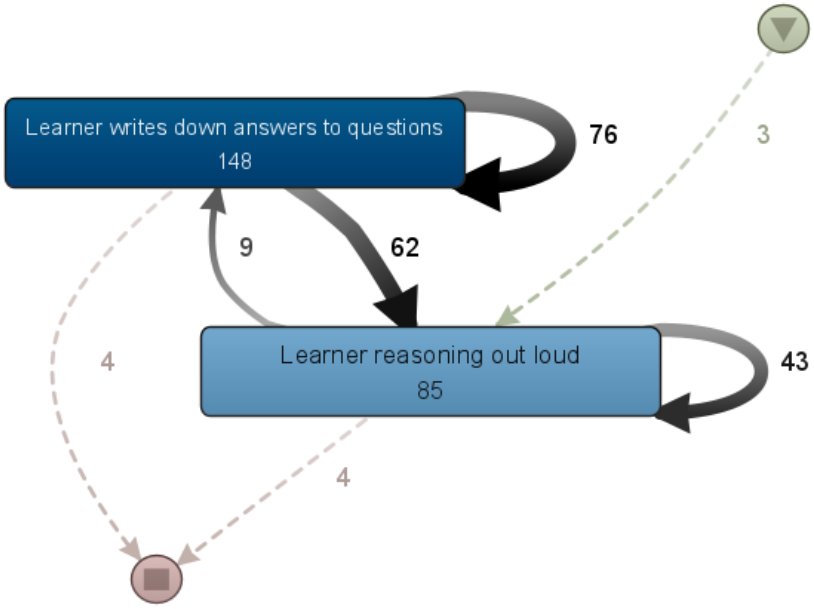


Figure 34. Process map 22 low APS tutors (filtered by learner behaviours)

The key differences between the high APS and low APS maps are summarised as follows:

- The high APS maps demonstrate a broader variety of tutoring practices, than the low APS maps, including open ended tutoring practices. The low APS maps are focused on tutor-led practices, such as instruction and micro questions. Although open-ended engagement is less prominent in the high APS versus high-ranking tutor groups, the findings are broadly similar regardless of whether learning outcomes or human evaluations are used as performance criteria. It is interesting to note that a **varied set of tutoring practices**, including **open-ended engagement**, appear to be **related** with **better learning outcomes**, versus purely tutor-led practices.

- There is a **mixture of shallow and deeper learner operations** demonstrated in the **high APS maps**, (including both learner repetition and learner reasoning) while the low APS maps show a greater focus on learner reasoning. On an initial examination, this is surprising, and contrasts with the results from process mining using evaluator rankings. The process maps based on evaluator rankings illustrated that learner reasoning was more prominent in the 'HIGH' versus 'LOW' group. However, on further consideration, this difference in prominent learner behaviours may be driven by the performance criterion used for process mining. When human evaluator rankings for SCT are used, it is reasonable to expect that there is a focus on learner led behaviours, such as learner reasoning. On the other hand, when learning outcomes are used, learner behaviours are likely to be more varied. It is possible that tutors in the high APS group may challenge learners to a greater extent, to attain progress in learning outcomes. As learners are scaffolded to engage with challenging content, they may demonstrate a mixture of shallower and more sophisticated learning operations.

5.1.7 Comparison of analysis from process mining, using different performance measures

The results from process mining for high-performing versus low-performing tutors with reference to learner progress scores, demonstrate distinct latent processes underlying these sessions. However, the **nature and extent of the differences** vary with the performance criteria used.

When learning outcomes are used as the performance criteria, the main differences relate to the range of tutoring practices relied upon. For example, **high-performing tutors demonstrate a distributed range of practices**, including tutor guided engagement, tutor boosting learner affect and motivation, tutor open ended engagement, tutor directive

engagement and tutor explanatory engagement. In contrast**, low-performing tutors heavily rely on a restricted range of practices**, including tutor guided engagement and tutor boosting learner affect and motivation.

While open-ended engagement features more prominently in sessions conducted by high-performing tutors, **the distinction between high and low-performing sessions is less than the distinction discovered when tutor rankings are used**. Similarly, the learning behaviours encompass both shallow and more sophisticated learning operations, in sessions conducted by high-performing tutors. It is likely that tutors in these sessions are more willing to adopt tutor-controlled practices as they scaffold learners through challenging learning activities. Learning outcomes are prioritised, and tutors rely on more heavily on tutor centred practices when considered necessary.

Following from the above, there appear to be **differences in how the VCE platform is used,** with high-performing tutors being more likely to use the platform as a foundation to build open ended forms of engagement. However, the distinction between groups of tutors is clearer, when tutor rankings are used as the criteria for performance, rather than learner progress scores.

In the following sections, in order to validate the findings generated from process mining, I will firstly consider the insights derived from process mining in light of the output from tutor interviews. I will then move on to discuss my findings and evaluate the use of process mining as a form of analytics, including its value and limitations.

5.1 Validation of process mining insights with tutor interviews

A set of qualitative interviews were conducted with 30 tutors, including 15 tutors who were classified as high ranking for SCT, and 15 tutors who were classified as low ranking for SCT. The details of how tutors were sampled for tutor interviews are set out in Section 3.2 of

this thesis. The interviews were conducted over Zoom over a one-month period, on a one-to-one basis, and were semi-structured in format.

Both the high-ranked and the low-ranked tutors are typically from non-teaching backgrounds. The low-ranked tutors have fewer years of tutoring experience compared to high-ranked tutors, with the average years of experience for low-ranked tutors being 0.5, and the average years of experience for high-ranked tutors being 1.5.

The tutors were interviewed about their perceptions of SCT and co-regulation, and the perceived importance of these practices. They were also asked about the tutoring practices they would use to co-regulate learners, and how these would be adapted to the needs of the learner. Finally, they were questioned on the role of the VCE platform as they co-regulated learners, and challenges experienced with co-regulation. The themes which emerged from the interviews are summarised below:

- Both high-ranked and low-ranked tutors recognise SCT as being very important. For example, low-ranked tutors recognised that SCT gave students '*good skills for life and thinking time to learners'*. Similarly, high-ranked tutors recognised SCT as a fundamental feature of their approach towards teaching, and as being key to '*improving student engagement, particularly for younger learners'*.
- High-ranked tutors had a greater degree of awareness about the concept of co-regulation, and its similarities and differences with SCT. All of the high-ranked tutors were familiar with the term 'self-regulation', while only 3 of the low-ranked tutors were familiar with this term. A high-ranked tutor explained that the difference between co-regulation and SCT as being that '*co-regulation teaches the learner how to take responsibility for their learning…SCT is about personalising the learning and getting the learner actively involved'*.

- High-ranked tutors were able to discuss a large range of tactics to co-regulate learners, while low-ranked tutors tended to focus on the student voice, and boosting learner affect. Examples of the tactics that high-ranked tutors cited as being used to co-regulate learners include different types of questions (ranging from micro questions to more open-ended questions), use of visual and learning aids, various types of reward and praise. In contrast, only 2 of the low-ranked tutors provided a detailed explanation of the tactics used to co-regulate learners, with the majority of tutors simply referring to encouraging the learners to actively participate within the session, and to encourage them when mistakes were made.
- One of the high-ranked tutors recognised that his practice has evolved as he gained more years of tutoring experience: *"when tutors are not that experienced, they have their own set of questions that they tend to deliver. As they get more experienced, they become more comfortable using a broader range of tactics. This is possibly why they become more comfortable with using open ended questions. When they are more nervous, they use a more scripted conversation".*
- High-ranked tutors were also more comfortable with recognising that co-regulatory practices need to be adapted to the specific needs of the students. Most (75%) of the high-ranked tutors interviewed acknowledged that they would shift to a more instructional tutor-led approach in the scenario that the learner was struggling with the activity. In contrast, only half of the low-ranked tutors agreed that there was need to adapt how they co-regulated learners, with one tutor stating that '*SCT and co-regulation work well for all types of learners'.*
- There were interesting differences in the perceived role of the VCE platform. 75% of high-ranked tutors agreed with the statement that 'tutors were responsible themselves for the structure of their tutoring sessions', but only 40% of low-ranked tutors did.

The majority of low-ranked tutors perceived the structure and flow of their sessions as being determined by the VCE platform.

- With regard to challenges in co-regulating learners, most tutors agreed that it was a challenge to shift towards student centred teaching when they were used to a teacher centred style. In addition, issues were highlighted with regards to accurately diagnosing learner engagement and affect online, which in turn made effective co-regulation challenging.

The above findings from tutor interviews **support the findings from process mining**. For example, high-ranked tutors are able to better articulate the concept of co-regulation, versus low-ranked tutors. In addition, they are able to discuss a wide range of tactics which they use to co-regulate learners, while the discussion with low-ranked tutors tends to be more limited. This aligns with the findings from process mining, which illustrate high-ranked tutors as deploying a wider range of tutoring practices, including open ended engagement, while low-ranked tutors tend to overly rely on a more restricted range of practices. It is interesting to note that low-ranked tutors do recognise the importance of SCT and encouraging learners to take an active role within sessions; however, the results from process mining illustrate that low-ranked tutors in fact tend to adopt tutor-controlled practices, and micro practices rather than more open-ended forms of engagement.

It is also interesting to note from the interviews that **high-ranked tutors** tended to be **more comfortable with adapting their co-regulatory practices**, in the scenario that learners were struggling, to a more tutor-led style of engagement. This aligns with the strong transitions demonstrated in the process maps for high-ranked tutors between guided engagement and open- ended engagement. High-ranked tutors appear to be comfortable with scaffolding learners through micro questions (guided engagement) where necessary, before moving onto open ended forms of engagement when appropriate. In contrast, half of low-

ranked tutors did not acknowledge the need to adapt their approach to the ability of the learner. This may explain the weak transitions between diverse types of tutoring practices for the low-ranked tutors process map, and the absence of open-ended forms of engagement from the map.

Finally, the tutor interviews also **validate the insights from process mining** regarding **the use of the VCE platform**. The process maps indicated that tutor-learner interactions for low-ranked tutors were closely structured around the format of the VCE activities, while tutor-learner interactions for high-ranked tutors were more likely to extend learning. Tutor interviews support this finding, with low-ranked tutors being more likely to perceive their sessions as being led by the structure of the platform.

Summary

In summary, the use of process mining has enabled the discovery of distinct co-regulatory processes for high-performing versus low-performing tutors. These processes vary, both in terms of structure and complexity, and with regard to the role of open-ended forms of engagement within the overall process. In addition, the extent of the differences between co-regulatory processes for high-performing versus low-performing tutors varied with the performance criteria used. A clearer co-regulatory process was demonstrated when tutor rankings for SCT were used versus the use of learning outcomes as performance criteria. While open ended engagement was more prominent in sessions conducted by tutors, who attained higher average learner progress scores, the overall difference in processes was less stark.

The results from process mining have been supported and extended through findings from tutor interviews. Both high and low-ranked tutors perceive co-regulation to be very important, but high-ranked tutors are better able to articulate and enact the concept of co-regulation, and the tactics used to co-regulate effectively. The interviews show that high-

ranked tutors tend to have more experience than low-ranked tutors. This could explain why high-ranked tutors are better able to implement complex co-regulatory processes, versus low-ranked tutors who overly rely on scripted questions. It is also interesting to note that low-ranked tutors perceive themselves as promoting the 'learner voice', while the process maps illustrate that low-ranked tutors overly rely on tutor-controlled practices, such as directive engagement. This points to a lack of self-awareness of teaching practice amongst low-ranked tutors. There could be an opportunity for granular process maps to be used to inform tutor training and assessment rubrics. This is further discussed in Chapter 7 of this thesis.

The results from process mining are limited by the scale and scope of the study. The results could be validated and extended by increasing the number of sessions analysed, as well as considering a broader range of tutor performance. In addition, ESL learners, and learners with special needs, could be included in future analysis. The results from this work are applicable for the study of maths problems in a particular VCE; it is possible that the study of other subjects (for example, English) would be associated with entirely different co-regulatory processes. In addition, the nature and design of the specific VCE used in this research study may impact the types of co-regulatory processes which were highlighted. For example, the design of the VCE did not incorporate any video of tutors or learners. It is possible that video would have resulted in different forms of co-regulation becoming more prominent e.g. the video cues may have resulted in a greater significance of tutor practices which prompted learner monitoring. Another factor which may have influenced my findings is the fact that learners may not retain the same tutor over the multiple VCE sessions. In some cases tutors may therefore be becoming more familiar with the learners strengths and weaknesses. This factor could have influenced the significance of various co-regulatory practices which encouraged different levels of cognitive processing.

The process maps produced are descriptive and cannot be statistically validated. However, the descriptive processes are very useful to explore how theory can unfold in an authentic learning context, and to inform discussions around tutor training and assessment. As a next step, it is also useful to identify statistically meaningful patterns, which occur within and across sessions. These patterns could potentially offer opportunities to support tutor training and assessment in real time or near real time (further discussed in Chapter 7 of this thesis). During the next stage of analysis, I therefore implemented sequential pattern mining and decision tree analysis of the tutor differences for my datasets (Sections 5.2 and 5.3).

5.2 Pattern mining

Introduction

In this Section, I analyse my dataset using pattern mining. The previous section showed advantages of process mining in providing holistic, temporal visualisations of SRL, which drive rich discussions about how SRL unfolds in the context of that learning environment. However, I also noted the limits of process mining, for example, the inability to statistically validate the output of a fuzzy miner model. The use of process mining can be extended by combining it with **pattern mining**, which can **detect isolated patterns of behaviour** that can be statistically validated. The output from pattern mining can **validate** the results from **process mining**. In addition, pattern mining can provide an **additional layer of granularity** by identifying significant patterns which occur across sessions, as well as within them.

In the following, I set out my approach towards implementing pattern mining, and analyse the results obtained. This Section follows the following structure:

- I start off by reviewing the range of pattern mining algorithms, and their relevance for my study context. This includes evaluation of which algorithm or combination of algorithms will be best suited to address my research questions.
- I implement the relevant algorithms for my datasets and highlight significant sequences. Algorithms were implemented for both dataset E (where tutor performance is determined using evaluator rankings), and dataset F (where tutor performance is determined based on standard measures of learner progress).
- Finally, I analyse the results obtained, and explore the differences between dataset E and dataset F.

5.2.1 Identifying relevant pattern mining algorithms for my study context

In order to select a pattern mining algorithm for my research, I reviewed types of pattern mining and **evaluated their relevance** for my research. Different types of pattern mining align with varying research foci. For example, it is important to consider whether my research questions require analysis of how frequently an isolated range of events occur in a particular order, or if the scope of the analysis should be broadened to also understand the probability of moving from one event to another. Similarly, considerations of whether the research should focus on the occurrence of patterns across sessions, or within sessions, or both, is also necessary.

I have analysed in Table 37 the different types of pattern mining approaches by the nature of the output produced. As illustrated, **sequential forms of pattern mining are applicable** when researching the order in which events occur, while item-set and association rule mining are relevant for unordered transaction databases. Within the category of sequential pattern mining, there exist a variety of different types of analysis, ranging from analytics which simply identify how regularly certain patterns occur, through to analytics

which operate under various constraints (e.g., time periods), and evaluate the conditional probability of patterns occurring.

Table 36. Types of pattern mining.

Type of pattern mining	**Order-based**	**Identifies the support of patterns across cases**	**Identifies the support of patterns within cases**	**Measures confidence of patterns**	**Description**
Sequential pattern mining	X	x			Discovers the frequency of item-sets across cases (i.e., sequences)
Episode mining	X		X		Discovers the frequency of item-sets within a case (i.e., sequence), or the rules of the pattern
Periodic pattern mining	X		X		Discovers patterns which occur frequently and periodically within a case (i.e., sequence). The user can specify constraints on period length, such as minimum and maximum average period length, and a maximum period length.
Sequential rule-based mining	X	X		X	Discovers sequential rules with a given confidence.
Item-set mining		X			Discovers frequent item-sets in an unordered transaction database.
Association rule mining		X		X	Discovers the rules of association in an unordered transaction database, with a given confidence.

My research focuses on understanding how tutors co-regulate learners, and how this varies with tutor performance. The order of tutor-learner interactions is therefore very

important, and forms of sequential pattern mining are therefore relevant for my research. I already implemented process mining to discover holistic processes, and prominent events/transitions. However, the results from process mining cannot be statistically validated, and also do not differentiate between the patterns that are occurring across sessions, versus within a session. I therefore now **aimed to identify statistically significant patterns that exist across sessions, and within sessions**. To supplement the analysis from process mining, I chose to use **sequential pattern mining** and **episode mining**, to analyse the support of patterns which occur within and across sequences. Process mining has already provided a description of the most prominent transitions, and mining the rules of association was therefore not prioritised. As I am focused on discovering differences in the order of tutor-learner interaction, rather than exploring the temporal differences that may occur, periodic pattern mining was not implemented. Rather, I chose to evaluate the patterns through analysing the statistical significance of the difference in the frequency of the pattern, for sessions conducted by high-performing versus low-performing tutors. My research focuses on discovering the differences in the co-regulatory processes adopted by high vs low-performing tutors, and this objective aligns with analysing the significance of the differences in frequent patterns of interaction.

There is a wide range of algorithms that can be used for sequential pattern mining and episode mining. A full description of the algorithms available, and how they work, is provided in the Appendix E. While the output for each of these algorithms will be identical, the process and algorithm performance will vary. For the purposes of my research, I selected **CM-SPAM for sequential pattern mining**, and **EMMA for episode mining**. Both algorithms are available from the open source SPMF platform. To mine patterns efficiently, I followed the below approach:

- I started with implementing CM-SPAM to identify frequent sequences, across sessions conducted by high and low-ranked tutors (the 's-support').
- I compared the difference in the s-support, for sessions conducted by high-ranked versus low-ranked tutors, and highlighted statistically significant patterns.
- I then implemented EMMA for the significant patterns to identity how frequently these patterns occurred within each session ('the i-support').
- Finally, the mean i-support was calculated for each pattern for sessions conducted by high-ranked versus low-ranked tutors.

5.2.2 Pattern mining using data sample E (using evaluator rankings as the criterion for tutor performance)

Implementation of CM- SPAM

I used the dataset collated from sessions conducted by high-ranked and low-ranked tutors for pattern mining. Please see Appendix I for a sample of the dataset collected. The data-set was coded into a format which can be used by CM-SPAM.

The input for CM-SPAM is a sequence database, in which each session corresponds to a case in the file, and the events in the session are chronologically ordered. A threshold named minsup (a value in [0,1] representing a percentage) is specified prior to implementation which dictates the frequency threshold a pattern needs to exceed to be highlighted. Moreover, the implementation in SPMF adds another parameter, which is the maximum sequential pattern length in terms of items.

I considered the level of granularity of the events that would be mined, and chose to mine events at the signifiers level, rather than the event category level. I **mined events at the signifiers level** to discover statistically significant patterns, which were fine-grained and represented the practical tutor-learner interactions in the VCE. The discovery of these patterns could have **potential operational implications**, such as real time or near real time

interventions to support tutor-learner interactions. These operational implications do not form part of the scope of this thesis, but could be the subject of future research (see Chapter 7).

With regard to the configuration of parameters for CM-Spam, the minimum pattern length was set as two items, and maximum pattern length was set to the maximum possible number of items. I did not specify any required items for sequential mining, and set the maximum gap between items to one. A 'minsup 'value of 0.5 was applied, which meant that a sequential pattern would be considered as frequent if it appeared in over half of the total number of sequences.

Results from implementation of CM-SPAM

The patterns generated are presented in Table 1, along with their support in each group:

*High-ranked tutors – S-frequent patterns**

Table 37. S frequent patterns for high-ranked tutors *Note: there are 29 sessions in total conducted by high-ranked tutors.

S-frequent patterns	**S-frequency**
Learner writing answers–tutor passive–learner writing answers	26
Learner reasons–tutor passive–learner reasons	25
Tutor asks open ended conceptual questions–learner reasons–tutor passive-learner reasons	21
Tutor asks open ended conceptual questions–learner reasons–tutor passive	21
Tutor praise–tutor instruction–learner repetition	18
Tutor passive–learner writes answers–tutor passive–learner writes answers	18
Tutor passive–learner writing answers–tutor praise	16
Tutor passive–learner reasons–tutor praise	16
Tutor passive–learner reasons–tutor passive	16
Learner writes answer–tutor praises learner–tutor asks learner a specific question	16
Learner writes answer–tutor passive–learner writes answer–tutor passive	16
Learner reasons–tutor passive–learner reasons–tutor praise	16
Tutor praises learner–tutor rewards learner–tutor instructs learner	15
Tutor praises learner–tutor instructs learner–learner writes answer	15
Learner writes answers–tutor passive–learner writes answers–tutor praise	15

Tutor asks learner a specific question–learner writes answer–tutor praise	15

Low-ranked tutors – S-frequent patterns*

Table 38. S frequent patterns for low-ranked tutors *Note: there are 30 sessions in total conducted by low-ranked tutors.

S-frequent patterns	**S-frequency**
Learner writing answers–tutor passive–learner writing answers	18
Learner reasons–tutor passive–learner reasons	16
Tutor asks specific question–tutor asks specific question–tutor asks specific question	16

Statistically significant patterns

I then conducted a Pearson chi-square test to analyse whether there was a relationship between the tutor ranking and the frequency of patterns observed. The null hypothesis was that there was no relationship between tutor ranking and the frequency of the patterns observed, and the confidence interval was set to 95% ($p < .05$). Table 40 presents the results of this analysis. The null hypothesis was rejected for the highlighted patterns, for which a probability smaller than 0.05 was calculated for the Chi square value. The frequency of the highlighted patterns appears to be related to whether sessions are conducted by high-ranked or low-ranked tutors.

Table 39. Significant s- frequent patterns for dataset E (* statistically significant $p<0.05$).

S-frequent patterns for high or low-ranked tutor sessions	**Frequency in sessions conducted by high-ranked tutors**	**Frequency in sessions conducted by low-ranked tutors**	**P(x)**
Learner reasons–tutor passive–learner reasons	25	16	0.02*
Tutor asks learner a specific question–learner writes answer–tutor praise	15	6	0.02*
Learner writing answers–tutor passive–learner writing answers	26	18	0.03*

Learner writes answer–tutor praises learner–tutor asks learner a specific question	16	8	0.04*
Tutor asks open ended, conceptual questions–learner reasons–tutor passive–learner reasons	21	12	0.04*
Learner writes answer–tutor passive–learner writes answer–tutor passive	16	14	0.70
Tutor passive–learner reasons–tutor praise	16	9	0.08
Tutor passive–learner reasons-tutor passive	16	9	0.08
Learner reasons–tutor passive–learner reasons–tutor praise	16	9	0.08
Tutor asks open ended conceptual questions–learner reasons–tutor passive	21	14	0.09
Tutor praise–tutor instruction–learner repetition	18	12	0.15
Tutor praises learner–tutor instructs learner–learner writes answer	15	10	0.23
Tutor passive–learner writing answers–tutor praise	16	11	0.24
Tutor passive–learner writes answers–tutor passive–learner writes answers	18	13	0.24
Learner writes answers–tutor passive–learner writes answers–tutor praise	15	10	0.24
Tutor asks specific question–tutor asks specific question–tutor asks specific question	13	16	0.36
Tutor praises learner–tutor rewards learner–tutor instructs learner	15	12	0.51

Analysis of results from implementation of CM-SPAM

Sessions conducted by high-ranked tutors demonstrate 16 s-frequent patterns, while sessions conducted by low-ranked tutors only demonstrate 3 s-frequent patterns. This implies that there are **a range of similarities** in the manner in which **high-ranked tutors interact**

with students, while the commonalities in the interaction generated by **low-ranked tutors** are **more limited.**

Reviewing the s-frequent patterns for high-ranked tutors highlights 3 main types of patterns of engagement:

- *Patterns in which the tutor encourages the learner to constructively engage in the learning, through passive observation of the learner while they work.* In these patterns, the learner is engaged in active or constructive modes of learning (e.g., writing an answer or reasoning), and the tutor is passively observing the learner as they work (e.g., through using filler words such as 'ok', or 'mm'). There are a range of patterns of varying length and frequency, which fall into this category.
- *Patterns in which the tutors attempt to scaffold or diagnose learner understanding, through asking specific questions prior to the learner solving a problem, or after the learner has attempted a problem.* Tutors use praise/reward as a tool to boost learner morale, as part of this interaction.
- *Patterns in which the tutor encourages constructive engagement, through using open-ended questions to elicit learner reasoning, and passively observing the learner.*
- *Patterns in which the tutor leads the learner through their activities in the VCE, through issuing instructions on how to approach the problems, combined with praise or reward.*

In contrast, the patterns of engagement demonstrated by low-ranked tutors can be categorised as follows:

- *Patterns in which the tutor encourages the learner to constructively engage in the learning, through passive observation of the learner while they work, or reason*. In

contrast to high-ranked tutors, only relatively short patterns of this type are frequent for low-ranked tutors.

- *Patterns in which the tutor scaffolds the learner through a string of specific questions*. This use of specific questions contrasts with high-ranked tutors, who are more likely to use specific questions as a precursor or follow up to the learner solving a problem in the VCE.

The **patterns exhibited by high-ranked tutors are diverse**, and tend to **encourage active participation by learners**, while **low-ranked tutors** demonstrate patterns in which the tutor appears to retain a **greater degree of control**. For example, high-ranked tutors frequently use open ended questioning to elicit learner reasoning, while these patterns are not frequent for low-ranked tutors.

Applying the chi-square test, with the significance level set to 5%, demonstrates that certain patterns are much more likely to appear in the high-ranked tutor sessions, versus the low-ranked tutor sessions. For example, **patterns in which the tutor passively observes the learner while they problem-solve or reason**, are frequent for both high and low-ranked tutors. However, the p-value demonstrates that these patterns are **much more likely to occur for high-ranked tutors**. Similarly, **high-ranked tutors** are much more likely to **use questioning in diverse ways**; open ended questions are used to elicit reasoning, and specific questions are used to scaffold learners or diagnose understanding pre or post problem solving. While sequences of specific questions are only s-frequent for low-ranked tutors, they also occur in the high-ranked group, and this is not a significant pattern.

As expected, the length of the pattern affects the results observed. For example, *'Tutor asks open-ended, conceptual questions–learner reasons–tutor passive'* is not statistically significant, but the longer pattern *'Tutor asks open ended conceptual questions–learner reasons–tutor passive–learner reasons'* is. This indicates that both high-ranked and

low-ranked tutors use open ended questions to elicit reasoning from learners, but **high-ranked tutors tend to elicit more extensive reasoning**. Further research is required to explore why high-ranked tutors elicit more extensive reasoning. For example, high-ranked tutors may be able to elicit a greater degree of reasoning due to the to appropriate pitching of the open-ended questions, both with regards to timing and framing of question, or due to appropriate pausing/monitoring to encourage further reasoning prior to moving onto the next activity, or due to a stronger subject matter expertise.

Implementation of EMMA algorithm

In the next stage of analysis, I further analysed the statistically significant patterns which occurred across sessions. I explored whether there were any differences in how regularly these patterns occurred **within** sessions, for high-ranked tutors versus low-ranked tutors.

In order to conduct this analysis, I implemented the EMMA algorithm for the high-ranked tutors and low-ranked tutors' dataset. EMMA is one of the fastest algorithms for frequent episode mining. It calculates the instance support of patterns using the concept of head frequency (i.e. the frequency or proportion of occurrences of a particular pattern at the start or "head" of a sequence).

The purpose of this analysis was to discover the instance support of patterns that had already been identified as statistically significant across sessions, to explore whether these patterns were also more frequent within sessions. I therefore implemented EMMA with a minimum support of 2, and explored the results to find the instance support for the relevant patterns.

Results from implementation of EMMA

The mean of the episode support for the identified patterns is presented in Table 41, along with the standard deviations. As illustrated, the **episode support for all the patterns**

skews higher for the high-ranked tutor group versus the low-ranked tutor group. However, the standard deviations are also very high which indicates that the results have limited value to be relied upon to establish statistical significance. A potential reason that the standard deviation may be high is due to differing learner needs, which resulted in tutors providing differing levels of scaffolding within sessions. In future research, I can potentially test this hypothesis by conducting the research and analysis for highly similar learner profiles and session profiles.

Analysis of results from implementation of EMMA

While the episode support cannot be relied upon for statistical significance, it is still interesting to note that the data for the high-ranked tutor group is skewed towards a higher mean of episode support for the frequent patterns. Sessions conducted by a group of high-ranked tutors appear to demonstrate a handful of patterns; whilst these same patterns may occur more often across the course of an individual session conducted by a high-ranked tutor (although there are high levels of variation between specific sessions so statistical comparison would require a greater amount of reliable data).

A review of the results also illustrates the points below:

- The use of the '*Learner writes answer – tutor praises learner – tutor asks learner a specific question'* pattern appears to distinguish high-ranked tutors from low-ranked tutors. The pattern has a low mean for low-ranked tutors, particularly when compared to high-ranked tutors. It also has a relatively low standard deviation. Considering both the s-support and the i-support of the pattern, it appears to be infrequently used by low-ranked tutors, and when it is used, the episode support is limited.
- The *'Learner reasons – tutor passive – learner reasons'* pattern has a higher mean episode support for high-ranked tutors, but also has a very high standard deviation.

Considering the s-frequency and i-frequency together, this pattern seems to be implemented across almost all high-ranked tutor sessions, but individual tutors vary with regards to how often they use this pattern.

Table 40. Mean and standard deviation of the episode support of patterns in high and low-ranking tutor groups.

Pattern	**High-ranked tutor group**		**Low raked tutor group**	
	Mean	**Stdev**	**Mean**	**Stdev**
Learner reasons–tutor passive–learner reasons	1.36	1.51	0.60	1.07
Tutor asks learner a specific question–learner writes answer–tutor praise	0.48	0.38	0.19	0.29
Learner writes answers–tutor passive–learner writes answers	2.23	1.50	0.65	1.65
Learner writes answer–tutor praises learner–tutor asks learner a specific question	0.40	0.41	0.09	0.17
Tutor asks open ended conceptual questions–learner reasons–tutor passive–learner reasons	0.50	0.46	0.16	0.30

5.2.3 Pattern mining using data sample F (using average progress scores as the criterion for tutor performance)

Implementation of CM-SPAM

I then implemented pattern mining for the sample of data collected using average progress scores as a criterion for performance (dataset F).

To start with, I implemented CM-SPAM, using the same parameters as those applied for dataset E. The minimum pattern length was set as two items, and maximum pattern length was set to the maximum possible number of items. I did not specify any required items for sequential mining, and set the maximum gap between items to one. A ‘minsup’ value of 0.5 was applied, which meant that a sequential pattern would be considered as frequent if it appeared in over half of the total number of sequences.

Results from implementation of CM-SPAM

The patterns generated are presented in Table 42 and 43, along with their support in each group:

High-performing tutors – S-frequent patterns

Table 41. S frequent patterns for high APS tutors *Note: there are 20 sessions in total conducted by high APS tutors.

S-frequent patterns	S-frequency
Tutor asks open ended, conceptual questions–learner reasons–tutor passive	14
Learner reasons–tutor passive–learner reasons	14
Tutor asks open ended conceptual questions–learner reasons–tutor passive–learner reasons	14
Tutor asks open ended conceptual questions–learner reasons–tutor praise	11
Tutor asks learner a specific question–tutor asks open ended conceptual questions–learner reasons	11
Learner writes answers–tutor asks open ended conceptual questions–learner reasons	11

Low-performing tutors – S-frequent patterns*

Table 42. S frequent patterns for high APS tutors *Note: there are 21 sessions in total conducted by high APS tutors.

S-frequent patterns	S-frequency
Learner reasons–tutor passive–learner reasons	11

Statistically significant patterns

I then conducted a Pearson chi-square test of independence to analyse whether there was a relationship between the tutor performance and frequency of patterns observed. The null hypothesis was that there was no relationship between tutor performance and the frequency of the patterns observed, and the significance level was set to 5%. Table 44 presents the results of this analysis. The null hypothesis was rejected for the highlighted patterns, for which a probability of under 5% was calculated for the Chi square value. The **frequency of the three highlighted patterns** appears to be **related to whether sessions are conducted by high-performing or low-performing tutors**, where performance is judged by the learning outcomes.

Table 43. Significant s-frequent patterns for dataset F.

S-frequent patterns for high or low-ranked tutor sessions	**Frequency in sessions conducted by high-performing tutors**	**Frequency in sessions conducted by low-performing tutors**	**P(x)**
Tutor asks learner a specific question–tutor asks open ended conceptual questions–learner reasons	11	3	0.01
Tutor asks open ended conceptual questions–learner reasons–tutor praise	11	4	0.03
Tutor asks open ended conceptual questions–learner reasons–tutor passive–learner reasons	14	7	0.04
Tutor asks open ended conceptual questions–learner reasons–tutor passive	14	9	0.16
Learner writes answers–tutor asks open ended conceptual questions–learner reasons	11	8	0.42
Learner reasons–tutor passive–learner reasons	14	11	0.44

Analysis of results from implementation of CM-SPAM

There are some similarities between the results for the dataset F (using Average Progress Scores as a criterion for performance), and dataset E (using evaluator rankings as a criterion for performance). For example, **high APS tutors demonstrate shared patterns of behaviour**, with 6 patterns being s-frequent, versus only 1 s-frequent pattern being demonstrated by low APS tutors. In addition, **open-ended tutoring practices feature more heavily in sessions conducted by high APS tutors**, with most frequent patterns integrating open-ended questions from the tutor. In contrast, the **only regular pattern** in sessions conducted by **low APS tutors** is **learner reasoning**, while the tutor passively observes. While this pattern of interaction is important to encourage learner autonomy, it is not

complemented by other patterns aimed at encouraging reasoning (as per sessions conducted by low APS tutors).

Examining the significance of the difference in pattern frequency illustrates **that high APS tutors** are distinguished by the **diverse manner** in which they use **open ended questions**. Both high and low APS tutors demonstrate patterns of using open ended questioning to elicit reasoning (with these patterns being less frequent for low-performing tutors). However, high APS tutors are significantly more likely to employ diverse patterns, such as **building open-ended questioning on top of scaffolding through micro questions**. They are also more likely to demonstrate longer patterns of engagement to elicit reasoning, and to reward reasoning with praise.

It should be noted that the learning outcomes dataset was collected to experiment with the use of an alternative performance criterion, and is smaller than the rankings dataset (dataset E). It is possible that the results achieved with this small dataset would be further magnified with the use of a larger dataset.

Implementation of EMMA algorithm

The EMMA algorithm was implemented for dataset F, with a minimum support of 2, and I explored the results to find the instance support for the relevant patterns. The instance support generated for each pattern was normalised for session duration, and the mean and standard deviation was calculated. The mean of the episode support for the identified patterns is presented in Table 45, along with the standard deviation.

Analysis of results from implementation of EMMA algorithm

As illustrated, the **episode support for the patterns is skewed higher for the high APS tutor group** versus the low APS tutor group. However, the standard deviations are also high, and the mean episode support can therefore not be relied upon to establish statistical significance. While the episode support cannot be relied upon for statistical significance, it is

still interesting to note that the data for the high APS group is skewed towards a higher mean of episode support for the s- frequent patterns.

While the episode mean does skew higher for the high APS groups, there also seems to be **significant variation in the extent to which s-frequent patterns occur within sessions**. A review of the results illustrates the points below:

- The *Tutor asks open ended, conceptual questions–learner reasons–tutor passive–learner reasons* pattern demonstrates a high standard deviation, as well as comparable episode means for high vs low APS tutors.
- While the mean of the *Tutor asks open ended, conceptual questions–learner reasons–tutor praise* pattern is much higher in the high versus low APS group, there is also a particularly high standard deviation.
- The *Tutor asks learner a specific question–tutor asks open ended, conceptual questions–learner reasons* pattern is the only pattern with a considerably higher mean in the high APS group, and relatively low standard deviation.

Table 44. Mean and standard deviation of the episode support of patterns in high and low APS tutor groups.

Pattern	**High-performing tutor group**		**Low-performing tutor group**	
	Mean	**Stdev**	**Mean**	**Stdev**
Tutor asks open ended conceptual questions–learner reasons–tutor praise	0.78	1.62	0.28	0.56
Tutor asks learner a specific question–tutor asks open ended, conceptual questions–learner reasons	0.35	0.55	0.19	0.54
Tutor asks open ended conceptual questions–learner reasons–tutor passive–learner reasons	0.58	0.67	0.44	0.61

Summary

I have implemented CM-SPAM (**sequential pattern mining)** and EMMA (**episode mining)** to identify significant patterns of interaction demonstrated in sessions conducted by high-performing vs low-performing tutors. Tutor performance is measured by human evaluator rankings for student centred teaching (dataset E), and average progress scores (dataset F).

The results for dataset E show that that there are **a range of similarities** in the manner in which **high-ranked tutors interact with students**, while the commonalities in the interaction generated by **low-ranked tutors** are **more limited.** For example, the sessions conducted by high-ranked tutors demonstrate 16 s-frequent patterns, while sessions conducted by low-ranked tutors only demonstrate 3 f s-frequent patterns. This is also true for dataset F (where tutor performance is measured by average progress scores), albeit to a lesser extent. For example, **high APS tutors demonstrate shared patterns of behaviour**, with 6 patterns being s-frequent, versus only 1 s-frequent pattern being demonstrated by low APS tutors.

The results for dataset E also show that the **patterns exhibited by high-ranked tutors are diverse**, and tend to **encourage active participation by learners**, while **low-ranked tutors** demonstrate patterns in which the tutor appears to retain a **greater degree of control**. For example, high-ranked tutors frequently use open ended questioning to elicit learner reasoning, while these patterns are not frequent for low-ranked tutors. The results for dataset F are similar in this regard. **Open-ended tutoring practices feature more heavily in sessions conducted by high APS tutors**, with most frequent patterns integrating open-ended questions from the tutor.

In addition, high-ranked tutors are much more likely to **use questioning in diverse ways**; open ended questions are used to elicit reasoning, and specific questions are used to scaffold learners or diagnose understanding pre or post problem solving. Similarly, high APS

tutors are significantly more likely to employ diverse patterns, such as **building open-ended questioning on top of scaffolding through micro questions**. For both dataset E and F, high-performing tutors are also more likely to demonstrate longer patterns of engagement to elicit reasoning, and to reward reasoning with praise.

Patterns in which the tutor passively observes the learner while they problem-solve or reason, are frequent for both high and low-ranked tutors, in dataset E. However, the results from the chi-square test show that these patterns are **much more likely to occur for high-ranked tutors**. While sequences of specific questions are only s-frequent for low-ranked tutors, they also occur in the high-ranked group, but this is not a significant pattern.

With regard to episode mining for dataset E and F, the **episode support for all the patterns skews higher for the high-performing tutor group** versus the low-performing tutor group. However, the standard deviations are also very high which indicates that the results have limited value to be relied upon to establish statistical significance. There is only one pattern in each dataset with a higher mean and relatively low standard deviation:

- For dataset E, the use of the '*Learner writes answer – tutor praises learner – tutor asks learner a specific question'* pattern appears to distinguish high-ranked tutors from low-ranked tutors. The pattern has a low mean for low-ranked tutors, particularly when compared to high-ranked tutors. It also has a relatively low standard deviation. Considering both the s-support and the i-support of the pattern, it appears to be infrequently used by low-ranked tutors, and when it is used, the episode support is limited
- For dataset F, the *Tutor asks learner a specific question – tutor asks open ended, conceptual questions – learner reasons* pattern is the only pattern with a considerably higher mean in the high APS group, and relatively low standard deviation.

As described above, pattern mining is useful for validating and extending the insight from process mining. Moreover, the output can also be used in another way; the significant patterns detected can be used as attributes for classification algorithms to predict whether tutors belong to the high versus low-performing class. For example, Cukurova et al. (2022) built decision trees using the sequences of tutor behaviours as attributes, attaining classification accuracies of approximately 88%. Predictive analytics have many operational implications, such as real time or near real time prompts for tutors when maladaptive patterns of engagement are detected. While the operationalisation of predictive analytics is not the subject of this thesis, the development of predictive analytics using sequences as attributes, is a ripe area for research.

In Section 5.3, I will explore whether the statistically significant sequences from pattern mining can be used to predict whether sessions have been conducted by high-performing versus low-performing tutors.

5.3 Decision trees

Introduction

Following the identification of statistically significant sequences in Section 5.2, I explore the use of these sequences to build predictive analytics which classify tutors into high and low-performing groups. This is a ripe area for research, with potentially significant operational implications. Prediction of tutor performance based on tutor-learner interactions could potentially enable early interventions to support tutors, for example through explainable feedback.

My research aims to build on previous research which utilises sequences of tutor behaviours as the basis for decision trees (Cukurova et al., 2022).

The structure of this Section is as follows:

- I start off by selecting predictive analytics, and pre-processing my data sample.

- I implement the algorithms for my dataset, experimenting with using binned and un-binned data. Classification models are built for both dataset E (with tutor performance determined by human evaluator rankings), and dataset F (with tutor performance determined by average progress scores).
- I analyse the results obtained for each data-set, and explore the differences between the classification models for each dataset. The operational implications of this work are discussed in chapter 7 of this thesis, but are otherwise outside the scope of my research.

5.3.1 Selection of classifiers

I have utilised the open-source WEKA software for my research. WEKA (Waikato Environment for Knowledge Analysis) has numerous in-built machine learning algorithms, including decision tree and rule-based classifier algorithm. I selected the JRIP and J48 algorithms for implementation, as these algorithms have been shown to be able to effectively classify data using pattern frequency as attributes (Cukurova et al., 2022). JRIP generates a set of rules which can be used to classify tutors in high and low-performing groups, based on the attributes of the data-set, while J48 generates a decision tree.

I have initially implemented the algorithms using the default parameters, with data being normalised for session duration. I have then experimented with binning the data, and tuning the algorithm parameters to improve classifier performance.

5.3.2 Implementation of predictive analytics for dataset E (evaluator rankings used as criterion for performance)

The data input used for the classifiers is the i-support of statistically significant patterns (set out in Table 41). I have named and listed these below for convenience:

- Pattern A: Tutor asks learner a specific question–learner writes answer–tutor praise (named as *'Guided questioning and praise')*

- Pattern B: Learner writes answer–tutor praises learner–tutor asks learner a specific question (named as *'Follow up praise and guided questioning'*)
- Pattern C: Tutor asks open ended, conceptual questions–learner reasons–tutor passive–learner reasons (named as *'Open-ended prompting and observe reasoning'*).
- Pattern D: Learner reasons–tutor passive–learner reasons (named as *'Observe reasoning'*).
- Pattern E: Learner writes answers–tutor passive–learner writes answers (named as *'Observe problem-solving'*).

JRIP algorithm

Implementation using the algorithm default parameters, and normalised data-set

The JRIP algorithm was trained using the normalised data-set, and default parameters. For example, the default setting for the number of folds for pruning was 3 folds, and the default setting for the minimum total weight of the instances in a rule was 2 instances.

Initially, all 5 patterns identified as having a significant difference in s-support between high and low-performing tutors, were used as input for the decision tree. However, patterns D (*'Observe reasoning')* and E (*Observe problem-solving')* demonstrated a high standard deviation from the mean i-support for both the high and low-ranked tutor groups. The inclusion of patterns D and E had a negative effect on the performance of the model, and they were removed as attributes. The resulting model, using patterns A-C as attributes**,** **generates 3 rules** which are set out below.

JRIP rules:

==========

(Pattern C Open-ended prompting and observe reasoning >= 0.353482) => Tutor ranking=High (25.0/6.0)

(Pattern B Follow up praise and guided questioning >= 0.321958) => Tutor ranking=High (10.0/2.0)

=> Tutor ranking=Low (24.0/2.0)

The model is tested through a 10-fold cross validation, and a **classification accuracy of 76.3%** is achieved. The model performs well across both classes of data, with comparable rates of precision and recall. **Precision** in a decision tree is the proportion of true positive instances among the instances that the model labeled as positive, while **Recall** is the proportion of true positive instances among all actual positive instances. **Recall is slightly lower for low-ranked** tutors versus high-ranked tutors, indicating that a minority of low-ranked tutors demonstrate pattern B *Follow up praise and guided questioning* and pattern C *Open-ended prompting and observe reasoning* in a manner which is more typical of high-ranked tutors. The **minority of low-ranked tutors** who are misclassified are **using practices to diagnose learner understanding, and elicit reasoning**, more frequently than expected.

Table 45. Stratified cross validation (JRIP, dataset E).

Correctly classified instances	45 (76.3%)
Incorrectly classified instances	14 (23.7%)
Kappa statistic	*0.53*

Table 46. detailed accuracy by class (JRIP, dataset E).

Class	TP rate	FP rate	Precision	Recall	F-measure	MCC	ROC area	PRC area
High	0.793	0.267	0.742	0.793	0.767	0.527	0.714	0.630
Low	0.733	0.207	0.786	0.733	0.759	0.527	0.714	0.705

=== Confusion Matrix ===

a b <-- classified as

23 6 | a = High

8 22 | b = Low

Implementation of the JRIP algorithm using normalised binned data, and experimenting with algorithm parameters

In order to further improve the performance of the JRIP model, I binned the normalised data into 3 groups:

- No pattern
- Low frequency
- High/medium frequency

The **resulting model generated 3 rules**, relying on pattern B and pattern C to classify the data.

JRIP rules:

==========

(Pattern B Follow up praise and guided questioning = High/medium frequency) => Tutor ranking=High (20.0/3.0)

(Pattern C Open-ended prompting and observe reasoning = High/medium frequency) => Tutor ranking=High (15.0/5.0)

=> Tutor ranking=Low (24.0/2.0)

The model was tested using ten-fold cross-validation, and resulted in **83% of classification accuracy**. This is an improvement in model performance using normalised data, without binning. **Binning the data** also results in a **significant improvement** in the ability of the model to classify high-ranked tutors, with both precision and recall improving for high-ranked tutors:

- The previous JRIP model had a precision of 74% and this improves to 77.1% for high-ranked tutors
- In addition, the previous JRIP model had a recall of 79.1% for high-ranked tutors which improved to 93.1%.

While the precision for **low-ranked tutors** improves (78.6% to 91.7%), **recall remains the same** at 73.3%. This reflects the variation that occur in a minority of the low-ranked tutor dataset, where the tutoring practices employed by the low-ranked tutors are more sophisticated than expected. It may indicate that low-ranked tutors have improved their practice following their last formal evaluation and ranking.

Table 47. Stratified cross validation (JRIP, dataset E, normalised and binned data).

Correctly classified instances	49 (83.1%)
Incorrectly classified instances	10 (16.9%)
Kappa statistic	*0.66*

Table 48. detailed accuracy by class (JRIP, dataset E, normalised and binned data).

Class	TP rate	FP rate	Precision	Recall	F-measure	MCC	ROC area	PRC area
High	0.931	0.267	0.771	0.931	0.844	0.676	0.808	0.734
Low	0.733	0.069	0.917	0.733	0.815	0.676	0.808	0.790

=== Confusion Matrix ===

a b <-- classified as

27 2 | a = High

8 22 | b = Low

In addition to binning the data, I experimented with adjusting the number of folds of data used by the algorithm for pruning. However, both increasing and decreasing the number of folds from 3.0 had a detrimental effect on the performance of the algorithm, and the default parameter settings were therefore retained. The default setting for the minNo parameter, which refers to the minimum weight of instances that must be covered by a rule, were also retained at 2.0; only 3 rules were generated by the JRIP model, and these did not require simplification through increasing the minNo value.

J48 algorithm

Implementation using the algorithm default parameters, and normalised data-set

In the next stage of the work, I experimented with using the J48 algorithm to classify tutors into high-ranked and low-ranked groups. Once again, the default parameters for the algorithm were used. The normalised and binned i-support of patterns A-C were used as attributes of the sessions. The resulting model is set out in Figure 35.

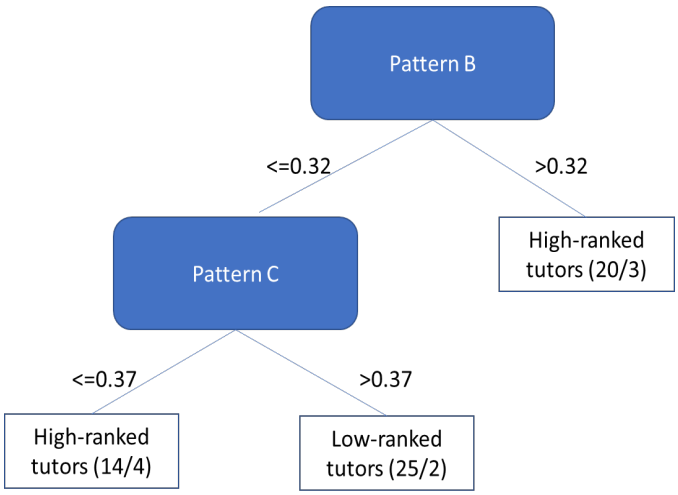


Figure 35. Decision tree for dataset E, using normalised, un-binned data

The model was tested with a stratified 10-fold cross validation, and a **classification accuracy of 76%** was achieved.

Table 49. Stratified cross validation (J48, dataset E).

Correctly classified instances	45 (76.3%)
Incorrectly classified instances	14 (23.7%)
Kappa statistic	*0.53*

Table 50. detailed accuracy by class (J48, dataset E).

Class	TP rate	FP rate	Precision	Recall	F-measure	MCC	ROC area	PRC area
High	0.793	0.267	0.742	0.793	0.767	0.527	0.710	0.630
Low	0.733	0.207	0.786	0.733	0.759	0.527	0.710	0.693

=== Confusion Matrix ===

a b <-- classified as

23 6 | a = High

8 22 | b = Low

Implementation of the J48 algorithm using normalised binned data, and experimenting with algorithm parameters

In the next stage, I experimented with binning the normalised data to explore whether this would further improve the performance of the model. The frequency data was binned into the following categories:

- High/medium frequency
- Low frequency
- No pattern

The J48 algorithm was implemented and the decision tree visualised in Figure 36 was obtained. The model was tested through 10-fold cross validation and a **classification accuracy of 77.9%** was achieved.

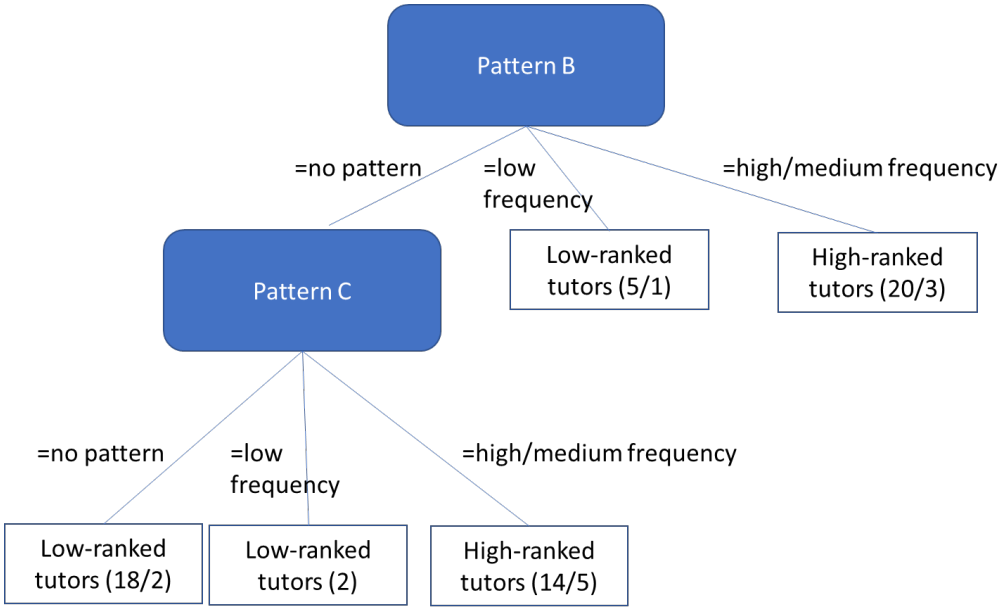


Figure 36. Decision tree for dataset E, using normalised, binned data

Table 51. Stratified cross validation (J48, dataset E, normalised and binned data).

Correctly classified instances	46 (77.9%)
Incorrectly classified instances	13 (22%)
Kappa statistic	*0.56*

Table 52. detailed accuracy by class (J48, dataset E, normalised and binned data).

Class	TP rate	FP rate	Precision	Recall	F-measure	MCC	ROC area	PRC area
High	0.828	0.267	0.750	0.828	0.787	0.563	0.783	0.722
Low	0.733	0.172	0.815	0.733	0.772	0.563	0.783	0.744

=== Confusion Matrix ===

a b <-- classified as

24 5 | a = High

8 22 | b = Low

I also experimented with adjusting the value of the **minNo parameter to 15 instances**. This resulted in a simpler tree as visualised in Figure 37, but also slightly reduced the classification accuracy to 75%. In particular, the ability of the tree to detect all high-ranked tutors is negatively affected, with **recall of high-ranked tutors falling to 59%.** Given the lower overall classification accuracy, and the diminished recall for high-ranked tutors, the **default parameters** for the J48 algorithm appeared **optimal.**

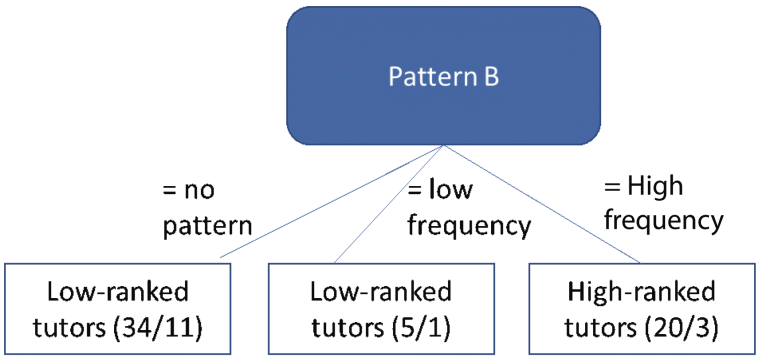


Figure 37. Decision tree for dataset E, using normalised, binned data (minNo parameter tuned to 15)

Table 53. Stratified cross validation ((J48, dataset E, normalised and binned data, minNo=15).

Correctly classified instances	44 (74.6%)
Incorrectly classified instances	15 (25.4%)

Table 54. detailed accuracy by class ((J48, dataset E, normalised and binned data, minNo=15).

Class	TP rate	FP rate	Precision	Recall	F-measure	MCC	ROC area	PRC area
High	0.586	0.100	0.850	0.586	0.694	0.513	0.662	0.653
Low	0.900	0.414	0.692	0.900	0.783	0.513	0.662	0.641

=== Confusion Matrix ===

a b <-- classified as

17 12 | a = High

3 27 | b = Low

Analysis of the results for dataset E (evaluator rankings used as criterion for performance)

The **classification algorithms perform well** for dataset E, with both JRIP and J48 achieving classification accuracies exceeding 75%. All implementations rely on patterns B *Follow up praise and guided questioning* and C *Open-ended prompting and observe reasoning* to classify data. Pattern B is the tutor practice of praising and then asking specific questions to test or extend understanding, following the learner solving a problem, and pattern C is the tutor eliciting learner reasoning through asking open ended questions, and then encouraging reasoning through passive observation.

The use of patterns B and C is **particularly effective in detecting high-ranked tutors**, with high rates of precision and recall being achieved across models. **Recall is slightly lower for low-ranked tutors** versus high-ranked tutors, indicating that a minority of low-ranked tutors demonstrate pattern B *Follow up praise and guided questioning* and C *Open-ended prompting and observe reasoning* in a manner which is more typical of high-ranked tutors. The minority of low-ranked tutors that are misclassified are using sophisticated practices to extend learner understanding, and elicit reasoning more frequently than expected.

These results could indicate that low-ranked tutors have **improved their practice** following their last formal evaluation and ranking.

The best performance is attained by an implementation of JRIP, using normalised and binned data, which has a classification accuracy of approximately 83%, versus 76% using normalised un-binned data. The improvement in classifier accuracy reflects the value of binning data to smooth out the effect of variations within the data, and enable the classifier to detect tutor performance more effectively. This is particularly true for high-ranked tutors, who demonstrate variation in the pattern frequency over a certain threshold.

Overall, the performance of classifiers is acceptable, and reflects the **strong relationship** between **patterns** of tutor-learner engagement and **tutor rankings**. The classifier performance is particularly good considering that the tutor rankings are determined by human evaluators and are subject to bias and inconsistency. In future research, it would be interesting to repeat this analysis using sessions that have been evaluated by human evaluators (rather than sessions randomly selected from ranked tutors). This may further improve classifier performance. In addition, classifiers trained using this data would be more effective in identifying high-performing versus low-performing practices from an operational perspective, and ensuring that tutors obtained the relevant support and training.

5.3.3 Implementation of predictive analytics for dataset F (using APS as performance criterion)

As per the analysis for the dataset E, I selected the JRIP and J48 algorithms for implementation for dataset F. The attributes for the dataset are the i-support for significant patterns, set out in in Section 5.2, and are set out below and named for convenience:

- Pattern A: Tutor asks open ended, conceptual questions–learner reasons–tutor praise (named as *'Prompt and praise reasoning'*).

- Pattern B: Tutor asks learner a specific question–tutor asks open ended conceptual questions–learner reasons (named as '*Prompt reasoning through guided and open-ended questions'*).
- Pattern C: Tutor asks open ended conceptual questions–learner reasons–tutor passive–learner reasons (named as *'Open-ended questioning and observe reasoning'*).

I initially implemented the algorithms using the default parameters, with data being normalised for session duration. I then experimented with binning the data, and tuning the algorithm parameters to improve classifier performance.

JRIP algorithm

Implementation using the algorithm default parameters, and normalised data-set

The JRIP algorithm was implemented using the default parameters, and normalised, un-binned data. **2 rules were generated**, relying on pattern B to classify tutors into high and low-performing groups.

(Pattern B Prompt reasoning through guided and open-ended questions = 0) => Tutor ranking=Low (27.0/10.0)

=> Tutor ranking=High (14.0/3.0)

The model was tested through stratified ten-fold cross validation, and a **classification accuracy of approximately 63%** was achieved. While recall for low APS tutors is fairly high, at 75%, the **recall for high APS tutors is low at only 52%.**

Table 55. Stratified cross validation (JRIP, dataset F).

Correctly classified instances	26 (63.4%)
Incorrectly classified instances	15 (36.6%)
Kappa statistic	0.27

Table 56. detailed accuracy by class (JRIP, dataset F).

Class	TP rate	FP rate	Precision	Recall	F-measure	MCC	ROC area	PRC area
High	0.524	0.250	0.688	0.524	0.595	0.281	0.555	0.571
Low	0.750	0.476	0.600	0.750	0.667	0.281	0.555	0.513

=== Confusion Matrix ===

a b <-- classified as

11 10 | a = High

5 15 | b = Low

Implementation of JRIP using normalised, binned data-set

In the next stage, JRIP was implemented with a normalised, binned data-set. The data was separated into 4 bins: no pattern, low frequency, medium frequency and high frequency.

To improve the classifier performance, pattern C *Open-ended questioning and observe reasoning* was removed as an attribute, given the high standard deviation for this pattern in both tutor groups. Two rules were generated from the resulting data-set, once again relying on pattern B.

(Pattern B Prompt reasoning through guided and open-ended questions = No pattern) => Category=Low (27.0/10.0)

=> Category=High (14.0/3.0)

Testing the model with 10-fold cross validation achieved a slightly lower **performance of 60.9%.** However, this model is better able to detect **high-performing tutors, with a recall of 57%** (versus the 52% using normalised, un-binned data and all attributes).

Table 57. Stratified cross validation (JRIP, dataset F, normalised and binned data).

Correctly classified instances	25 (60.9%)
Incorrectly classified instances	16 (39%)
Kappa statistic	0.22

Table 58. detailed accuracy by class (JRIP, dataset F, normalised and binned data).

Class	TP rate	FP rate	Precision	Recall	F-measure	MCC	ROC area	PRC area
High	0.571	0.350	0.632	0.571	0.600	0.222	0.548	0.557
Low	0.650	0.429	0.591	0.650	0.619	0.222	0.548	0.505

=== Confusion Matrix ===

a b <-- classified as

12 9 | a = High

7 13 | b = Low

Implementation of J48 algorithm

Implementation using the algorithm default parameters, and normalised data-set

I implemented the J48 algorithm, using normalised (un-binned) data, and the default parameters for the algorithm. The resulting tree is visualised in Figure 38.

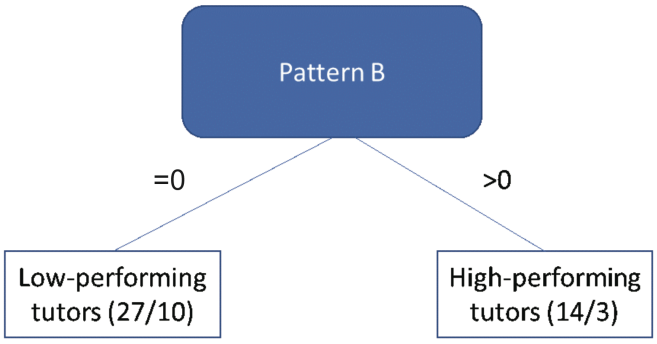


Figure 38. Decision tree for dataset F, using normalised, un-binned data

However, on testing the model with ten-fold cross validation, the classifier only achieves a **classifier accuracy of 54%.** It should be noted that there are only two classes of tutors, and the accuracy achieved is only slightly higher than through the baseline of chance. Further, the rate of recall is low, particularly for high-performing tutors (less than 48%). This indicates that solely relying on pattern B *Prompt reasoning through guided and open-ended questions* to classify the data is not effective to detect tutor performance. In particular, many high APS tutors may not demonstrate pattern B, and will not be detected based on this classification approach.

Table 59. Stratified cross validation (J48, dataset F).

Correctly classified instances	22 (53.6%)
Incorrectly classified instances	19 (46.3%)
Kappa statistic	0.07

Table 60. detailed accuracy by class (J48, dataset F).

Class	TP rate	FP rate	Precision	Recall	F-measure	MCC	ROC area	PRC area
High	0.476	0.400	0.556	0.476	0.513	0.077	0.500	0.506
Low	0.600	0.524	0.522	0.600	0.558	0.077	0.500	0.486

=== Confusion Matrix ===

a b <-- classified as

10 11 | a = High

8 12 | b = Low

Implementation using the normalised and binned data-set

In order to improve the performance of the J48 algorithm, the data was binned into 4 bins (no pattern, low frequency, medium frequency and high frequency). The algorithm was implemented again using default parameters and the decision tree in Figure 39 was obtained.

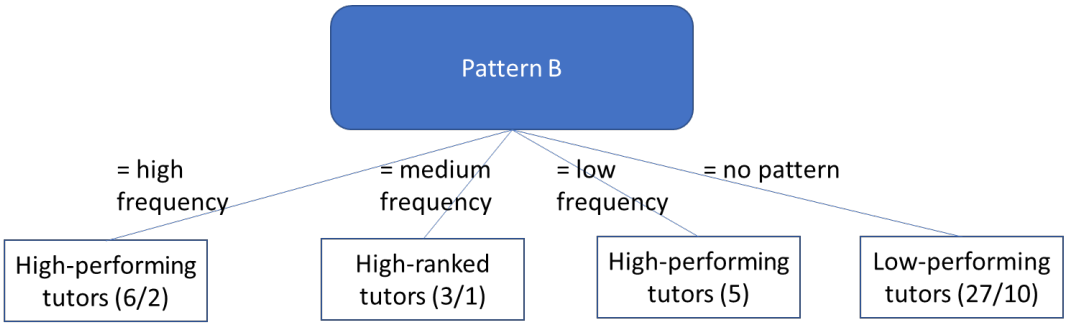


Figure 39. Decision tree for dataset F, using normalised, binned data

Testing the model with 10-fold cross validation demonstrates a **classifier accuracy of 68%,** a significant improvement from the classification accuracy achieved when data was used without binning. However, the recall for high-performing tutors remains low, at 52%, indicating that there are many high-performing tutors who do not demonstrate pattern B

Prompt reasoning through guided and open-ended questions. On the other hand, the absence of pattern B does seem to be an effective indicator for detecting low-performing tutors (as shown by a recall of 85%).

Table 61. Stratified cross validation (J48, dataset F, normalised and binned).

Correctly classified instances	28 (68.3%)
Incorrectly classified instances	13 (31.7%)
Kappa statistic	0.37

Table 62. Detailed accuracy by class (J48, dataset F, normalised and binned).

Class	TP rate	FP rate	Precision	Recall	F-measure	MCC	ROC area	PRC area
High	0.524	0.150	0.786	0.524	0.629	0.394	0.583	0.654
Low	0.850	0.476	0.630	0.850	0.723	0.394	0.583	0.527

=== Confusion Matrix ===

a b <-- classified as

11 10 | a = High

3 17 | b = Low

In an attempt to improve the recall for high-performing tutors, I **experimented with removing pattern B** *Prompt reasoning through guided and open-ended questions* as an attribute. The resulting tree is presented below.

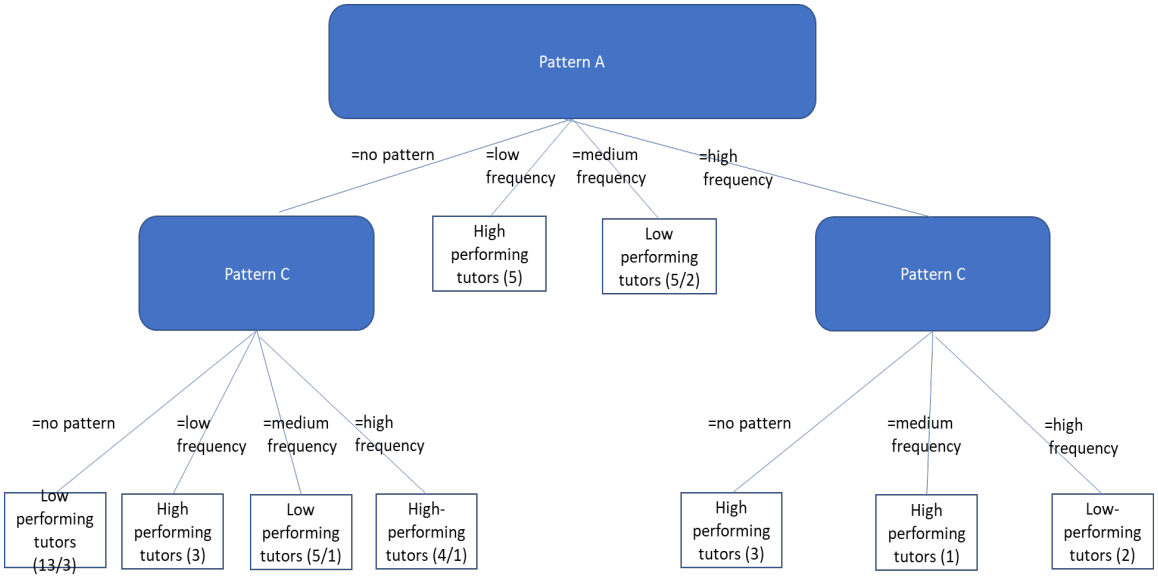


Figure 40. Decision tree for dataset F, using normalised, binned data (pattern B removed a attribute)

Table 63. Stratified cross validation (J48, dataset F, normalised and binned, remove pattern B).

Correctly classified instances	29 (70.7%)
Incorrectly classified instances	12 (29.3%)
Kappa statistic	0.41

Table 64. detailed accuracy by class (J48, dataset F, normalised and binned, remove pattern B).

Class	TP rate	FP rate	Precision	Recall	F-measure	MCC	ROC area	PRC area
High	0.714	0.300	0.714	0.714	0.714	0.414	0.681	0.723
Low	0.700	0.286	0.700	0.700	0.700	0.414	0.681	0.597

=== Confusion Matrices ===

a b <-- classified as

15 6 | a = High

6 14 | b = Low

While this tree achieves a better recall (~70%) for high APS tutors, it is much bigger and may risk overfitting the dataset. While this classifier may not be suited for the dataset, it does **illustrate the significant variations within the data-set for patterns A** *Prompt and*

praise reasoning **and C** *Open-ended questioning and observe reasoning*. Both low and high-performing tutors seem to use patterns which involve open ended questioning in a variable manner.

Analysis

The J48 classifier performs best for dataset F, when the data is normalised and binned. Even in this scenario, the **highest classification accuracy attained** through the **implementation of J48 using normalised and binned data is only 68%.** The resulting model uses pattern B *Prompt reasoning through guided and open-ended questions* to classify the tutors into groups. Pattern B involves the tutor asking specific questions to scaffold the learner, prior to asking open ended questions to elicit reasoning. While JRIP also generates rules using pattern B, the performance of the algorithm with this dataset is weaker (61% accuracy).

The use of **pattern B** to classify tutors into high and low-performing groups is **more effective for low APS tutors versus high APS tutors**, as demonstrated by the higher recall rates for low APS tutors. Most low APS tutors do not demonstrate pattern B, therefore being detected by the model. However, there is also a significant minority of high APS tutors which also do not demonstrate pattern B, and these tutors are misclassified as being low APS. There is also significant variation for both low and high APS tutors in the frequency of patterns A *Prompt and praise reasoning* and C *Open-ended questioning and observe reasoning*, demonstrated by the large, overfitted models generated when these patterns are used as the sole attributes.

The performance of the classifiers reflects the relationship between the significant sequences and tutor performance. While the absence of certain sequences is a strong indicator that a tutor has low APS, it is harder to detect high-performing tutors based on the frequency of the identified sequences. It is worth noting that tutor performance in this dataset is

determined by learner progress, rather than tutor ranking for student-centred learning. High APS tutors may be demonstrating tutor-led (rather than student-centred) practices to **scaffold learners** where the learner requires additional support. This is reflected in the **high variation** in the frequency of **patterns of engagement**, with respect to self-regulation.

Summary

In this Section, I have explored whether patterns of tutor-learner interactions can be used to predict tutor performance, with tutor performance measured by either (1) evaluator rankings or (2) average progress scores. I have experimented with the implementation of the JRIP and J48 algorithms, and with the use of binned and un-binned datasets.

Patterns B (*Follow up praise and guided questioning)* and C (*Open-ended prompting and observe reasoning)* were automatically selected to classify tutors into high versus low-performing groups, when performance is measured by evaluator rankings. The best performance is attained by an implementation of JRIP, using normalised and binned data. This implementation has a classification accuracy of approximately 83%, versus 76% using normalised un-binned data. The improvement in classifier accuracy reflects the value of binning data to smooth out the effect of variations within the data, and enables the classifier to detect tutor performance more effectively. **Recall is slightly lower for low-ranked tutors** versus high-ranked tutors, indicating that a minority of low-ranked tutors are using sophisticated practices to extend learner understanding, and elicit reasoning, more frequently than expected. These results could indicate that some low-ranked tutors have **improved their practice** following their last formal evaluation and ranking.

Pattern B (*Prompt reasoning through guided and open-ended questions)* was automatically selected to classify tutors into high versus low-performing groups, when performance is measured by average progress scores (APS). The **highest classification accuracy attained** is through the **implementation of J48 using normalised and binned**

data (68%). The use of **pattern B** to classify tutors into high and low-performing groups is much **more effective for low APS tutors versus high APS tutors**, as demonstrated by the higher recall rates for low APS tutors. Most low APS tutors do not demonstrate pattern B, therefore being detected by the model. However, there is also a large minority of high APS tutors which also do not demonstrate pattern B, and these tutors are misclassified as being low APS.

It is interesting to note that a higher level of accuracy is achieved for classifiers when performance is measured by evaluator rankings (dataset E), versus being measured by average progress scores (dataset F). This is most likely because the identified patterns of tutor-learner interaction are based on a framework of signifiers of co-regulation (as set out in Section 4.2), which is closely related to the construct of student-centred teaching. While a positive relationship between SRL and learning outcomes is well-established (see van de Graff, 2022), limited research has explored the relationship between co-regulatory teaching practices and learning outcomes. It is possible that **high APS tutors** implement tutor-led (rather than student centred) practices to **scaffold learners** and boost learning outcomes. This may result in greater variation in the patterns of engagement with respect to SRL.

Chapter 6 Discussion

Introduction

My research explored the differences in how high versus low-performing tutors co-regulate learners in a virtual classroom environment. My research was intended to be both theoretically grounded, as well as a practical representation of how tutors co-regulate learners in practice. In order to ensure that the research had a strong theoretical foundation, I contextualised an established model of self-regulated learning (SRL), and used this as the basis for a framework of signifiers. Selected forms of learning analytics were used to implement the framework of signifiers, and develop a rich understanding of how tutors co-regulate learners in practice.

In this chapter, I will start by reviewing the contextualised model of SRL developed through my research, and the associated framework of signifiers. In particular, I will consider the benefits and challenges of using an existing architecture which is designed to study SRL across generalizable contexts, and the effectiveness of the resulting framework of signifiers.

I will move on to reflect on the effectiveness of the selected forms of learning analytics in addressing RQ3 and identifying key differences between high versus low-performing tutors with respect to co-regulation. Specifically, I will consider the extent to which the implementation of learning analytics enabled both holistic and granular, statistically significant differences to be identified.

Finally, I will discuss the key differences detected between high versus low-performing tutors with respect to co-regulation, considering datasets where tutor performance is measured by evaluator rankings, and where tutor performance is measured by average progress scores. I will consider the differences in the broad processes adopted by tutors, as well as the significant isolated patterns of interaction demonstrated by high-performing versus low-performing tutors, and the associated operational implications.

6.1 Theoretical implications

As described in Section 2.3 of this thesis, there is a need for theoretically based frameworks which enable the exploration of co-regulation in a VCE. The majority of the current literature either hasn't applied a theoretical framework, or has developed frameworks tailored for a traditional classroom environment (using large grained forms of data). My research developed a contextualised model of SRL, which provided the foundation for theoretical frameworks and enabled an exploration of co-regulation, utilising the fine-grained data generated from online contexts.

I developed a contextualised model of SRL, which is based on the Winne and Hadwin (1998) model of SRL (set out in Figure 41 and detailed in Section 4.1). The Winne and Hadwin model is structured by a component of SRL (COPES) rather than by phase of SRL. The components of SRL include Conditions, Operations, Products, Evaluations, and Standards. The model emphasises the relationship between components of SRL, with SRL being visualised as a weakly sequenced and recursive process (see Section 2.1 for a more detailed overview of the model).

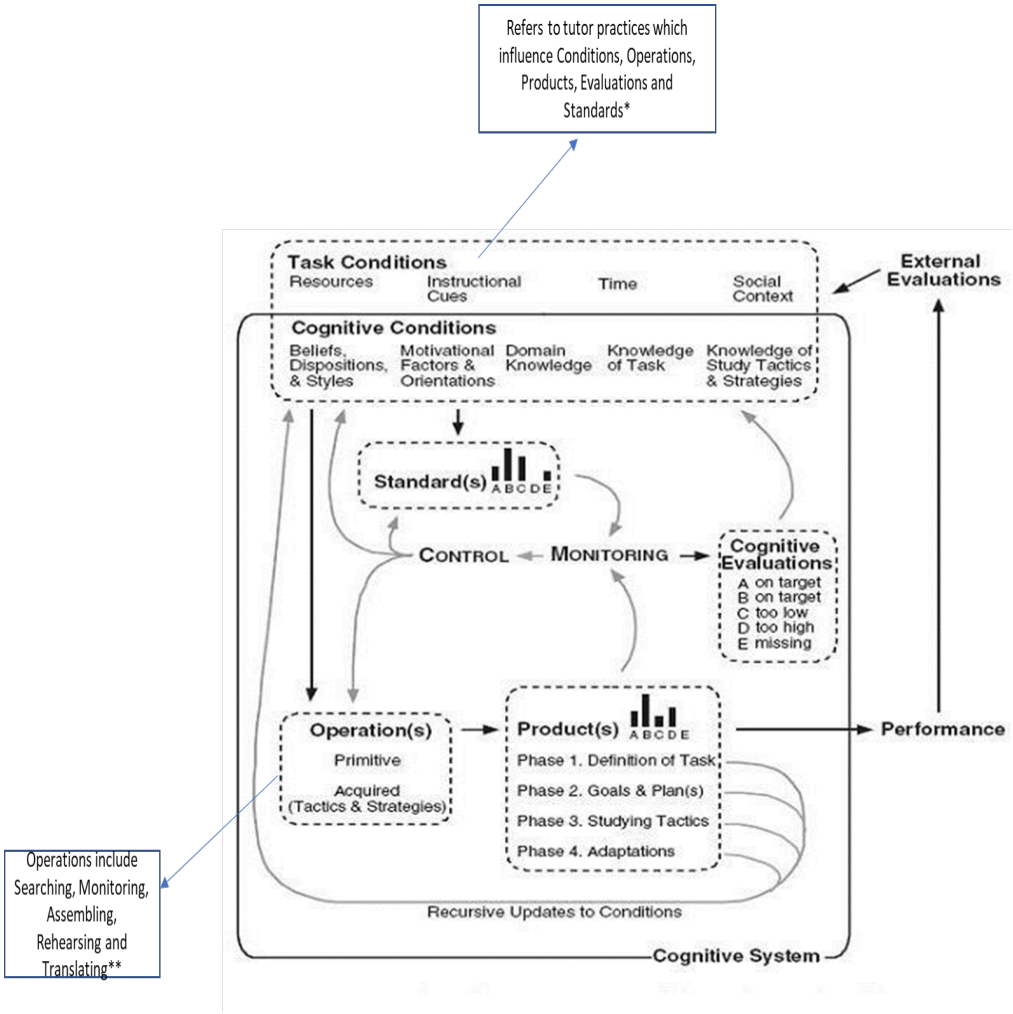


Figure 41. Contextualised model of SRL

Suitability of the Winne and Hadwin (1998) model to the study context: holistic model structure

The holistic structure of the Winne and Hadwin (1998) model made it well suited to exploring SRL in an under-researched study context, such as tutor co-regulation. The architecture of the model enabled it to act as a comprehensive foundation for ensuring that the relevant elements of tutor-learner interactions were captured, despite the limited research conducted into co-regulation to date. For example, the architecture of the Winne and Hadwin model suggests that tutor practices (a type of Condition) can influence learner Operations (i.e. the cognitive processes adopted by learner). Specific examples of tutor co-regulatory

practices were identified, by conducting a literature review and exploratory data analysis that identify tutor co-regulatory practices which may influence learner cognitive processing.

Similarly, the model architecture can be used to identify additional co-regulatory practices by mapping the relationship between tutor practices to other components of the model. In addition to learner Operations, the other components of the model include Conditions (e.g. learner motivation), Products (e.g. learner goals), Evaluation (e.g. learner assessment of their performance), and Standards (e.g. learner perception of high- quality work). Tutor practices can be mapped to other components of the model through a combination of a literature review and exploratory data analysis. While research into tutoring practices has been relatively limited, there is a substantial body of literature with regard to each of these components in related contexts (e.g. classrooms, and intelligent tutoring systems). This literature was drawn on and interpreted for my study context, to hypothesise about potential tutoring practices which may influence aspects of learner SRL. For example, the steps below were followed to identify relevant tutoring practices with respect to internal Conditions:

- The architecture of the model suggests that there is a relationship between tutor practices and motivation
- A literature review demonstrated that there is a substantial body of literature which shows that motivation can be promoted through rewarding effort (e.g. Dweck, 2012)
- An exploratory data analysis showed that tutors can reward effort through praise, and through awarding effort points or stickers.
- A tutoring co-regulatory practice to boost learner motivation through specific types of praise, and awarding effort points/stickers, was highlighted

The holistic structure of the Winne and Hadwin (1998) model therefore enabled the elicitation of insights about a broad set of tutoring practices with respect to co-regulation. In an under-researched area, such as tutor co-regulation, this was important in discovering key differences between groups of tutors. For example, identifying tutor co-regulatory practices with respect to learner cognitive engagement (Operations), as well as learner motivation (Conditions), enabled the identification of distinct patterns; in contrast to high-performing tutors, low-performing tutors tended to use motivational practices to reward shallow forms of cognitive engagement, such as learner rehearsing.

Suitability of the Winne and Hadwin (1998) model to the study context: information processing architecture

The architecture of the model was suited to my study context, as it is oriented around information processing through components of SRL rather than by phase of SRL. My research demonstrated that this orientation aligns better with the context of unstructured tutor-learner interactions than a model oriented by phase of SRL. For example, the model architecture maps a relationship between tutor practices (Conditions), and learner cognitive processing (Operations). It therefore acts as a foundation for analysis of how tutors can cue different depths of cognitive processing by learners, within the broader context of co-regulation. As my research findings demonstrate, tutor co-regulatory practices with respect to cognitive engagement are a significant factor in differentiating high-performing tutors from low-performing tutors. High-performing tutors tend to use a variety of open-ended practices, while low-performing tutors rely on practices which tend to be tutor led, and transfer limited autonomy to the learner for cognitive engagement. The relationship between Conditions and Operations in the model architecture is therefore key to my research results.

While high-performing tutors are more likely to encourage open ended forms of engagement, my results indicate that they do not demonstrate distinct practices with regard to

the explicit prompting of SRL products across phases of SRL (such as task definition or goal setting). The Winne and Hadwin model may therefore be better suited to exploring co-regulation in the context of authentic human tutoring, versus models which are structured by phase of SRL.

Challenges of contextualising the Winne and Hadwin (1998) model

While the COPES architecture of the model enables a granular approach towards analysing relationships between components of SRL, the components themselves are still relatively macro. A significant amount of research was required to contextualise the components to a specific study context, through extensive literature review and data analysis. For example, while instructional cues are included as a potential external condition within the COPES model, the model does not include examples of the types of instructional cues that may influence SRL. Similarly, the model refers to the instructional materials as a relevant external condition, but does not specifically address factors such as online learning design. While acknowledging the significant amount of research required to contextualise the Winne and Hadwin (1998) model, the macro nature of the model components is somewhat inevitable given the intention to make the model generalizable across contexts, and the increasingly varied nature of online learning environments.

Summary

Overall, my results demonstrate that the Winne and Hadwin model provides a strong foundation for exploring co-regulation in a VCE. The holistic architecture of the model, focused on information processing, aligns well with real-life tutor-learner interactions. While a significant amount of work is required to contextualise the model to my study context, this is inevitable given that the model is designed to study SRL across generalizable contexts.

The results from my research indicate potential for theoretical development in further work (outside of the scope of this thesis). For example, future research can explore whether

the model could potentially be further built upon, to show favourable and maladaptive pathways, as process-oriented research into SRL advances. The process maps generated by high-performing tutors show that open-ended tutoring practices often lead to sophisticated learner operations (such as reasoning). Further research (outside of the scope of this thesis) can explore the subsequent impact of learner reasoning on inter-subjectivity and learner motivation/sense of well-being, to highlight favourable tutor-learner loops, with respect to SRL.

6.2 Implications for framework of signifiers

I used the contextualised model of SRL as the foundation for a framework of signifiers, representing how tutors co-regulate learners in a VCE. As described in Section 6.1, a mixed methods approach was used to design the framework, including both literature review and exploratory data analysis. The framework of signifiers for tutor co-regulatory practices is set out in Table 29, and the framework of signifiers for learner operations is set out in Table 31.

Mixed methods approach to design framework of signifiers

Through adopting a mixed methods approach, the framework of signifiers identifies features of tutor-learner interaction that are both theoretically grounded, as well as being practically relevant. For example, my results show that key differences between the process maps for the two groups of sessions relate to the forms of tutor engagement deployed. These forms of engagement include tutor directive, explanatory, guided, open ended and passive engagement. A mixed methods approach was necessary to identify these categories of engagement, including both literature review (to ensure the categories of engagement were theoretically grounded) and observations of sessions being conducted in the VCE (to ensure that the categories of engagement were tailored to the study context). This approach enabled

the identification of nuanced events, which may not be highlighted in theoretical models of SRL, for example directive engagement versus explanatory engagement.

Framework structure and design

The frameworks were designed to include two levels of granularity; categories and observable signifiers. The categories map to theoretical constructs within the contextualised model, while the signifiers are the observable actions taken, which map to the categories. For example, the contextualised model acknowledges that learner internal conditions, such as motivation, influence SRL. This theoretical construct is mapped to the category of tutoring practices, 'tutor boosting learner motivation'. This category is then mapped to a number of signifiers, such as tutor praising the learner or offering rewards, such as effort points and stickers. These two levels of granularity allow the frameworks to be implemented in a manner which (1) immediately aligns with theoretical models, and (2) are a practical representation of tutor-learner interactions.

The results from the learning analytics also support the inclusion of a wide range of categories of signifiers within the framework of signifiers. The distinct processes discovered are informed by the diverse range of events included in the framework. For example, the inclusion of both nuanced forms of engagement, and motivational co-regulatory practices enabled significant patterns of tutor-learner interaction to be identified. The high number of events and associated transitions mean that process maps need to viewed at high levels to discover a coherent process. It would be possible in later research, to focus the analysis on certain domains of the framework (e.g., metacognitive or affective). This would limit the number of events/transitions and enable a more detailed process to be visualised. However, it is important to recognise that focusing on specific domains of the framework would obscure features of the overall process, such as structure or reliance on certain events/transitions.

It is also notable that process maps produced using categories, versus signifiers, differ from each other with regard to the inclusion of certain events. The process maps produced using categories as events include tutor metacognitive practices, and learner searching operations, within the high-level process maps. In contrast, the process maps produced using signifiers as events do not include the signifiers associated with metacognitive practices or learner searching operations. This could be due to an imbalance in the number of types of signifiers across event categories, which means that some event categories are more likely to be displayed than the associated signifiers. Further research could experiment with limiting the number of types of signifiers that can be used per event category.

Data types explored, in the framework of signifiers

The frameworks benefit from the use of natural language dialogue as a rich source of signifiers. Natural language dialogue is key to developing a rounded understanding of tutor-learner interactions. This was evidenced by the nature of the prominent signifiers highlighted through the implementation of learning analytics, with key signifiers being drawn from natural language dialogue (e.g. open-ended prompts). The signifiers were labelled using human judgment, rather than automated processes, as my research focused on identifying and exploring key patterns of behaviour with respect to co-regulation. However, advances in foundational large language models (LLMs) make automated labelling of signifiers an interesting topic for future research. Automating the signifier labelling process would enable data to be collected at scale. In addition, it would potentially add further layers of granularity to the data collected, such as start and end time stamps for events.

Summary

Overall, the mixed methods approach enables a representation of tutor-learner interactions, which is both practical and theoretically aligned. The design of the framework of signifiers, and the inclusion of a diverse range of events, allows key differences between

sessions conducted by high versus low-performing tutors to be identified. In addition, natural language dialogue has proven to be a rich source of data which informs our understanding of how co-regulation unfolds in practice in a VCE. Future research could explore opportunities to automate the labelling process using LLMs, which both enable data to be collected at scale using the framework of signifiers.

6.3 Approach towards the use of learning analytics

My research aimed to identify (1) holistic and (2) statistically significant differences between sessions conducted by high-performing tutors versus low-performing tutors (see RQ3). In order to meet both of these research goals, a combination of different types of learning analytics were implemented. The learning analytics implemented explored key differences between the co-regulatory practices of high and low-performing tutors from different perspectives. In order to identify holistic differences between high-performing versus low-performing tutors, process mining was used to discover the latent processes shaping sessions conducted by high versus low-performing tutors. On the other hand, pattern mining and decision trees enabled a more granular analysis of the key differences between tutor classes, and identified statistically significant differences.

In the following section, I reflect on the benefits of adopting a layered approach towards analysing co-regulation in a VCE. I start off by considering the benefits of implementing process mining as the foundation of my work (versus simple frequency-based analytics), prior to considering the value of implementing statistical forms of analysis.

Benefits of implementing process mining

Process mining was implemented to discover the holistic, structural differences between sessions conducted by high versus low-performing tutors. Process mining was fundamental to my analysis, as it discovered distinct, theoretically grounded processes for co-regulation in sessions conducted by high-performing versus low-performing tutors. These

holistic visualisations can be interpreted in light of the contextualised model of SRL (set out in Section 4.1). They can therefore act as the basis for a rich exploration of how co-regulation occurs in practice, generating insights which extend beyond simple frequency-based analytics.

Process mining is able to extend the insights obtained via simple frequency-based analytics through driving an understanding of the underlying process. For example, analysing the relative frequency of events in sessions conducted by low-ranked versus high-ranked tutors illustrated that high-ranked tutors rely on a more diverse range of practices than low-ranked tutors. For example, the pie charts in Figures 42 and 43 show that the events comprising 80% of all events in sessions conducted by high-ranked tutors include both tutor-controlled practices (e.g., tutor guided engagement), and open-ended practices (e.g., learner translating and tutor open-ended engagement). On the other hand, the events comprising 80% of all events in sessions conducted by low-ranked tutors focus mainly on tutor-controlled practices (e.g., tutor guided engagement and tutor directive engagement) or passive engagement. However, while this frequency-based analysis is useful, it does not enable a detailed exploration of the role of different events within the overall process.

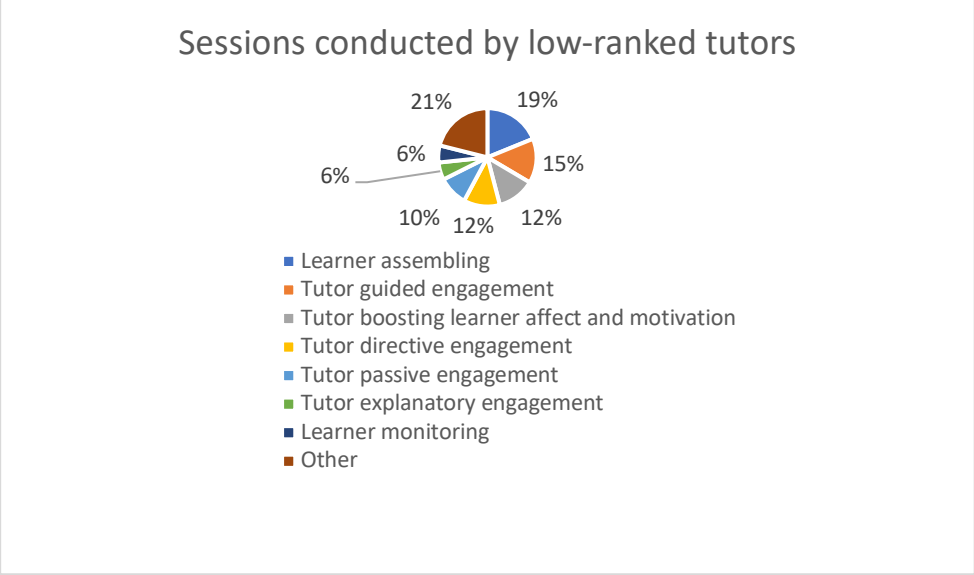


Figure 42. Relative frequency of events in sessions conducted by low-ranked tutors

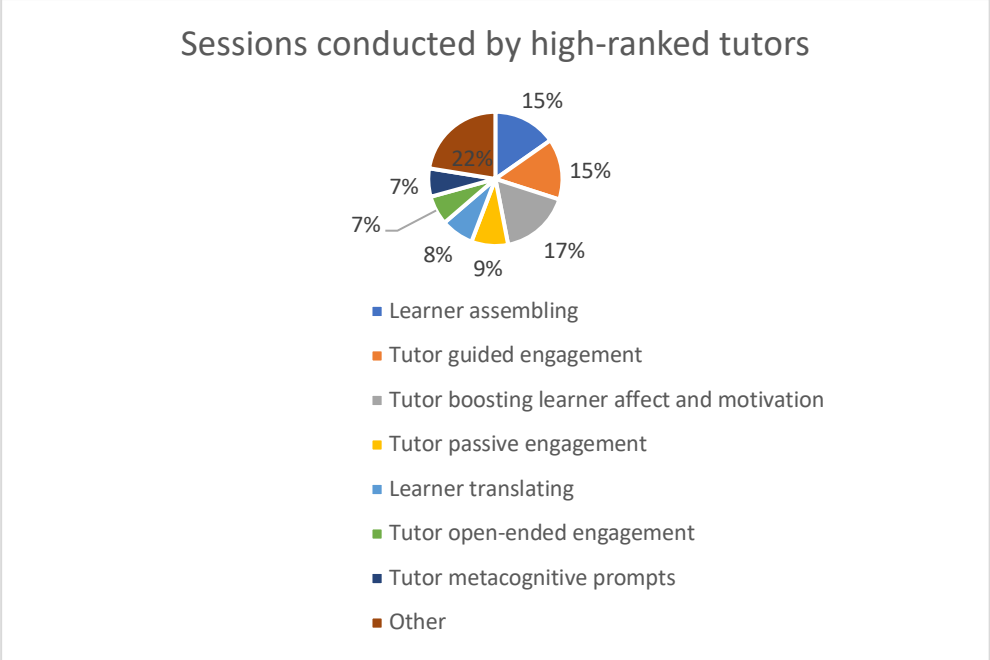


Figure 43. Relative frequency of events in sessions conducted by high-ranked tutors

For example, the process maps for high-ranked tutors demonstrate that tutor open-ended engagement has a prominent role within the process, with multiple transitions to other events. Similarly, the process maps for low-ranked tutors demonstrate that low-ranked tutors are frequently stuck in a repetitive path of tutor-controlled events, with transitions between tutor directive engagement-tutor guided engagement-tutor explanatory engagement being common. The holistic nature of the output from process mining has operational implications. For example, the results from process mining could be used to enable the development of detailed, theoretically founded rubrics for tutor assessment, and guide tutor training. Opportunities for future research are discussed in Chapter 7.

Benefits of implementing sequence mining and decision trees

The descriptive nature of process mining means that the results cannot be statistically validated. In order to identify statistically meaningful, granular patterns of interactions, which can be used as the basis for operational intervention, other forms of learning analytics were required. Additional layers of analytics, comprising of sequence mining and decision trees, were therefore implemented. Sequence mining enabled the identification of isolated patterns

of co-regulation, which can be statistically validated as being significant. In addition, the results from sequence mining provided an additional layer of granularity, by identifying patterns both across and within sessions. The identified sequences were used to build decision trees, which classified tutors as high or low-performing, highlighting the predictive power of various combinations of sequences.

Combined benefits of the learning analytics implemented

In addition to highlighting the holistic and statistically significant differences between high versus low-performing tutors, the combined insight from each form of learning analytics creates a deeper understanding of how tutors co-regulate learners in a VCE, with each form of analysis validating and extending the other.

Triangulation and validation of results for dataset E (where evaluator rankings are criterion for performance)

The structure of the results from process mining and sequence mining align, validating each other. For example, the complex process maps produced for high-ranked tutor sessions validate the diverse range of frequent patterns identified by sequence mining (compared with only 3 frequent patterns in sessions conducted by low-ranked tutors). The results from process mining also emphasise the role of open-ended questioning by tutors in the sessions conducted by high-ranked tutors. This aligns with the significant pattern *Tutor asks open ended, conceptual questions – learner reasons – tutor passive – learner reasons* detected through pattern mining for the high-ranked tutor dataset. Patterns featuring open-ended questioning are significantly less frequent for low-ranked tutor sessions, with less than half of the sessions conducted by low-ranked tutors demonstrating these patterns.

Similarly, the accuracy of the classifiers reflects the strength of the correlation between isolated sequences of tutor-learner engagement, and tutor performance, as well as validating the findings from process mining. With regard to dataset E (where the criteria for

tutor performance is human evaluator rankings), the highest classifier accuracy achieved is 83% (through the JRIP algorithm). The model is relatively simple, consisting of only 3 rules, and relies on two types of tutor practices; pattern A, which comprises of the tutor using praise and follow up questions to confirm and extend learner understanding, and pattern B, in which the tutor combines open ended questions with passive observation of the learner, to elicit reasoning. This model is particularly strong at detecting high-ranked tutors, with precision of 77% and recall of 93%. However, the model performs worse for low-ranked tutors, with a precision of almost 92% but recall of only 73%. The classifier accuracy can be understood in light of the results from process mining. Clear differences were discovered in the holistic process for high-ranked tutors versus low-ranked tutors. This aligns with the high levels of classification accuracy achieved. In particular, the process maps for high-ranked tutors clearly demonstrate that open ended questioning and learner reasoning are integral to the overall process for high-ranked tutors, while this is not the case for low-ranked tutors. This aligns with the role of pattern B in the classification model, with the frequency of patterns revolving around tutor open-ended questioning and learner reasoning, being key to predicting tutor performance.

With regard to low-ranked tutors, it is interesting to note that the process maps clearly describe the overall process as being tutor led and closely aligned to the structure of the platform. This aligns with the results from sequential pattern mining, which show that there are limited significant patterns of engagement, indicating that low-ranked tutors do not proactively deploy pedagogical tactics in a similar manner to high-ranked tutors.

Classification models thus detect low-ranked tutors, based on the absence of patterns of engagement found amongst high-ranked tutors. This can lead to lower levels of recall, when low-ranked tutors gain experience and training, and begin to improve their practice.

Triangulation and validation of results for dataset F (where Average Progress Scores are criterion for performance)

With regard to dataset F (which uses average progress scores as a measure of tutor performance), the process maps can be validated through the output from pattern mining. For example, the process maps emphasise the more prominent role of open-ended engagement for high APS tutors versus low APS tutors. This aligns with the s-frequent patterns detected for high APS tutors which include forms of open-ended engagement. The only s-frequent pattern of behaviour for low APS tutors is learner reasoning, while the tutor passively observes. In addition, the process map for high APS tutors shows a more distributed range of tutor practices, while low APS tutor practices are primarily focused on forms of praise and guided engagement. This aligns with the larger number of s-frequent patterns detected for high APS tutors versus low APS tutors.

The lower levels of classifier accuracy align with the nature of the process maps for dataset F. While the process maps for high-performing versus low-performing tutors differ in the degree of emphasis on open-ended engagement, the contrast between the groups of tutors is less stark (versus when dataset E was used). This is due to the relatively high level of variation in how tutors boost self-regulation, across the high-performing and low-performing groups.

Summary

As illustrated in my literature review (Section 2.4), the combined use of descriptive forms of learning analytics (such as process mining) with techniques that can be statistically validated (e.g. sequence mining), is relatively rare. Most studies have either focused on process mining on a stand-alone basis, or occasionally combined with pattern mining.
My research demonstrates the value of combining process mining, with sequence mining and decision trees. Adopting a layered approach towards learning analytics enables both holistic

and granular, statistically significant differences to be identified, between high and low-performing tutors. A richer understanding of the differences between the sessions conducted by high versus low-performing tutors, with respect to co-regulation, is developed.
In addition, the results from a specific form of learning analytics (e.g. process mining) can be used to interpret the findings from the other forms of learning analytics (e.g. sequence mining and decision trees). In addition to deepening our understanding of the stand-alone results from that specific form of learning analytics, this also enables a type of validation of the results obtained.

From an operational perspective, developing both a holistic and granular understanding of co-regulation in a VCE is valuable, as macro and micro level operational interventions can be aligned. The holistic understanding enables the high-level development of tutor training, and evaluation approaches and rubrics. On the other hand, the identification of granular, significant differences acts as the foundation for specific micro-level operational interventions.

6.4 Implications of results

As discussed above, the framework of signifiers was implemented through a selection of learning analytics to develop a rounded and granular understanding of co-regulation in a VCE. I discuss below the results generated from the implementation of each type of learning analytics for dataset E and dataset F in light of the literature, and consider the implications of my findings for learning design and operations. I start by considering the implications from the process mining results for dataset E and F, before moving onto analysing the implications of the results from pattern mining and decision trees.

Implications of process mining results for dataset E where evaluator rankings are used as the criterion for tutor performance)

The results from the implementation of process mining for dataset E extend the relevant literature, and offer new contributions. As my literature review (set out in Chapter 2), demonstrates, limited prior research has been conducted on the co-regulatory practices of human tutors. However, there is a substantial body of work on how teachers can promote SRL in classroom environments, which can be interpreted in light of my findings. Similarly, my work contributes towards interpretations of related learning theories (such as the Chi & Wylie ICAP framework) for my study context. Finally, I am not aware of any studies which discover the processes underlying sessions conducted by high and low-performing tutors, where tutor performance is defined with respect to SRL. My process mining results for dataset E offer new contributions to the literature.

My findings extend the literature with regard to the types of teacher practices which promote SRL. Previous studies have shown that while teachers can promote SRL in their classrooms SRL directly or implicitly, implicit instruction in SRL tends to be more likely (Dignath & Büttner, 2018). My results show that these findings also apply to the study context of tutors conducting sessions in a VCE, with implicit techniques such as open-ended prompting featuring prominently in the process maps rather than direct prompting of SRL.

Similarly, there is a substantial body of literature on learning theories relevant to SRL, which can be extended through my research. For example, Chi and Wylie (2014) set out the ICAP framework of cognitive engagement, defining various types of learner engagement. As part of designing the framework of signifiers, I interpreted the ICAP framework to identify tutor practices that may influence learner Operations. The implementation of process mining using the signifiers identified, demonstrates that there are key differences between high versus low-ranked tutors, with respect to the types of cognitive engagement encouraged. High-ranked tutors are more likely to use open-ended practices that encourage interactive and constructive forms of cognitive engagement by learners. On the other hand, low-ranked tutors

tend to use directive and explanatory forms of engagement, which are more likely to promote passive or active engagement by learners. In addition to relating tutor performance with respect to SRL, to the ICAP framework, my findings also extend the literature by demonstrating that the importance of considering patterns of engagement. For example, high-ranked tutors are more likely to transition from passive forms of engagement (where they diagnose gaps in learning) to guided forms of engagement (in which learners are scaffolded through narrow questions). On the other hand, low-ranked tutors tend to transition from passive forms of engagement to explanatory forms of engagement (where they attempt to plug gaps in learning through explanations).

Finally, I offer a new contribution to the literature by implementing process mining for dataset E, and examining tutor-learner interactions at a much more granular level than human observation would allow. I am not aware of any studies which have implemented process mining to discover co-regulatory processes, for tutors who are evaluated by humans as being high/low performing for SRL. Process mining enables patterns which are key to the overall structure of tutor-learner interactions to be identified, which could not be done through human observation and frequency-based forms of analysis. Prominent patterns of tutor-learner interaction are identified, which differentiate high-ranked tutors from low-ranked tutors.

High-ranked tutors appear to use a more diverse range of tutor practices, and transition fluidly between these. For example, high-ranked tutors appear to deploy a range of practices to build on from learner solving problems within the VCE, as demonstrated by the multiple transitions between learner assembling to interactive and open-ended forms of engagement. In addition, my findings also demonstrate a clear distinction between the nature of the prominent events and transitions featured in sessions conducted by high-ranked versus low-ranked tutors. The sessions conducted by low-ranked tutors tend to center on a small

number of practices, out of which most are tutor-centred and cede limited autonomy to the learner, for example, forms of directive engagement. Open ended, conceptual dialogue does not appear to be a significant feature of tutor-learner interactions. In contrast, high-ranked tutors utilise a more diverse range of practices, with open-ended engagement featuring more prominently. In addition, high versus low-ranked tutors also seem to use the same tutoring practice to different effect. For example, praise and reward are heavily embedded in the maps for both categories of tutor. However, these practices are linked to shallow learning operations in the low-ranked tutor map, and to more sophisticated learning operations, such as reasoning, for the high-ranked tutor map.

Implications of process mining results for dataset F (where average progress scores are used as the criterion for tutor performance)

My findings with regard to dataset F (where Average Progress Scores are used as the criterion for tutor performance) extend and offer new contributions to the literature. As my literature review (set out in Section 2.4) illustrates, there have been some previous studies exploring the use of process mining for high versus low-performing learners, where learner performance is defined with respect to attainment (Bannert et al., 2014). However, these studies have focused on learners working independently, and have not considered how tutors can co-regulate learners. In addition, there is no work that I am aware of which examines how high-performing tutors differ from low-performing tutors with respect to co-regulation, when the criterion for performance is academic attainment.

The results from my study are consistent with previous studies which have implemented process mining for high-performing versus low-performing learners. For example, Bannert et al. (2014) found that high-performing learners at universities adopted more complex forms of information processing than low-performing learners, when attainment was used as the criterion for evaluating performance. My research extends these

findings by demonstrating that high-performing tutors adopt a more diverse range of tutoring practices, including open ended forms of engagement which encourage more complex forms of cognitive processing. My findings also extend the literature by highlighting the importance of flexibility in the patterns of cognitive engagement encouraged. The learning behaviours encompass both shallow and more sophisticated learning operations, in sessions conducted by high APS tutors. It is likely that tutors in these sessions are more willing to adopt tutor-controlled practices as they scaffold learners through challenging learning activities. Learning outcomes are prioritised, and tutors rely on more heavily on tutor centred practices when considered necessary.

Operational implications of process mining results

The results from process mining have implications for the design of the VCE platform, as well as approaches towards tutor training and evaluation.

The structural differences in the discovered processes for the sessions conducted by high versus low-performing tutors point to a difference in the way that the VCE platform is used, by the two groups. High-performing tutors appear to use the platform as a starting point, from which tutor-learner interactions are built, with learning being extended through rich, open-ended forms of engagement, such as sophisticated dialogue. In contrast, low-performing tutors appear to be constrained by the platform, and tend to restrict their practices towards guiding the learners through the activities set out in the platform. This is supported by the findings from the tutor interviews; 75% of high-ranked tutors agreed with the statement that 'tutors were responsible themselves for the structure of their tutoring sessions', but only 40% of low-ranked tutors did, with the majority of low-ranked tutors perceiving the structure and flow of their sessions as being determined by the VCE platform. From an operational perspective, this indicates the need for scaffolding tutors on how to use the learning resources set out in the VCE. For example, the learning design could incorporate

prompts for tutors, on the types of tutoring practices that can be deployed once the learner has solved a problem set out in the VCE, such as a diagnostic question to confirm understanding or an open-ended question to prompt reasoning and conceptual thinking. Similarly, another area for further research is the embedding of real-time recommendations into the platform; for example, the learning design could be further developed to suggest intelligent pathways for tutor-learner interaction, depending on learner progress during the session.

The key patterns of engagement identified also have implications for the development of more detailed approaches and rubrics for tutor training and evaluation. Based on my results, I have suggested an extension to the rubric used by TSL for evaluating student-centred teaching. The suggested extension to the TSL rubric is presented in Table 66.

Table 66: TSL rubric for student-centred teaching, with suggested extensions based on my research findings. (Highlighted in grey)

TSL Student-centred teaching rubric	**Overlaps with co-regulatory tutoring practices**
How effectively (quality, appropriateness, necessity) did the tutor pre-assess the student to identify student needs?	Effective diagnosis enables the tutor to pitch the learning at the appropriate level, enabling a degree of learner independence within the learning process
How accurately did the tutor understand the student needs?	
How well did the tutor adapt to the student needs? - Breaking down concepts - Scaffolding e.g. using learning aids - Using different methods - Adjusting pace - Addressing gaps in knowledge - Follow up questions	Adapting to the learner needs can support deeper cognitive processing by learners, which forms part of effective self-regulation
How effectively (quality, appropriateness, necessity) did the tutor post assess the student to check for understanding?	Post-assessment enables the tutor to pitch the learning at the appropriate level, enabling a degree of learner independence within the learning process
How well did the tutor promote conceptual understanding over computational methods?	Promotion of conceptual understanding over computational methods relates to deeper cognitive processing which forms part of effective self-regulation
How well did the tutor promote pupil independence and autonomy?	Promotion of learner independence is a fundamental aspect of co-regulation
How well did the tutor use specific praise/motivation to positively impact the student mindset? (Including use of effort points)	Learner affect, well-being and mindset are recognised as key contextual factors which influence how effectively a learner can self-regulate

How well did the tutor demonstrate listening skills in the session? (Is the tutor aware of everything that happens during the session?)	Effective listening skills are relevant for appropriate diagnosis, adaptation, and personalisation of learning, all of which can influence learner engagement and likelihood of effective self-regulation.
How well did the tutor discuss real-life situations in the lesson?	Learner motivation is recognised as a key contextual factor which influences how effectively a learner can self-regulate. Discussing real-life situations can enhance the perceived value of learning and thus motivation.
How much value was the tutor able to add to the student's learning (new learning, new strategies, confidence, fluency)?	Modelling learning strategies can help a learner to effectively self-regulate
How well did the tutor use variation in tone to keep the student engaged?	Learner motivation is recognised as a key contextual factor which influences how effectively a learner can self-regulate. Discussing real-life situations can enhance the perceived value of learning and thus motivation.
How well was the tutor able to use interests and previous knowledge about the student to impact interaction?	
How well was the tutor able to extend (challenge, making links) the student's learning? (If applicable)	Extending learning where appropriate can boost learner motivation and engagement, and encourage self-regulation
How well did the tutor manage any SEN needs of the student? (If applicable)	N/A
How well did the tutor demonstrate different engagement techniques? (If applicable)	The range of engagement techniques used by the tutor can influence depths of cognitive processing, and effective self-regulation
How well was the tutor able to 'turn around' engagement and interest of the student? (If applicable)	Learner motivation is recognised as a key contextual factor which influences how effectively a learner can self-regulate.
How well did the tutor personalise learning based on the student interests? (If applicable)	Personalising learning to student interests can boost student motivation and effective self-regulation.
Suggested extensions based on my research findings	**Overlaps with co-regulation**
How often did the tutor transition from asking specific narrow-ended questions towards asking open-ended questions?	Encourages deeper cognitive processing by learners, which forms part of effective self-regulation.
How often did the tutor transition from passive forms of engagement (where they diagnose gaps in learning) to guided forms of engagement (in which learners are scaffolded through narrow questions)? How frequently did this transition occur, versus a transition from passive forms of engagement to explanatory forms of engagement?	Signals the level of learner independence and constructive engagement encouraged by tutors, after diagnosing learner gaps.

What was the range of practices employed by tutors after the learner had addressed a question in the VCE? How varied were these practices, and to what extent did they encourage constructive dialogue?	Signals whether tutors are using the VCE platform to extend learning, and encourage deeper cognitive processing by learners.
Which learning behaviours were praise and reward used to encourage? For example, were praise and reward used to encourage sophisticated learning behaviours such as reasoning, or shallow learning behaviours such as repetition?	The use of praise and reward to encourage sophisticated learning behaviours will result in deeper cognitive processing by learners, and more effective self-regulation.
How effectively were passive tutor behaviours used to encourage learner reasoning?	Encourages deeper cognitive processing by learners, which forms part of effective self-regulation.

It is interesting to note that the tutor-led patterns of behaviour observed for low-ranked tutors do not align with tutor self-reports generated through exploratory interviews; in these interviews, low-ranked tutors recognise the importance of encouraging learner reasoning and active participation. It is possible that low-ranked tutors have been trained on the concept of student-centred pedagogy, but do not fully understand how to enact this during their interactions with learners. This has implications for tutor training; for example, low-ranked tutors may potentially benefit from visualisations of tutor-learner interactions being built into the platform, to support tutor reflection and CPD.

Finally, process mining for dataset illustrates that learning behaviours encompass both shallow and more sophisticated learning operations, in sessions conducted by high APS tutors. It is likely that tutors in these sessions are more willing to adopt tutor-controlled practices as they scaffold learners through challenging learning activities. Any development of learning design and approaches towards tutor training/evaluation will need to take account of this flexibility in patterns of engagement.

Summary of implications of findings from process mining

In this section, I have considered the implications of my results from process mining for dataset E (where evaluator rankings are the criterion for tutor performance), and dataset F

(where average progress scores are the criterion for tutor performance). I have shown how my findings interpret the current body of literature for my study context, with regard to the teaching practices that promote SRL and influence different forms of cognitive processing. In addition, I have analysed how my process mining results offer a new contribution to the field, by highlighting structural and granular aspects of co-regulatory processes in a VCE. The key structural differences between high versus low-performing tutors are highlighted, and the differences in prominent events and transitions are noted. The impact of using learner attainment as a performance criterion rather than evaluator rankings is also explored. For example, the use of attainment as a performance criterion shows that high-performing tutors use more varied patterns of engagement, than if evaluator rankings are used as performance criterion. This is potentially due to high-performing tutors (with respect to attainment) shifting to more directive styles of engagement as they scaffold the learners through challenging content. Finally, the potential operational implications of my findings are considered, including the development of tutor training approaches and rubrics, and opportunities for intelligent recommendations to scaffold tutors.

Following the discussion of the implications of my findings from process mining, I will move on to discuss the implications of the findings generated from pattern mining and decision trees.

Implication of results from pattern mining and classifiers, for dataset E (where evaluator rankings are criterion for tutor performance)

The implementation of pattern mining and decision trees for dataset E builds on findings from previous studies, such as Cukurova et al. (2022). As described in my literature review, in this study, statistically significant patterns were identified for high-performing and low-performing tutors, and used to build classifiers. Tutor performance was determined through a combination of factors including evaluator rankings. The study found that significant patterns

of tutor behaviour could be used to predict tutor performance with relatively high levels of accuracy. Examples of key patterns for high-performing tutors include patterns in which they monitored learners as they reasoned or worked, with longer patterns of tutor behaviour being more significant. My findings align with this study, in terms of the types of patterns identified. For example, patterns in which the tutor passively observes the learner while they problem-solve or reason are much more likely to occur for high-ranked tutors. In addition, my research also aligns with the previous findings that the length of significant patterns tends to be longer for high-performing tutors versus low-performing tutors. For example, *'Tutor asks open ended conceptual questions–learner reasons–tutor passive'* is not statistically significant, but the longer pattern *'Tutor asks open ended conceptual questions–learner reasons–tutor passive–learner reasons'.*

My research extends the findings from previous studies in multiple ways. Firstly, tutor-learner patterns of interaction are considered, while previous studies have only considered tutor behaviours in isolation. Considering tutor-learner interactions in the round opens up new avenues for research. For example, the significance of the longer pattern *'Tutor asks open ended conceptual questions–learner reasons–tutor passive–learner reasons'* indicates high-ranked tutors tend to elicit more extensive reasoning than low-ranked tutors. Further research is required to explore why high-ranked tutors elicit more extensive reasoning. For example, high-ranked tutors may be able to elicit a greater degree of reasoning due to the to appropriate pitching of the open-ended questions, both with regards to timing and framing of question, or due to appropriate pausing/monitoring to encourage further reasoning prior to moving onto the next activity.

My findings also extend the literature by identifying new significant patterns of engagement. For example, my results show that use of the '*Learner writes answer– tutor praises learner–tutor asks learner a specific question'* pattern distinguishes high-ranked

tutors from low-ranked tutors. The pattern has a low mean for low-ranked tutors, particularly when compared to high-ranked tutors, and a relatively low standard deviation. The pattern appears to be infrequently demonstrated by low-ranked tutors, and when it is used, the episode support is limited. These results indicate that the patterns of tutor-learner engagement, after the learner has attempted to solve a problem, are important. The significance of this pattern also means that it can be used to help predict tutor performance, via the classifiers implemented, with an acceptable level of classifier accuracy of 83% being achieved.

Implication of results from pattern mining and decision trees, for dataset F (where average progress scores are criterion for tutor performance)

My findings for dataset F contribute to the literature by identifying significant co-regulatory patterns of engagement, and analysing the extent to which they can be used to predict tutor performance with respect to attainment. While the relationship between SRL and learning outcomes is well established (e.g. van de Graff, 2022), I am not aware of any research that has explored whether statistically significant patterns of co-regulation can be used to predict tutor performance with respect to attainment.

My results highlight a small number of significant patterns which distinguish between high versus low-performing tutors. However, the standard deviation for the mean episode support (i.e. frequency of patterns within sessions) is very high. This indicates that sessions conducted by high-performing tutors vary in the nature of tutor-learner interactions with respect to SRL. The only pattern with a considerably higher mean in the high APS group, and relatively low standard deviation is *Tutor asks learner a specific question – tutor asks open ended, conceptual questions – learner reasons.* This pattern suggests that high-performing tutors will scaffold learners through micro-questions, possibly diagnosing learner domain knowledge, before encouraging the learner to engage in open-ended dialogue. These

findings align with an established body of research on the Zone of Proximal Development (Vygotsky, 1978), which suggests that content/instruction needs to be pitched at the appropriate level of challenge for learners. High-performing tutors, with respect to learning outcomes, seem to be careful to calibrate their style of engagement to learner needs, adjusting the sophistication of the engagement as needed.

A relatively low level of classifier accuracy is achieved for dataset F. The highest classifier accuracy achieved when average progress scores are used as the performance measure is only 68%. This is most likely because the identified patterns of tutor-learner interaction are based on a framework of signifiers of co-regulation, which is closely related to the construct of student-centred teaching. While there is a substantial body of research on the relationship between learning outcomes and SRL (e.g. van de Graff, 2022), limited research has explored the relationship between co-regulatory teaching practices and learning outcomes. It is possible that high APS tutors implement tutor-led (rather than student centred) practices to scaffold learners and boost learning outcomes. This may result in greater variation in the patterns of engagement with respect to SRL, reducing the accuracy of the classifiers.

Operational implications of results from pattern mining and decision trees

The nature of the significant patterns identified illustrate that key differences tend to emerge between high versus low-performing tutors after the learner has attempted a question, or when the tutor embarks on a course of questioning. The learning design of the VCE platform could therefore be enhanced to provide scaffolding to tutors at these specific points within the overall problem-solving activity.

The level of classifier accuracy achieved has implications for tutor training and evaluation processes occurring outside of the VCE platform. For example, the acceptable classification accuracy achieved where evaluator rankings are used as the performance

measure, indicates that there is potential to inform the evaluation of SCT by tracking patterns of interaction. The classifiers highlighted two patterns as being most effective in predicting tutor rankings: the patterns identified were '*Learner writes answer – tutor praises learner – tutor asks learner a specific question',* and *'Tutor asks open-ended question – learner reasons – tutor passive – learner reasons.* The absence of these patterns in tutor sessions over a period of time could potentially indicate that a tutor requires support and training. Further research is required to establish the significance of the highlighted patterns in a scaled dataset and validate their predictive capability, before they can be used to potentially inform tutoring training processes.

In addition to identifying struggling tutors, predictive patterns of engagement could also potentially be used to track tutors who have improved their practice and require less support than previously. For example, while the overall classifier accuracy is relatively high, recall is slightly lower for low-ranked tutors versus high-ranked tutors, indicating that a minority of low-ranked tutors are using sophisticated practices to extend learner understanding more frequently than expected. These results could indicate that some low-ranked tutors have improved their practice following their last formal evaluation and ranking. As described above and in Section 5.3, a lower classifier accuracy was achieved for dataset F (where the tutor performance is measured by average progress scores). The *Tutor asks learner a specific question – tutor asks open ended, conceptual questions – learner reasons* pattern was selected by classifiers to predict the average progress scores of tutors. However, the use of this pattern to classify tutors into high and low-performing groups is more effective for low APS tutors versus high APS tutors, as demonstrated by the higher recall rates for low APS tutors. Most low APS tutors do not demonstrate the pattern, therefore being detected by the model. However, there is also a significant minority of high APS tutors which also do not demonstrate this pattern, and these tutors are misclassified as being low APS. As previously

discussed, high APS tutors may use a mix of tutor-led and student-centred practices, as they scaffold the learner through challenging content. It is therefore challenging to accurately detect high APS tutors using a framework of signifiers for co-regulation.

Summary of implications of findings from implementation of pattern mining and decision trees

In this section, I have analysed the implications of my results from pattern mining and decision trees. I have shown how the existing body of literature can be interpreted in light of my findings, and highlighted where I have made new contributions to the literature.

For example, my findings for dataset E (where evaluator rankings are criterion for performance) are consistent with findings from previous studies, such as Cukurova et al. (2022). The results emphasise the importance of tutoring practices such as monitoring learners, as they reason or problem solve, as well as the significance of the length of the pattern. My findings extend the literature by considering patterns of tutor-learner interaction, rather than solely focusing on learner behaviours. My findings also identify significant patterns such as '*Learner writes answer – tutor praises learner – tutor asks learner a specific question'* These patterns of interaction can be used to build classifiers with acceptable levels of accuracy.

My findings for dataset F contribute to the literature by identifying significant co-regulatory engagement patterns and analysing the extent to which they can be used to predict tutor performance for attainment. While the relationship between SRL and learning outcomes was well-established (e.g. van de Graff, 2022), I am unaware of any research that has explored statistically significant patterns of co-regulation that can be used to predict tutor performance for attainment. My research has found that a small number of statistically significant patterns can be identified for high-performing versus low-performing tutors, in dataset F. However, the significant patterns exhibit high levels of standard deviation in terms

of their frequency in sessions conducted by high-performing tutors. This is likely because high-performing tutors calibrate their level of engagement to the learner's needs, as they scaffold the learner through challenging content. This means that significant co-regulatory patterns cannot be used to predict whether tutors are high-performing or low-performing (with respect to learner attainment) with high levels of accuracy.

The results from pattern mining and decision trees also have a number of operational implications for micro-level interventions which can scaffold tutors at specific points within the session. For example, the nature of the patterns identified indicates that key differences tend to emerge between high versus low-performing tutors after the learner has attempted a question, or when the tutor embarks on a course of questioning. Therefore, these are key points within the session where scaffolding could be introduced into the learning design (e.g., through intelligent prompts). In addition, the reasonable classifier accuracy (83%) for dataset E suggests that there may be potential to monitor how effectively tutors co-regulate learners through machine learning techniques. On the other hand, the lower levels of classifier accuracy for dataset F indicate that tutor performance (with respect to learner attainment) cannot be effectively monitored using co-regulatory patterns of engagement.

Conclusion

My research has aimed to explore how tutors co-regulate learners in a VCE, and specifically the differences between high and low-performing tutors. In this chapter, I have reflected on my approach towards developing a theoretical foundation for my work and implementing frameworks through learning analytics. I have also considered the implications of my results, both with regard to the existing literature and from an operational perspective.

In order to develop a theoretical foundation for this work, I have contextualised the Hadwin and Winne (2001) model of SRL. The architecture of the Winne and Hadwin model, specifically its information-processing orientation, make it better suited to an unstructured

study context, versus a model which is structured by phase of SRL. Further, its holistic nature works well, given the under-researched nature of the study context, as it enables the identification of the most relevant signifiers. The generalizable nature of the model does mean that a significant amount of research is required to develop the model to a sufficient degree of granularity, but this is somewhat inevitable.

The frameworks were designed using a mixed methods approach, which enable a theoretically grounded and practical representation of co-regulation within a VCE. The framework is designed to include two levels of granularity: categories and observable signifiers. The categories map to theoretical constructs within the contextualised model, while the signifiers are the observable actions taken, which map to the categories. In addition, the mixed methods approach allows nuanced events to be captured within the framework (e.g. with respect to the ways in which tutors can influence cognitive engagement), and key differences between processes to therefore be accurately detected. Natural language dialogue is used to identify signifiers and has proven to be a rich source of data which informs our understanding of how co-regulation unfolds in practice in a VCE.

In order to implement the framework of signifiers, a layered approach towards analytics is adopted. Different forms of analytics were implemented to fully address Research Question 3, by identifying both holistic and statistically significant differences in how tutors co-regulate learners in a VCE. Process mining is fundamental to my analysis, by generating holistic visualisations which align to the contextualised model of SRL (set out in section 4.1), and act as the basis for a rich exploration of how co-regulation occurs in practice. However, the descriptive nature of process mining means that the results cannot be statistically validated. In order to identify statistically meaningful, granular patterns of interactions, additional layers of analysis, comprising of sequence mining and decision trees, are implemented. This approach supports a macro and micro level understanding of co-regulation

in a VCE, which can be used both for the high-level development of evaluation rubrics, and specific micro-level operational interventions. Further, the combined insight from each form of learning analytics creates a deeper understanding of how tutors co-regulate learners in a VCE, with each form of analysis validating and extending the other.

I moved on to discuss the implications of my findings from the implementation of learning analytics, both from the perspective of the existing literature, and from an operational perspective. With regards to process mining, my findings interpret the current body of literature for my study context, with regard to the teaching practices that promote SRL and influence different forms of cognitive processing. In addition, I have analysed how my process mining results offer a new contribution to the field, by highlighting structural and granular aspects of co-regulatory processes in a VCE. The impact of using learner attainment as a performance criterion rather than evaluator rankings is also explored. For example, the use of attainment as a performance criterion shows that high-performing tutors use more varied patterns of engagement, than if evaluator rankings are used as performance criterion. This is potentially due to high-performing tutors (with respect to attainment) shifting to more directive styles of engagement as they scaffold the learners through challenging content.

With regard to pattern mining and decision trees, I have again analysed the implications of my results for the existing literature, and highlighted where I have made new contributions. For example, my findings for dataset E (where evaluator rankings are criterion for performance) are consistent with findings from previous studies, such as Cukurova et al. (2022). The results emphasise the importance of tutoring practices such as monitoring learners, as they reason or problem solve, as well as the significance of the length of the pattern. My findings extend the literature by considering patterns of tutor-learner interaction, rather than solely focusing on learner behaviours. My findings also identify significant patterns such as '*Learner writes answer–tutor praises learner–tutor asks learner a specific*

question' These patterns of interaction can be used to build classifiers with acceptable levels of accuracy. My findings for dataset F (where average progress scores are the criterion for tutor performance) contribute to the literature by identifying significant co-regulatory patterns of engagement, and analysing the extent to which they can be used to predict tutor performance with respect to attainment. While the relationship between SRL and learning outcomes is well established (e.g. van de Graff, 2022), I am not aware of any research that has explored whether statistically significant patterns of co-regulation can be used to predict tutor performance with respect to attainment. My research finds that a small number of statistically significant patterns can be identified for high-performing versus low-performing tutors, in dataset F. However, low levels of classifier accuracy suggest that significant co-regulatory patterns cannot be used to predict whether tutors are high-performing or low-performing (with respect to learner attainment).

Finally, I have considered the operational implications of my results. The differences in the structure and key features of process maps indicate that process mining could be used to support the development of tutor training approaches and rubrics, as well as micro-level interventions. The results from pattern mining and decision trees also have a number of operational implications for the point within a tutoring session at which scaffolding may be introduced (e.g. after the learner has attempted a question, or when the tutor embarks on a course of questioning). In addition, reasonable levels of classifier accuracy (83%) for dataset E suggest that there may be potential to monitor how effectively tutors co-regulate learners through machine learning techniques. On the other hand, the lower levels of classifier accuracy for dataset F indicate that it would be challenging to monitor tutor performance (with respect to learner attainment) using co-regulatory patterns of engagement.

6.5 Limitations

In this section, I discuss limitations of my research.

The research is limited by the sample size used for the implementation of learning analytics. I have collated data from the highest performing and lowest performing tutors only in order to identify distinguishing patterns of behaviour (as outlined in chapter 3). Data has been collected from 30 tutors for dataset E, and from 20 tutors for dataset F, with at least two sessions being analysed per tutor. While this is not an extensive sample of tutors, it aligns with my aim of understanding the differences in how high versus low-performing tutors influence learner SRL. The size of this dataset is also comparable with published studies focusing on teaching activities. For example, the sample size of Cukurova et al. (2022) was 44 teachers; Prieto et al. (2018) was two teachers in 12 classroom sections; Camacho et al. (2020) was three teachers in 18 learning activities; Prieto et al. (2016) was four teachers in four sessions of 35 to 40 minutes; Riel et al. (2018) was 42 teachers. Future research can confirm my findings through larger sample sizes, in addition to exploring patterns of behaviour for tutors with mid-range performance.

Secondly, this research was intended to identify potential signifiers of SRL in a specific tutoring context (primary online mathematics), and explore whether these can be used to predict tutor performance. Further research is required to explore whether the approach and findings could be applicable in other subject domains. It is possible that that significant patterns of tutor-learner interaction could be identified in similar problem-solving domains (e.g., physics), but less likely in subject domains with less structured dialogue (such as arts and humanities).

Thirdly, this research was conducted in partnership with an industrial supplier of online tutor training (Third Space Learning). The tutors engaged by Third Space Learning are professional tutors, who are centrally trained using company rubrics and training approaches that emphasise the importance of student-centred teaching. It is possible that these findings

would not apply to sessions conducted by online tutors, who are not trained with similar pedagogical approaches.

Fourthly, my research relied on manual coding of the natural language dialogue between tutors and learners. This approach is limited in scale, and further research is required whether LLMs will be able to reliably label signifiers derived from natural language dialogue.

Fifthly, the results from this work are applicable for the study of maths problems in a particular VCE; it is possible that the study of other subjects (for example, English) would be associated with entirely different co-regulatory processes. In addition, the nature and design of the specific VCE used in this research study may impact the types of co-regulatory processes which were highlighted. For example, the design of the VCE did not incorporate any video of tutors or learners. It is possible that video would have resulted in different forms of co-regulation becoming more prominent e.g. the video cues may have resulted in a greater significance of tutor practices which prompted learner monitoring. Another factor which may have influenced my findings is the fact that learners may not retain the same tutor over the multiple VCE sessions. This factor could have influenced the significance of various co-regulatory practices which encouraged different levels of cognitive processing.

Further, the overarching approach that I adopted towards my research is not grounded in an explicit epistemological or ontological perspective. This can result in narrow interpretations and perspectives. For example, future research could pay more attention to the relationship between SRL and the real-life learning contexts of children, and consider the implications of these relationships for my findings.

Finally, the findings from process mining, sequential pattern mining and classification models can be further validated through triangulation with learner and teacher feedback. I have attempted to validate the findings through tutor interviews. However, organisational

constraints have meant that the sample size for tutor interviews is relatively small. In addition, organisational constraints have meant that learner interviews and school teacher interviews have been out of scope. Future work can gain a richer understanding of the influence of tutors on learner SRL, through a broader set of qualitative and quantitative interviews with tutors and school teachers of the relevant pupils. While learner self-reports may also be an interesting avenue to explore, there is likely to be some noise in this data given the young age of the learners.

Chapter 7 Conclusions

Introduction

The aim of this chapter is to summarise the research set out in this thesis and draw some conclusions, highlighting key contributions.

I will start by recapping the broader aims and objectives, and the key research questions which guided my work. I will then move on to summarise the research conducted and conclusions drawn for each research question. This is followed by a statement of the key contributions made, recognising the limitations on my work. Finally, I will suggest some directions for future research.

7.1 Recap of aims of research

Limited research has previously been conducted to explore how co-regulation unfolds in practice, specifically in the context of online learning environments. The proliferation of online tutoring has created new opportunities to monitor how tutors co-regulate learners from a granular, process-oriented perspective. Monitoring effective versus maladaptive patterns of tutor-learner interaction has multiple implications, both from a research and operational perspective. From a research perspective, it can enable a richer understanding of how theoretical models apply to authentic learning situations. From an operational perspective, monitoring effective tutor-learner interactions can potentially be the foundation for embedded tutor scaffolds and personalised training. This would improve the quality of delivery of tutoring sessions, and therefore learning outcomes.

My research addresses the gap in the literature by exploring tutor co-regulation of learners in a virtual classroom environment, benefitting from the associated rich sources of data. Specifically, I propose a granular, theoretically grounded framework which is implemented through learning analytics, to identify key patterns of tutor-learner interaction with respect to co-regulation.

There were three main objectives guiding my research. Firstly, I aimed to develop theoretically grounded frameworks which can effectively track co-regulatory tutoring practices in a virtual classroom environment. Secondly, I aimed to explore the effectiveness of process mining, pattern mining and decision trees in building a sophisticated understanding of co-regulation in a VCE, both on a stand-alone and combined basis. Finally, I aimed to identify the key pedagogical differences between high versus low-performing tutors, with respect to co-regulation in a VCE.

My work was shaped by the following overarching research question: *How can we use learning analytics to detect the influence of human tutoring practices on learner SRL processes, for high and low-performing tutors in a VCE?*

The overarching RQ was mapped to sub questions to address the set of aims outlined above. The goals and contributions for each question are set out in Figure 44.

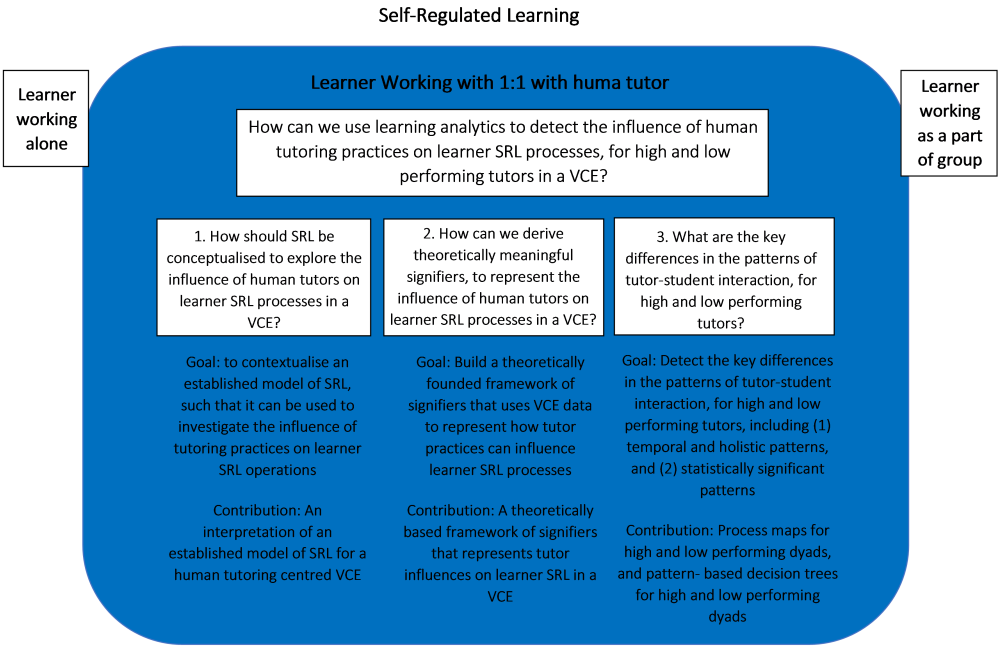


Figure 44. Research questions mapped to goals and contributions

In the following sections, I will summarise my findings for each research question, and draw some conclusions.

7.2 Research summary and conclusions

RQ: How should SRL be conceptualised to explore the influence of human tutors on learner SRL processes in a VCE?

Goal: To develop a contextualised model of SRL, which can be used to investigate the influence of human tutoring practices on learner SRL in a VCE

I have developed a contextualised model of SRL in Section 4.1 of this thesis, and it is set out in Figure 44.

Contribution: An interpretation of an established model of SRL for a human tutoring centred VCE

I have contextualised the Winne and Hadwin (1998) model of SRL to the study context of exploring how tutors co-regulate learners in a Virtual Classroom Environment. I have contextualised this model through a literature review and exploratory data analysis.
I have adopted a dynamic, process-oriented conceptualisation of SRL to explore the influence of human tutors on learner SRL. This conceptualisation aligns with the evolution of SRL over the course of a learning activity. Further, it enables an exploration of the shaping influence of contextual factors (such as tutor cues) on learner SRL processes.

The Winne and Hadwin (1998) model of SRL has been adopted for contextualisation, as the architecture of the model has been assessed as being well-suited to my study context (see section 4.1). The model sets out an information processing architecture, which considers the transition of information between macro components (COPES: Conditions, Operations, Products, Evaluation and Standards). This focus on information processing, rather than SRL phase, is well suited to an unstructured context such as tutor-learner interactions. Further, the holistic nature of the model provides a comprehensive foundation in an under-researched context, allowing key tutoring practices to be detected.

An extensive literature review and exploratory data analysis are used to develop the contextualised model. The resulting model enables the influence of tutoring practices on learner operations to be tracked in a holistic manner, which accounts for the indirect influence of other related factors in the model, such as the influence of tutoring practices on learners' cognitive state, and evaluations of their own learning.

The architecture of the model enables SRL as a process to be traced. For example, the interconnectivity between tutor practices and various components of the model can be evaluated. My research findings demonstrate that directive practices appear to encourage shallower operations, versus open ended practices, which encourage more sophisticated learner operations (e.g., reasoning). The model indicates that tutoring practices will also have an impact on internal conditions, such as learner mindset.

The contextualised model supports a holistic and dynamic conceptualisation of SRL. It is therefore a suitable foundation for investigating the influence of human tutors on learner SRL in a VCE. However, my research has identified a number of areas for further theoretical development. For example, empirical research could test whether certain variables, such as tutor evaluations, should be included in the model, and the relationship between these, and tutor mindset/practices. Another area for theoretical development is considering tutor self-regulation, and the relationship between tutor SRL and learner SRL. Addressing these theoretical questions would help build a more tailored model to represent how tutors co-regulate learners.

The model could also be further developed with regard to time and order assumptions underpinning the conceptualisation of SRL as a process. For example, what is the influence of time on the evolution of co-regulatory tutor practices, both within a learning activity and across learning activities? Similarly, is a certain order of events and transitions more likely within particular classes of tutors and learners? I have partly addressed this, in my

contribution to RQ3, where I have considered the different patterns of engagement for high vs low-performing tutors. Further research is needed to help develop contextualised models of SRL with regard to order and time, based on empirical results.

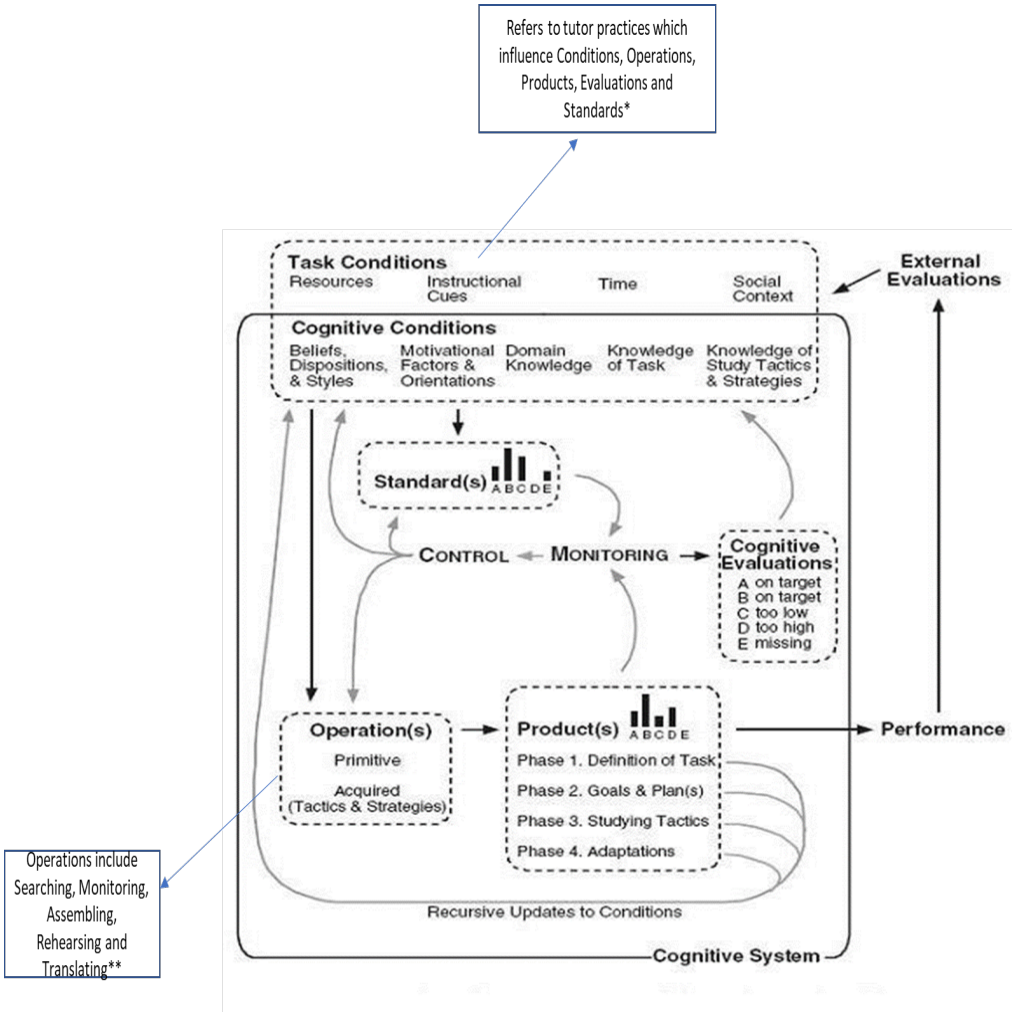


Figure 45. Contextualised Winne and Hadwin model of SRL, for exploring how tutors co-regulate learner operations in a VCE. Adapted from "Studying as Self-Regulated Learning," by P. H. Winne and A. F. Hadwin, 1998, in D. J. Hacker, J. Dunlosky, and A. C. Graesser (Eds.), *Metacognition in Educational Theory and Practice* Routledge (https://doi.org/10.4324/9781410602350-19).

*The full list of tutor practices which can theoretically co-regulate learners in a VCE is set out in Table 22.

** The definition of the SMART operations with reference to a VCE is set out in table 23.

RQ2: How can we derive theoretically meaningful signifiers to represent the influence of human tutors on learner SRL processes in a VCE?

Goal: Build a theoretically grounded framework of signifiers, which uses VCE data to represent how tutor practices can influence learner SRL processes

I have built a theoretically grounded framework of signifiers for tutor practices and learner operations. The framework of signifiers is set out in Section 4.2 of this thesis, and is presented in Tables 29 and 31.

Contribution: A theoretically grounded framework of signifiers, which uses VCE data to represent how tutor practices can influence learner SRL processes

I designed a framework of signifiers for tutor co-regulatory practices and learner SRL using a mixed methods approach. The framework of signifiers identifies data proxies using VCE data, including natural language dialogue, to provide a theoretically grounded and practical representation of co-regulation in a VCE.

The frameworks were built using a mixed methods approach (Cukurova et al., 2016), which suited my research aims of developing frameworks which are both theoretically grounded and practical. The value of the mixed methods approach was demonstrated through the development of signifiers relating to cognitive engagement. The literature review stage identified that tutors can prompt various forms of cognitive engagement; for example, tutors can adopt tutor-led practices, which encourage the learner to adopt a passive state. However, it was the review of empirical data which refined my understanding of 'tutor-led' engagement in the context of a VCE. In particular, the empirical review indicated that 'tutor-led' engagement was a broad category which encompassed situations in which the tutor was directing learners to take a specific action to solve a problem, as well as situations in which the tutor was explaining a concept to the learner to scaffold understanding. In order to more

accurately represent the various forms of tutor-learner engagement, the 'tutor-led' engagement category was replaced by 'directive engagement', and 'explanatory engagement'. This distinction was later shown as being important, when detecting distinguishing patterns of tutor-learner interaction, for high versus low-performing tutors. While both classes of tutors rely on explanatory engagement in various manners, low-performing tutors demonstrate a greater reliance on directive engagement.

The framework of signifiers relied on natural language dialogue, as well as the tutor-learner interaction observed in the VCE (e.g., use of VCE tools, and interactive whiteboard). The use of natural unstructured dialogue between learner and tutor to signify co-regulation is under-researched. While the interactions between learners and pedagogical agents have been subject to some research, the use of human-to-human dialogue can enable analysis that can later support these interactions. Further, natural language dialogue can be analysed from the perspective of understanding tutoring practices that support and inhibit learner SRL, rather than primarily focusing on the development of learner SRL skills (as is the case with most studies on ITS).

My findings have demonstrated the value of analysing natural dialogue between tutor-learners to understand how co-regulation occurs in practice. This thesis goes beyond previous research which analyses natural dialogue, as it considers the dialogue as a whole rather than specific aspects of it (e.g., use of open-ended questions, see Hrastinski et al., 2021). Previous research focused on specific aspects of dialogue (e.g., use of certain key phrases), due the ease of automating these features. However, my work takes a broader approach towards analysing natural dialogue, with the aim of identifying significant features with respect to co-regulation. By adopting a broad approach towards the analysis of natural dialogue, my work highlights the differing use of a range of features, such as directives, explanatory statements and question types. The holistic structure and specific patterns of natural dialogue distinguish

high-performing tutors from low-performing tutors (when performance is measured with respect to SRL).

The use of natural dialogue is particularly useful in the context of online human led tutoring, as natural dialogue is a key pedagogical tool in this environment. In my research, I combined the use of natural dialogue with other data sources, such as learner activity on the interactive whiteboard (e.g., learner writing on the whiteboard and erasing). This enables a holistic and contextualised analysis of co-regulatory tutoring practices. For example, 'learner erasing' is a signifier that the learner has monitored information and chosen to correct it. Combining these signifiers with natural dialogue signifiers enables exploration of the types of tutor communication that encourage monitoring.

While the application of a mixed methods approach to natural dialogue resulted in frameworks that can highlight key co-regulatory practices, the approach is limited in scale. Significant work was required to identify signifiers that could be reliably coded, and the manual coding of signifiers was a labour/time intensive process. In addition, the ability to identify signifiers that could be reliably coded was partly due to the central training received by tutors, which created similarities in the types of dialogue adopted. Despite these limitations, this research has merit in identifying key co-regulatory patterns using natural dialogue. Future research can explore the opportunities to scale this work, potentially using foundational LLMs to automate the detection of signifiers.

RQ3: What are the key differences in the patterns of tutor-student interaction, for high and low-performing tutors?

Goal: Detect the key patterns of interaction for high and low-performing tutors, including (1) holistic and (2) statistically significant patterns

I have identified holistic differences between high and low-performing tutors through process mining, and these are set out in Section 5.1. I have identified statistically significant differences between high and low-performing tutors using pattern mining and decision trees. These differences are set out in Section 5.2.

Contribution: Analysis of process maps, statistically significant pattern and pattern-based decision trees for high and low-performing tutors

I have identified differences in the structure of the co-regulatory processes adopted by high versus low-performing tutors, as well as differences in the prominent events and transitions within the process. Through pattern mining and decision trees, I have deepened my understanding of the differences between high and low-performing tutors by identifying significant isolated patterns that could potentially be used as the basis for interventions, and predicting tutor performance. With regard to the criterion for performance criteria for tutors, I have experimented with using average progress scores, as well as evaluator rankings for student-centred teaching.

Key holistic differences between high versus low-performing tutors

The use of process mining has discovered distinct latent processes underlying the sessions conducted by high-ranked versus low-ranked tutors. These differences include (1) the **range and patterns** of tutor-learner interactions and, (2) the **varied use of the VCE** by the groups of tutors with differently ranked performance.

With regard to the patterns of tutor-learner interactions, the sessions conducted by low-ranked tutors tend to center on a small number of practices, out of which most are tutor-centred and cede limited autonomy to the learner. Low-ranked tutors repeatedly engage in tutor-controlled paths of interaction with learners, with forms of directive engagement being particularly prominent within the overall process. Open-ended, conceptual dialogue does not appear to be a significant feature of tutor-learner interactions. In contrast, high-ranked tutors

utilise a more diverse range of practices, with open-ended engagement featuring prominently. High-ranked tutors appear to deploy a range of practices to build on from learner-solving problems within the VCE, as demonstrated by the multiple transitions between learner assembling and interactive and open-ended forms of engagement.

High-ranked versus low-ranked tutors also seem to use the same tutoring practice to different effects. For example, both high-ranked and low-ranked tutors utilise the practice of observing learners ('tutor passive engagement') to diagnose gaps in learning. However, in the case of low-ranked tutors, this practice most often leads to tutor explanations to plug the gaps in learning. In the case of high-ranked tutors, a more interactive form of engagement (micro questions) is used to plug diagnosed gaps. Similarly, praise and reward are heavily embedded in the maps for both categories of tutor. However, these practices are linked to shallow learning operations in the low-ranked tutor map, and to more sophisticated learning operations, such as reasoning, for the high-ranked tutor map.

While exploring the key differences between tutor classes, I have also considered the impact of changing the measure of tutor performance on the results. My results demonstrate that there are clearer differences with respect to co-regulation when human evaluator rankings for SCT are used, versus average progress scores (APS) as measures of tutor performance. For example, when average progress scores are used as the performance criteria, the main differences relate to the range of tutoring practices relied upon. While open-ended engagement features more prominently in sessions conducted by high-performing tutors, the distinction between high and low-performing sessions is less than the distinction discovered when tutor rankings are used. Similarly, the learning behaviours encompass both shallow and more sophisticated learning operations, in sessions conducted by high-performing tutors, when performance is measured via learning outcomes. It is possible that

tutors in these sessions are more willing to adopt tutor-controlled practices as they scaffold learners through challenging learning activities.

Key statistically significant differences between high versus low-performing tutors

A number of patterns are identified which are significantly more likely to occur in sessions conducted by high-performing tutors, versus low-performing tutors, as set out in Section 5.2.

High-ranked tutors are significantly more likely to more likely to use questioning in diverse ways; open ended questions are used to elicit reasoning, and specific questions are used to scaffold learners or diagnose understanding. Using average progress scores as the tutor performance measure produces similar results. High APS tutors are significantly more likely to employ diverse patterns, such as building open-ended questioning on top of scaffolding through micro questions. Similarly, patterns in which the tutor passively observes the learner while they problem-solve, or reason are much more likely to occur for high-ranked tutors.

In contrast, there are no patterns which are significantly more likely to occur in sessions conducted by low-performing tutors (whether performance is measured by tutor ranking or APS. In general, low-performing tutors appear to show more limited commonalities in tutor-learner interactions. For example, the sessions conducted by high-ranked tutors demonstrate 16 s-frequent patterns, while sessions conducted by low-ranked tutors only demonstrate 3 s-frequent patterns. Similarly, high APS tutors demonstrate shared patterns of behaviour, with 6 patterns being s-frequent, versus only 1 s-frequent pattern being demonstrated by low APS tutors.

Differences in patterns of interaction can be more reliably captured across tutor-learner sessions, rather than within a session. While the episode support for all the patterns skews higher for the high-performing tutor group versus the low-performing tutor group, the

standard deviations are also very high. There is only one pattern in each dataset with a higher mean and relatively low standard deviation:

- For dataset E, the use of the '*Learner writes answer – tutor praises learner – tutor asks learner a specific question'* pattern appears to distinguish high-ranked tutors from low-ranked tutors. The pattern has a low mean for low-ranked tutors, particularly when compared to high-ranked tutors. It also has a relatively low standard deviation. Considering both the s-support and the i-support of the pattern, it appears to be infrequently used by low-ranked tutors, and when it is used, the episode support is limited
- For dataset F, the *Tutor asks learner a specific question – tutor asks open ended, conceptual questions – learner reasons* pattern is the only pattern with a considerably higher mean in the high APS group, and relatively low standard deviation.

The significant patterns identified for dataset E can act as the basis for classifying tutors as high ranking versus low ranking, with an acceptable level of accuracy. Classifiers can predict whether tutors are high or low ranking, with an accuracy of 83%.

In contrast, the significant patterns identified for dataset F can only classify tutors as high versus low APS with a classifier accuracy of 68%. *Tutor asks learner a specific question – tutor asks open ended, conceptual questions – learner reasons* pattern was selected by classifiers to predict the average progress scores of tutors. However, the use of this pattern to classify tutors into high and low-performing groups is **more effective for low APS tutors versus high APS tutors**, as demonstrated by the higher recall rates for low APS tutors. Most low APS tutors do not demonstrate the pattern, therefore being detected by the model. However, there is also a significant minority of high APS tutors which also do not demonstrate this pattern, and these tutors are misclassified as being low APS.

7.3 Key contributions

My research has resulted in a number of contributions to the field. Firstly, I have contributed to the field by proposing a contextualised model of SRL. Secondly, I have contributed to the field by designing a framework of signifiers which represents how tutors co-regulate learners in a VCE, utilising VCE data including natural language dialogue. Thirdly, I have contributed to the field by exploring the value of combining descriptive and statistical forms of learning analytics. Finally, I have contributed to the field by identifying key holistic and granular differences between the co-regulatory practices adopted by high versus low-performing tutors.

I have contextualised the Winne and Hadwin (1998) model of SRL to the study context of exploring how tutors co-regulate learners in a Virtual Classroom Environment. I have contextualised this model through mapping relevant literature to components of the model to identify the various ways in which tutors practices can influence learner SRL. In addition, I have refined the contextualised model of SRL through exploratory data analysis, to ensure that the representation of how tutor practices can influence learner SRL is aligned to the practical context.

I have designed a framework of signifiers using a mixed methods approach, to ensure that the framework is both theoretically grounded and practically relevant. I have utilised VCE data within the framework of signifiers, including natural language dialogue. The resulting framework demonstrates the value of analysing natural dialogue between tutor-learners to understand how co-regulation occurs in practice. This thesis goes beyond previous research which analyses natural dialogue, as it considers the dialogue as a whole rather than specific aspects of it (e.g., use of open-ended questions [Hrastinski et al., 2021]). By adopting a broad approach towards the analysis of natural dialogue, my work highlights the differing use of a range of features, such as directives, explanatory statements and question types, for high versus low-performing tutors.

My research contributes towards the field through exploring the value of combining descriptive and statistical forms of learning analytics to build a rich understanding of co-regulation. Process mining (a descriptive form of learning analytics) discovers the latent processes underlying sessions conducted by high-performing versus low-performing tutors and highlights structural differences such as the complexity of tutoring practices, and prominent events/transitions. The results from process mining are triangulated and validated by statistical forms of learning analytics (sequence mining and decision trees), which identify significant isolated patterns of tutor-learner interaction. These patterns of interaction can be interpreted in light of the results from process mining, creating a richer understanding of how co-regulation occurs in practice.

Finally, my research contributes to the field by highlighting the differences between co-regulatory tutoring practices, between high and low-performing tutors. Effective tutoring practices are identified as being complex, diverse, and inclusive of open-ended forms of engagement. On the other hand, maladaptive tutoring practices tend to follow a more restricted range which aligns closely to the structure of the VCE and focuses on directive forms of engagement. Changing the performance measure from evaluator rankings for student centred teaching, to average progress scores, does not substantially change the nature of the results. The key differences between high and low-performing tutors are the same, regardless of the performance measure. However, the extent of the differences in co-regulatory patterns of interaction is greater when performance is measured with respect to evaluator rankings versus average progress scores, which makes sense given the indirect relationship between average progress scores and co-regulation.

7.4 Future Research Directions

The results from my research indicate potential for theoretical development in further work. For example, future research can explore whether the contextualised model of SRL

proposed in this thesis could potentially be further built upon, to show favourable and maladaptive pathways, as process-oriented research into SRL advances. The process maps generated by high-performing tutors show that open-ended tutoring practices often lead to sophisticated learner operations (such as reasoning). Further research can explore the subsequent impact of learner reasoning on inter-subjectivity and learner motivation/sense of well-being, to highlight favourable tutor-learner loops, with respect to SRL.

With regard to the framework of signifiers, the high number of events included in the framework mean that process maps need to view at high levels to discover a coherent process. It may be possible in later research, to focus the analysis on certain domains of the framework (e.g., metacognitive or affective). This would limit the number of events/transitions and enable a more detailed process to be visualised. However, it is important to recognise that focusing on specific domains of the framework would obscure features of the overall process, such as structure or reliance on certain events/transitions. Another area of research with regard to the framework of signifiers is the automation of labelling of signifiers derived from natural language dialogue, using LLMs. The approach set out in this thesis uses manual coding of natural language dialogue, but this approach would be difficult to scale.

My research explores an approach towards detecting the differences in how high-performing tutors co-regulate learners in a VCE, versus low-performing tutors. As such, the research has a range of potential operational implications, including intelligent tools that can support tutor evaluation, tutor training, and the learning design of the platform.

Future research could explore the influence of learning design on tutor co-regulatory patterns. This thesis focuses on tutor-learner interactions in the problem-solving stages of VCE sessions. Future research could consider tutor-learner interactions over the entire session, including the introduction, warm up, recap, problem solving and reflection stages. It

would be interesting to explore how tutor-learner interactions evolve over the course of the session, and how this varies with tutor performance.

Another area for further research is to consider the role of time within tutor-learner interactions. This thesis focuses on the order of events and transitions, to describe a holistic process. Future research could also consider how much time is spent on various activities, and how this varies with tutor performance.

A further area for potential future research is the development of intelligent tutor evaluation and training tools, building on the research findings presented in this thesis. This future research could build on recent developments in the field of AI, such as the development of foundational large language models (e.g., ChatGPT). Foundational large language models could potentially enable the automated labelling of natural language dialogue, a key data source in human-led online tutoring. This avenue for research is further described in the following section.

Tutor training and evaluation tools

Tutor training and quality assurance pose a significant challenge to scaling online human to human tutoring. In order to be able to successfully scale online human to human tutoring, it is necessary to be able to effectively monitor tutor-learner interactions at scale. Research, such as this thesis, can help to identify significant patterns of tutor-learner interaction, which can classify tutors into high and low-performing groups. This research can inform intelligent tools to support tutor evaluation and training.

Figure 46 depicts a potential tool to support tutor evaluation. The aim of this tool is to analyse and monitor tutor-student interactions in the VCE, directing the QA teams to focus on the tutors who are most in need of support. Similarly, we aim to explore a potential tool to support personalised professional development for tutors, depicted in Figure 47. This tool

will make suggestions for training and development, and visualise the tutors' skills and areas for improvement, mapping to real data from tutoring sessions conducted.

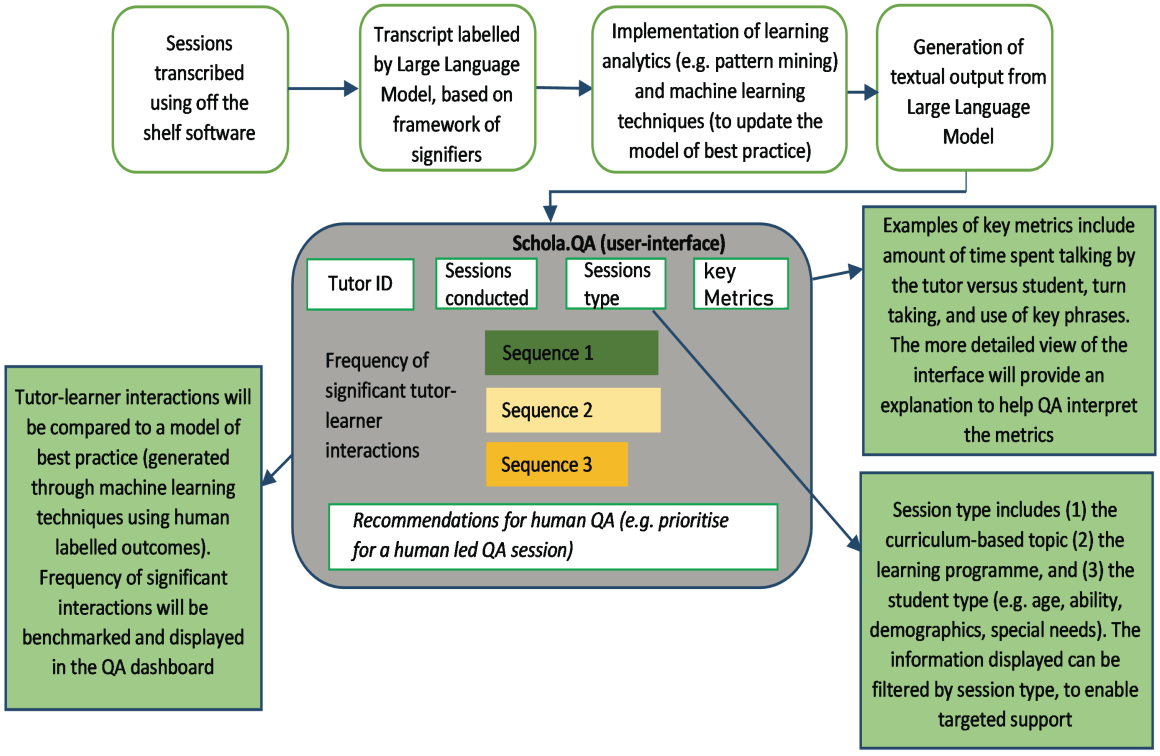


Figure 46. Intelligent tool to support QA

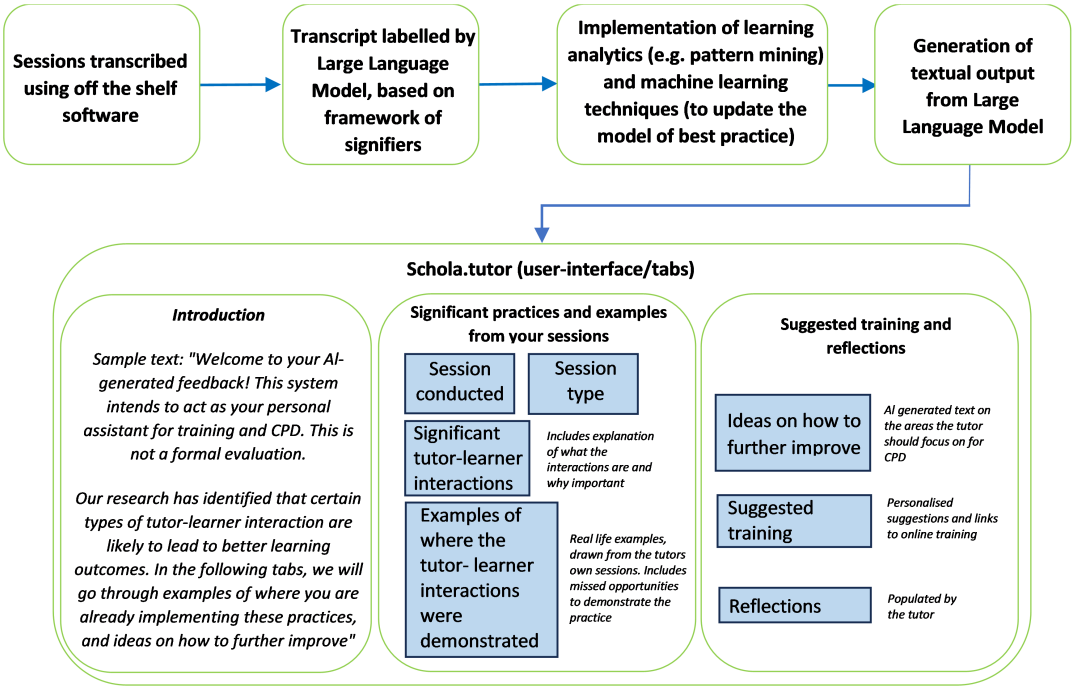


Figure 47. Intelligent tool to support tutor training

Appendix

Appendix A

Process for exploratory literature review

I followed the below steps to carry out my exploratory literature review:

1. I searched UCL databases using the following key-words and combinations of key-words, between the years 2000-2022:
    a. 'Tutor' AND 'co-regulation
    b. 'Teaching' AND 'co-regulation'
    c. 'Co-regulation' AND 'learning'
    d. 'Co-regulation' AND 'education
    e. 'Virtual classroom environment' AND 'co-regulation
    f. 'Online tutoring'
    g. 'One-to-one tutoring'
2. Results were filtered to exclude any papers that have not been peer reviewed.
3. I further evaluated the relevance of the paper by reading the paper abstracts. Papers which dealt with co-regulation in dissimilar contexts, such as small groups rather than one to one co-regulation, were judged to be out of scope.
4. Papers that were judged to be relevant to my research question were read in depth, and the quality of the paper was assessed based on the robustness of the research design, methods of data analysis, validation of results, and depth and breadth of the bibliography. Any papers that were judged to be of insufficient quality were excluded from the review.

5. I reviewed the bibliographies of the selected papers for additional references which appeared relevant to my studied context. These papers then passed though review steps (3) and (4) as stated above.

Appendix B

Evaluation of the reliability of the framework

Following the development of the framework of signifiers, I developed an approach towards coding the signifiers in real time. Given that manual coding was required, this approach followed the approach set out in Cukurova et al. (2022). The authors of this study developed a simple tool which enabled them to manually code events observed in real time, providing data on the event observed, timestamp, and session ID. I developed a similar tool to enable the coding of sessions, the screenshot of which is provided below (insert screenshot). In addition to the development of a coding tool, I also developed a set of rules for coding, based on studies which adopted a similar approach towards coding data. Further details of the coding rules adopted are provided in chapter x of this thesis. The reliability of the framework of signifiers was evaluated through the coding of 10 randomly selected sessions, with 2 trained coders, and an inter-rater reliability of ordinal k alpha 0.724. The framework was iterated on the basis of the results of the inter-rater reliability test, and the inter-rater reliability was evaluated again on the basis of coding another 15 randomly selected sessions. An inter-rater reliability of ordinal k alpha 0.842 was achieved, which was judged sufficient to proceed with coding the main dataset.

Appendix C

Description of the fuzzy miner algorithm

The fuzzy miner algorithm uses significance and correlation to judge the importance of events and relations. Significance refers to the level of interest we have in event classes

(i.e. activities) and of their relations (i.e. edges). One example of a measure for significance is frequency, with more frequent events and edges being regarded as more significant. On the other hand, correlation is relevant only for relations and measures how closely related the starting point and end point of a relation are, with regard to proximity, event name or data attributes associated with the event. Significance and correlation are calculated using the primary metrics set out in the table below. It is important to note that for each of the three primary types of metrics (unary significance, binary significance and binary correlation), different implementations can be used, ranging from log-based metrics to derivative metrics. In highly unstructured logs, there may be a high number of exceptional events which occur between key events. In order to measure the causal dependency between the desired event classes, indirect relations are measured in addition to direct relations, with the measured value attenuated with the increasing length of the relationship.

Metric	Description	Implementation (log-based)	Implementation (derivative)
Unary significance	Measures the relative importance of an event class	Relative frequency	Routing significance. Calculates the importance of the node/event, based on the difference between the incoming and outgoing arcs for the node, and thus its importance in the overall structure of the process
Binary significance	Measures the relative importance of a relation between two event classes	Relative frequency	Distance significance: The difference between the frequency of a relation, and its start and end point. The lower this difference, the higher the distance significance value will be, as the implication of that the relation is important for its start

			and end points. This metric is particularly useful in unstructured logs, with many competing relations, in removing the less important relations
Binary correlation	Measures the distance of events connected by the relation, where distance can be considered the magnitude of context change		Proximity correlation Data attribute correlation Event name correlation

The fuzzy miner algorithm differs from other process mining algorithms in that it does not follow an interpretative approach by attempting to map event log data to a pre-defined process design (for example, AND-or XOR-semantics). Instead, the fuzzy miner algorithm seeks to map all behaviour in the event log, prior to the level of detail being adapted depending on the configuration of the algorithm. In addition, the fuzzy miner algorithm seeks to represent a factual ordering of relations, and does not accommodate concurrency within the process. The algorithm follows the below steps to create the process model, as described by:

- A process model is generated, which maps all observed behaviour from the event log. All events classes are translated to activity nodes, with a description of their unary significance. Similarly, for every observed relation, a corresponding edge is added to the process model, with a description of its binary significance and binary correlation
- Three transformation methods are applied to simplify the model:
    - Conflict resolution. When two nodes in the initial model are connected by edges in both directions, they are deemed to be in conflict. This conflicting relation may either be an exception, reflect concurrent behaviour (i.e., the two nodes are on distinct parallel paths), or the conflicting relation may reflect a

loop which occurs in reality. The conflicting relation only needs to be preserved if it reflects a loop in the process model. To determine whether the relation should be preserved or discarded, the algorithm calculates the relative significance of both relations. If the relative significance of both conflicting relations exceeds a certain threshold, this indicates that there is a loop within the process model which needs to be preserved. If at least one of the relations has a relative significance less than the threshold, then the offset (i.e. the difference between the two relations is calculated). In the scenario that the offset value is greater than a pre-determined amount, this indicates that one relation is more important than another, and the latter is an exception and should be discarded. In the scenario that offset value is below the threshold, this indicates that both conflicting relations are equally important and likely to be concurrent (i.e. on different threads of the process).

- Edge filtering. Edges are filtered using a utility ratio, which is a weighted sum of binary significance and correlation. Filtering is not conducted using significance alone as this can result in small clusters of highly frequent behaviours, and prevent a holistic representation of the process The utility ratio can be configured to prefer either binary significance or correlation as a metric for preserving edges.
- Aggregation and abstraction. Nodes with a unary significance below a certain parameter become ‘victims’, and are either aggregated or abstracted. In the first stage, the model seeks to build clusters by aggregating nodes with highly correlated neighbour. In the second stage, isolated and singular clusters are abstracted from the process. The most significant relations of the removed

nodes can be artificially preserved and connected to the successor nodes if required.

Appendix D

Development of event log and importing into Disco

A challenge identified during the development of the event log, was the potential issue of signifiers occurring in parallel. For example, a tutor may write on the whiteboard while the learner is talking. This posed an issue in the representation of the process, in the process maps. Simply timestamping the parallel behaviours as separate events within the process map would risk representing the behaviours as occurring on parallel threads of the process and being discarded by the model. It was decided that the most accurate representation of the process would be to depict the behaviours as a single event that consists of two behaviours, rather than 2 isolated behaviours. For example, 'tutor writing' in parallel with 'learner speaking' was processed into 'tutor writing and learner speaking'. For the purposes of the subsequent pattern mining and statistical analysis, the parallel events were treated as a combined event, as well as separate events, with the aim of detecting all relevant significant patterns.

Following the development of the event log, the file was imported into Disco, and can be inspected using the tool. The Disco log analyses and presents log statistics (e.g., frequency of events). It also evaluates the number of case variants which occur within the log, where variant refers to a specific sequence of events. Variant analysis can be useful when analysing structured processes to detect variations in the process. It is less applicable for an unstructured dataset, such as the one that I am using for my research, as every case follows a unique sequence of events. The value of process mining is in detecting the strength of the direct, as well as indirect transitions, to describe the fundamental features of the process. I

therefore focused my work on implementing the fuzzy miner algorithm in Disco, rather than using the variant analysis also offered by the tool.

Appendix E

Overview of the SPMF pattern mining database

SPMF is an open-source project, started in 2009 to address the lack of large open-source data mining library specialized in frequent pattern mining.

SPMF provides more than 55 algorithms for pattern mining. Implementations of most of these algorithms can only be found in SPMF. For example, only three algorithms from SPMF appear in Weka and Knime (Apriori, FPGrowth and GSP), only one in Mahout (FPGrowth), two in LUCS-KDD (Apriori, FPGrowth), and eight in Coron.

Using SPMF

SPMF is implemented in Java and is cross-platform. The only requirement to run SPMF is to have Java 7 or higher installed. There are two versions of SPMF. The source code version offers all algorithms from SPMF. The documentation provides an example of how to run each algorithm. It explains the input and output of each algorithm, its main characteristics and where to obtain more information about the algorithm. Moreover, a sample program and input file are provided in the source code of SPMF to show how to execute each example from Java code. Running an algorithm is just a few lines of code

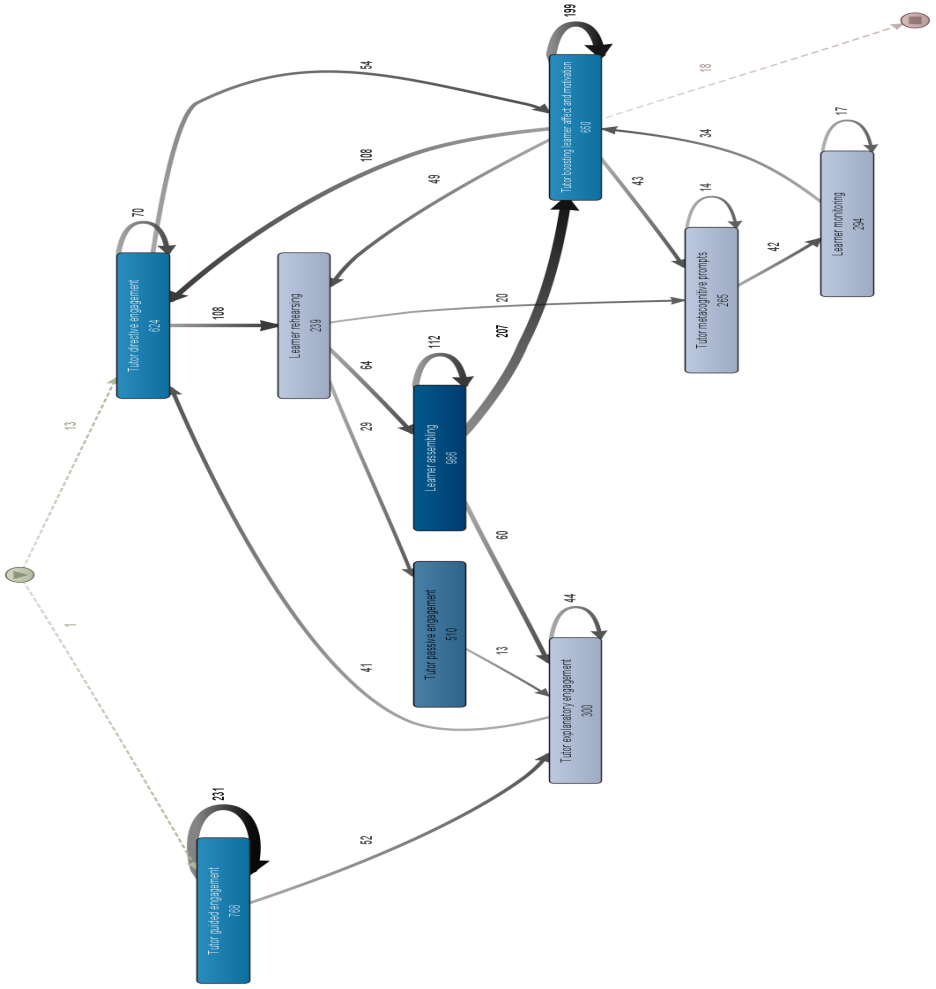
The release version of SPMF is a runnable JAR file that can be launched with a doubleclick. It provides a minimalist user interface (see Figure 1), designed to allow quickly testing the behaviour of the algorithms. The graphical user interface allows one to select an input file and an output file, choose an algorithm, enter its parameters and run it.

Input files for the algorithms are text files. The format that is used is the one from frequent pattern mining competitions such as FIMI and used by researchers in this domain (files where items are represented by integers). But, to allow greater interoperability, the GUI

and command line version of SPMF can also read the popular ARFF file format for itemset and association rule mining (used by Weka and Knime), and tools are provided to convert some selected formats to the SPMF format

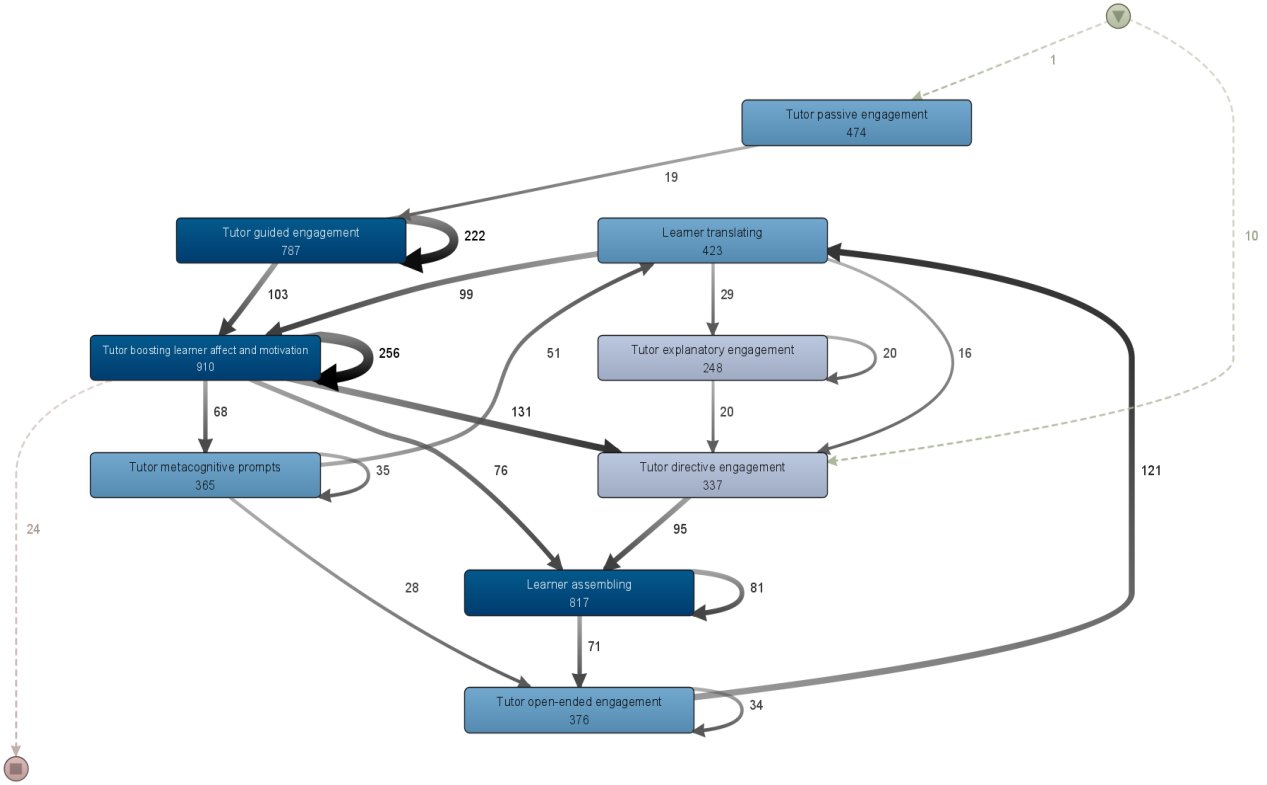
Appendix F Process maps

1. Process map 3 LOW HE tutors, event categories, paths sliders at 25 percent



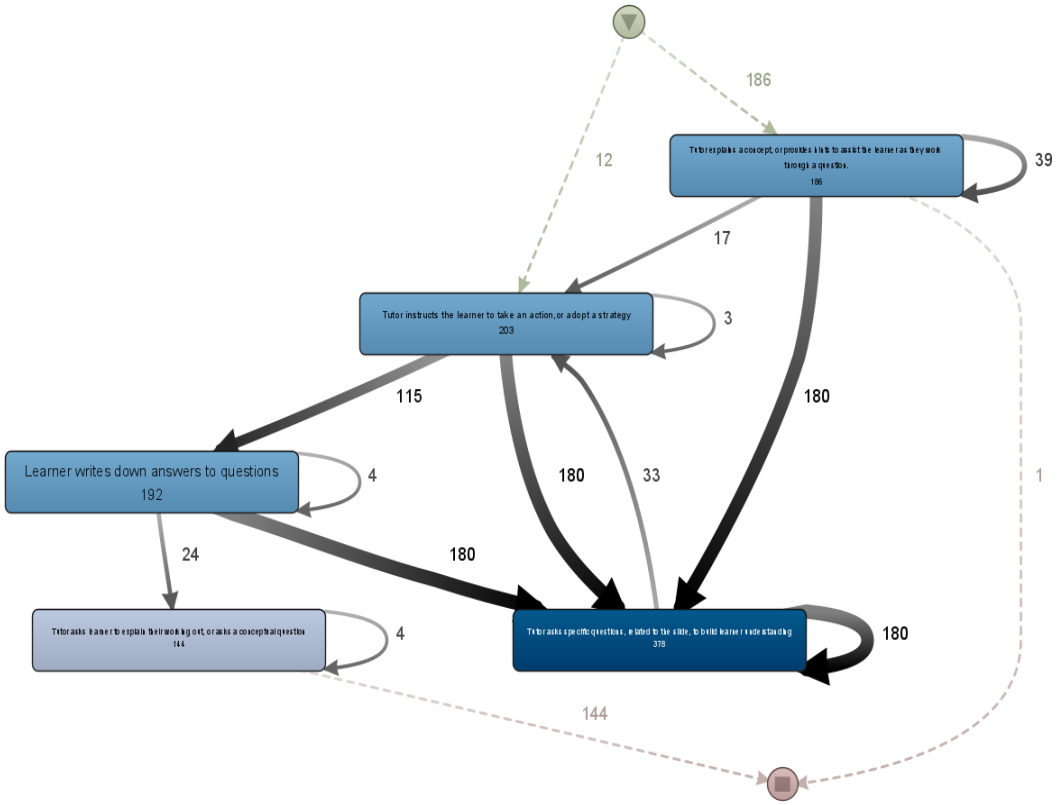
Process map 3 low-ranked tutors (path sliders at 25%)

2. [Process map 4 HIGH HE tutors, event categories, paths sliders at 25 percent]

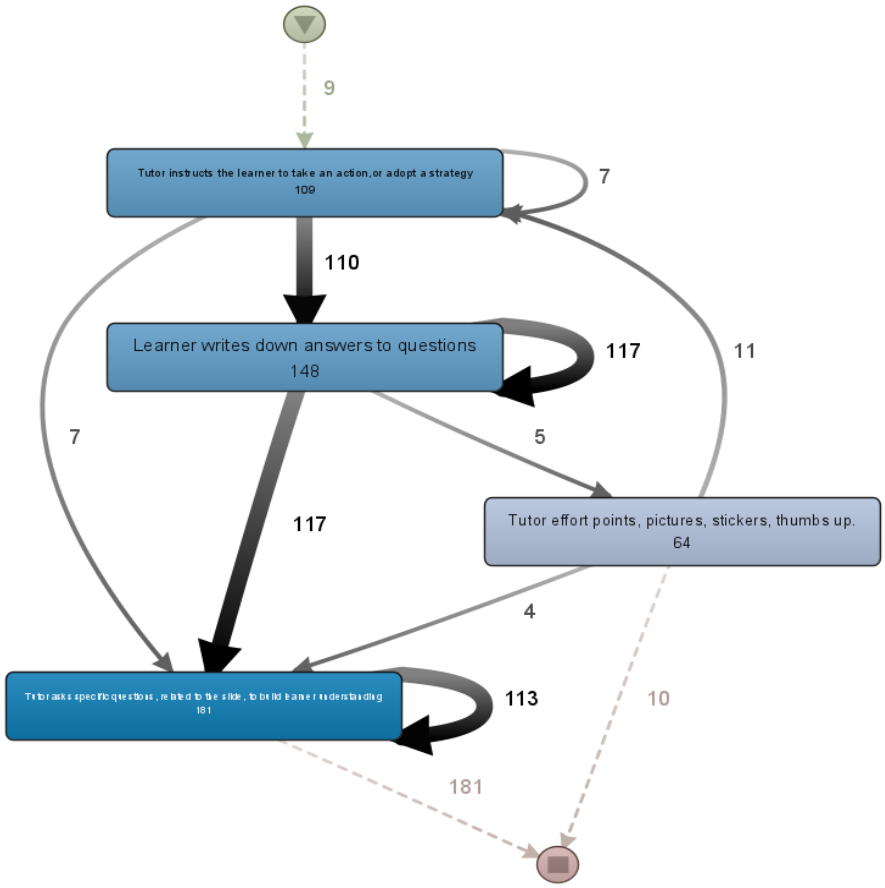


Process map 4 High-ranked tutors, event category view (sliders set

3. Process map 17_HIGH APS_signifiers map_event sliders at 10% and path sliders at 25 percent



4. [Process map 18_LOW APS_signifiers map_event sliders at 10 percent and path sliders at 25 percent]



Appendix G Signed ethics document (screenshot)

Institute of Education

UCL

Doctoral Student Ethics Application Form

Anyone conducting research under the auspices of the Institute of Education (staff, students or visitors) where the research involves human participants or the use of data collected from human participants, is required to gain ethical approval before starting. This includes preliminary and pilot studies. Please answer all relevant questions in simple terms that can be understood by a lay person and note that your form may be returned if incomplete.

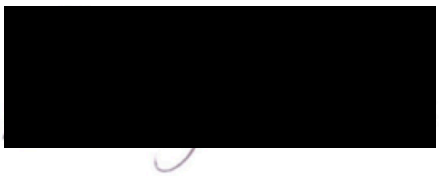
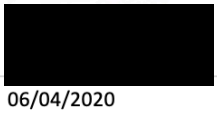
Registering your study with the UCL Data Protection Officer as part of the UCL Research Ethics Review Process

If you are proposing to collect personal data i.e. data from which a living individual can be identified **you must be registered with the UCL Data Protection Office before you submit your ethics application for review**. To do this, email the complete ethics form to data-protection@ucl.ac.uk. Once your registration number is received, add it to the form* and submit it to your supervisor for approval. If the Data Protection Office advises you to make changes to the way in which you propose to collect and store the data this should be reflected in your ethics application form.

Please note that the completion of the UCL GDPR online training is mandatory for all PhD students. The link is here: https://www.ucl.ac.uk/legal-services/ucl-general-data-protection-regulation-gdpr/gdpr-online-training

Section 1 Project details				
a.	Project title			How can we use models of Self-Regulated Learning (SRL) to inform the use of Learning Analytics to detect patterns of SRL in a virtual classroom environment?
b.	Student name and ID number (e.g. ABC12345678)			Madiha Khan, ID 840374
c.	***UCL Data Protection Registration Number**			No Z6364106/2020/03/96
c.	Supervisor/Personal Tutor			Professor Rose Luckin Dr Mutlu Cukurova
d.	Department			Culture, Communication and Media
e.	Course category (Tick one)	PhD	☒	EdD ☐
		DEdPsy	☐	
f.	**If applicable,** state who the funder is and if funding has been confirmed.			Third Space Learning, funding confirmed
g.	Intended research start date			September 2019
h.	Intended research end date			June 2020
i.	Country fieldwork will be conducted in			United Kingdom

November 2018

Reviewer 1		
Supervisor/first reviewer name	Rose Luckin	
Do you foresee any ethical difficulties with this research?	no	
Supervisor/first reviewer signature		
Date		
Reviewer 2		
Second reviewer name	Mutlu Cukurova	
Do you foresee any ethical difficulties with this research?	No	
Supervisor/second reviewer signature		
Date	06/04/2020	
Decision on behalf of reviews		
Decision	Approved	☐
	Approved subject to the following additional measures	☐
	Not approved for the reasons given below	☐
	Referred to REC for review	☐
Points to be noted by other reviewers and in report to REC		
Comments from reviewers for the applicant		

Once it is approved by both reviewers, students should submit their ethics application form to the Centre for Doctoral Education team: IOE.CDE@ucl.ac.uk.

Appendix H Petri Net models

Components of a Petri Net

Places (Circles)

- Represent conditions or states within the process.
- Can hold tokens, which represent the state of the process at a particular point.
- Often, places are used to model the preconditions and postconditions of transitions.

Transitions (Squares or Rectangles)

- Represent activities or events in the process.
- Come in two forms:
    - Black squares: Silent transitions that do not correspond to any observable activity in the event log but are used for routing and synchronization purposes.
    - White squares with words: Labeled transitions that correspond to actual activities or events recorded in the event log.

Tokens

- Represent the current state of the process.
- Tokens reside in places and move around the net, enabling and firing transitions according to the defined rules.
- When a transition fires, it consumes tokens from its input places and produces tokens in its output places.

Arcs (Arrows)

- Represent the flow between places and transitions.
- Connect places to transitions (indicating preconditions) and transitions to places (indicating postconditions).
- Arcs define the relationship and flow of tokens within the Petri net.

How It Works

- Initial Marking: The starting state of the Petri net, defined by the initial distribution of tokens in places.
- Enabled Transitions: A transition is enabled when all its input places contain the required number of tokens.
- Firing of Transitions: When an enabled transition fires, it consumes tokens from its input places and produces tokens in its output places, thereby moving the process forward.

Example Walkthrough

- Starting State: Tokens are placed in initial places.
- Activity Execution: A labeled transition (white square with words) is enabled and fires, consuming tokens from its input places and producing tokens in its output places.
- Silent Transition: A black square (silent transition) may fire to synchronize or route tokens without representing any observable activity.
- Process Continuation: Tokens continue to move through the Petri net according to the flow defined by arcs and transitions until the process reaches its final state.

Appendix I Sample coded data (Dataset E)

I have provided below a snapshot of the data collected from coding sessions using the framework of indicators.

Case ID	Time	Event code	Event	Category	Category code	Learner/tutor action
1122224	11:05:33 AM	1	Tutor Praise	Tutor boosting learner affect and motivation	1	Tutor
1122224	11:06:02 AM	12	Tutor instructs the learner to take an action,or adopt a strategy (statements using command words), e.g. can you please read this, or fill out this colum chart. Excluding process related commands e.g. can you go to next slide/questions, or lets start questions	Tutor directive engagement	4	Tutor
1122224	11:07:21 AM	37	Learner reasoning out loud	Learner translating	12	Learner
1122224	11:08:00 AM	10	Tutor prompts learner to plan approach (Forward looking statements 'How will you do this'?))	Tutor metacognitive and reflective prompts	3	Tutor
1122224	11:09:00 AM	36	Learner planning out approach	Learner translating	12	Learner
1122224	11:10:03 AM	1	Tutor Praise	Tutor boosting learner affect and motivation	1	Tutor
1122224	11:10:52 AM	20	Tutor asks learner to explain their working out, or a conceptual question e.g. how did you do that? Why do you think that is the case?	Tutor open-ended engagement	6	Tutor
1122224	11:11:02 AM	36	Learner planning out approach	Learner translating	12	Learner
1122224	11:11:49 AM	10	Tutor prompts learner to plan approach (Forward looking statements 'How will you do this'?))	Tutor metacognitive and reflective prompts	3	Tutor

Appendix J Sample coded data (observation protocol)

I have provided below a snapshot of the data collected from observing learners during sessions using the observation protocol.

School ID	Learner ID	Time	Observation code	Observation
D	123	10:06:00 AM	1	Eyes scanning the screen
D	123	10:07:00 AM	3	Learner is typing on the keyboard
D	123	10:09:10 AM	2	Learner pauses after writing or speaking and shows signs of thought (e.g. frowns, postural changes increase)
D	123	10:10:23 AM	4	Eyes tracking along screen while learner speaks
D	123	10:11:13 AM	3	Learner is typing on the keyboard
D	123	10:12:15 AM	4	Eyes tracking along screen while learner speaks
D	123	10:13:04 AM	5	Explanatory gestures combined with talking in turns
D	123	10:14:50 AM	3	Learner is typing on the keyboard
D	123	10:15:05 AM	1	Hands moving the mouse for a sustained period of time
D	123	10:15:10 AM	4	Eyes tracking along screen while learner speaks
D	123	10:15:40 AM	3	Learner is typing on the keyboard

Appendix K Snapshot of process mining event log

I have provided below a snapshot of the event log used for process mining.

Case ID	Time	Event code	Event	Category	Category code	Learner/tutor action
1122224	11:05:33 AM	1	Tutor Praise	Tutor boosting learner affect and motivation	1	Tutor
1122224	11:06:02 AM	12	Tutor instructs the learner to take an action,or adopt a strategy (statements using command words), e.g. can you please read this, or fill out this colum chart. Excluding process related commands e.g. can you go to next slide/questions, or lets start questions	Tutor instructional engagement	4	Tutor
1122224	11:07:21 AM	37	Learner reasoning out loud	Learner translating	12	Learner
1122224	11:08:00 AM	10	Tutor prompts learner to plan approach (Forward looking statements 'How will you do this'?))	Tutor metacognitive and reflective prompts	3	Tutor
1122224	11:09:00 AM	36	Learner planning out approach	Learner translating	12	Learner
1122224	11:10:03 AM	1	Tutor Praise	Tutor boosting learner affect and motivation	1	Tutor
1122224	11:10:52 AM	20	Tutor asks learner to explain their working out, or a conceptual question e.g. how did you do that? Why do you think that is the case?	Tutor open-ended engagement	6	Tutor
1122224	11:11:02 AM	36	Learner planning out approach	Learner translating	12	Learner
1122224	11:11:49 AM	10	Tutor prompts learner to plan approach (Forward looking statements 'How will you do this'?))	Tutor metacognitive and reflective prompts	3	Tutor

Appendix L Tutor interview questionnaire

Introduction/context of work

Background questions

1. How long have you been working for TSL:
    - Less than 1 year
    - 1-5 years
    - 5-10 years

2. Prior to working for TSL, was your primary job within teaching?
    - Yes
    - No

3. How many years' experience do you have teaching in online and offline environments?
    - Less than 1 year
    - 1-5 years
    - 5-10 years
    - More than 10 years

4. How many years' experience do you have teaching in online environments?
    - Less than 1 year
    - 1-5 years
    - 5-10 years
    - More than 10 years

Teacher approach towards student centred learning (SCT)

5. Do you understand what is referred to as student centred learning (if no, then interviewer is to provide an explanation):
    - Yes
    - No
    - Unsure

6. What is your understanding of SCT? (open text)

7. How important is SCT within a tutoring session?
    - Not important
    - Somewhat important
    - Very important

8. Does the importance of SCT vary depending on the ability of the learner? If yes, then how?
    - Yes
    - No

9. How do you seek to promote student centred learning? (open text)

10. Can you please rate the importance of the below activities in promoting SCT in TSL sessions, from a scale of 1-10, where 1 is not important, and 10 is the most important?
    - Developing a rapport with learner through asking about personal life, hobbies etc
    - Frequent praise and reward
    - Encouraging the learner to set goals for themselves during a TSL session
    - Asking the learner specific, clear questions
    - Asking the learner open ended questions, or prompting them to explain their answers and reasoning

Tutor approach towards CPD

11. What is the role of the TSL platform during your tutoring sessions?
    - Flexible. I tailor my approach towards the learner, and will use the platform to support my own approach
    - Fixed role. My sessions are led by the structure of the platform

12. How do you adapt your teaching to the learner, depending on the levels of engagement and ability? (open text)

13. How often do you reflect on and evaluate your teaching, independently of the TSL evaluations?
    - After every session
    - Monthly
    - Quarterly
    - Annually

14. What resources do you use to evaluate and reflect upon your teaching? (open text)

15. What are the challenges of developing your skills with regard to SCT? (open text). Please rate the below factors from 1-10:
    - Lack of evidence of SCT development amongst learners
    - Lack of clarity on what SCT is
    - Lack of training on how to promote SCT
    - Low priority of SCT amongst other training goals

References

Ahmad Uzir, N., Gašević, D., Matcha, W., Jovanović, J., & Pardo, A. (2020). Analytics of time management strategies in a flipped classroom. *Journal of Computer Assisted Learning*, *36*(1), 70–88. https://doi.org/10.1111/jcal.12392

Aleven, V. (2010). Rule-based cognitive modeling for intelligent tutoring systems. In R. Nkambou, J. Bourdeau, & R. Mizoguchi (Eds.), *Advances in intelligent tutoring systems* (pp. 33–62). Springer. https://doi.org/10.1007/978-3-642-14363-2_3

Allal, L. (2020). Assessment and the co-regulation of learning in the classroom. *Assessment in Education: Principles, Policy & Practice*, *27*(4), 332–349. https://doi.org/10.1080/0969594X.2019.1609411

Allal, L., & Pelgrims Ducrey, G. (2000). Assessment *of*—or *in*—the zone of proximal development. *Learning and Instruction*, *10*(2), 137–152. https://doi.org/10.1016/S0959-4752(99)00025-0

Andrade, H. L., & Brookhart, S. M. (2020). Classroom assessment as the co-regulation of learning. *Assessment in Education: Principles, Policy & Practice*, *27*(4), 350–372. https://doi.org/10.1080/0969594X.2019.1571992

Andrade, H. L., & Dugan, R. F. (2011). Exploring the construct validity of academic self-regulation using a new self-report questionnaire – The Survey of Academic Self-Regulation. *International Journal of Educational and Psychological Assessment*, *7*(1), 45–63.

Araya, R., & Aljovin, E. (2018). The effect of teacher questions on elementary school students' written responses on an online STEM platform. In T. Andre (Ed.), *Advances in Human Factors in Training, Education, and Learning Sciences* (pp. 372–382). Springer International Publishing. https://doi.org/10.1007/978-3-319-60018-5_36

Ayres J., Gehrke J., Yiu T., & Flannick J. (2002). Sequential pattern mining using a bitmap representation. In *Proceedings of the eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 429–435). Association for Computing Machinery.

Azevedo, R. (2014). Issues in dealing with sequential and temporal characteristics of self- and socially-regulated learning. *Metacognition and Learning*, *9*, 217–228. https://doi.org/10.1007/s11409-014-9123-1

Azevedo, R., & Hadwin, A. (2005). Scaffolding self-regulated learning and metacognition—Implications for the design of computer-based scaffolds. *Instructional Science*, *33*, 367–379. https://doi.org/10.1007/s11251-005-1272-9

Azevedo, R., Johnson, A., Chauncey, A., & Graesser, A. (2011). Use of hypermedia to assess and convey self-regulated learning. In B. J. Zimmerman & D. H. Schunk (Eds.), *Handbook of self-regulation of learning and performance* (pp. 102–121). Routledge. https://doi.org/10.4234/9780203839010.ch7

Azevedo, R., Mudrick, N. V., Taub, M., & Bradbury, A. E. (2019). Self-regulation in computer-assisted learning systems. In J. Dunlosky & K. A. Rawson (Eds.), *Handbook of cognition and education* (pp. 587–618). Cambridge University Press. https://doi.org/10.1017/9781108235631.024

Baird, L. L. (1983). *Attempts at defining personal competency* (Publication No. 83-15). Educational Testing Service.

Baker, R. D., Goldstein, A. B., & Heffernan, N. T. (2011). Detecting learning moment-by-moment. *International Journal of Artificial Intelligence in Education*, *21*(1–2), 5–25. https://doi.org/10.3233/JAI-2011-015

Bandura, A., & Cervone, D. (1983). Self-evaluative and self-efficacy mechanisms governing the motivational effects of goal systems. *Journal of Personality and Social Psychology*, *45*(5), 1017–1028. https://doi.org/10.1037/0022-3514.45.5.1017

Bandura, A., & National Institute of Mental Health. (1986). *Social foundations of thought and action: A social cognitive theory*. Prentice-Hall.

Bandura, A., & Schunk, D. H. (1981). Cultivating competence, self-efficacy, and intrinsic interest through proximal self-motivation. *Journal of Personality and Social Psychology*, *41*(3), 586–598. https://doi.org/10.1037/0022-3514.41.3.586

Bannert, M., Hildebrand, M., & Mengelkamp, C. (2009). Effects of a metacognitive support device in learning environments. *Computers in Human Behavior*, *25*(4), 829–835. https://doi.org/10.1016/j.chb.2008.07.002

Bannert, M., Molenaar, I., Azevedo, R., Järvelä, S., & Gašević, D. (2017). Relevance of learning analytics to measure and support students' learning in adaptive educational technologies. In *Proceedings of the seventh international learning analytics & knowledge conference* (pp. 568–569). Association for Computing Machinery. https://doi.org/10.1145/3027385.3029463

Bannert, M., Reimann, P., & Sonnenberg, C. (2014). Process mining techniques for analysing patterns and strategies in students' self-regulated learning. *Metacognition and Learning*, *9*, 161–185. https://doi.org/10.1007/s11409-013-9107-6

Bannert, M., Sonnenberg, C., Mengelkamp, C., & Pieger, E. (2015). Short- and long-term effects of students' self-directed metacognitive prompts on navigation behavior and learning performance. *Computers in Human Behavior*, *52*, 293–306. https://doi.org/10.1016/j.chb.2015.05.038

Beheshitha, S. S., Gašević, D., & Hatala, M. (2015). A process mining approach to linking the study of aptitude and event facets of self-regulated learning. In *Proceedings of the Fifth International Conference on Learning Analytics and Knowledge* (pp. 265–269). Association for Computing Machinery. https://doi.org/10.1145/2723576.2723628

Bell, C. V., & Pape, S. J. (2014). Scaffolding the development of self-regulated learning in mathematics classrooms. *Middle School Journal*, *45*(4), 23–32. https://doi.org/10.1080/00940771.2014.11461893

Bober, M. J., & Dennen, V. P. (2001). Intersubjectivity: Facilitating knowledge construction in online environments. *Educational Media International*, *38*(4), 241–250. https://doi.org/10.1080/09523980110105150

Boekaerts, M. (2002). Bringing about change in the classroom: Strengths and weaknesses of the self-regulated learning approach—EARLI Presidential Address, 2001. *Learning and Instruction*, *12*(6), 589–604. https://doi.org/10.1016/S0959-4752(02)00010-5

Boekaerts, M., & Corno, L. (2005). Self-regulation in the classroom: A perspective on assessment and intervention. *Applied Psychology*, *54*(2), 199–231. https://doi.org/10.1111/j.1464-0597.2005.00205.x

Bogarín, A., Cerezo, R., & Romero, C. (2018). A survey on educational process mining. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *8*, Article e1230. https://doi.org/10.1002/widm.1230

Braund, H., DeLuca, C., Panadero, E., & Cheng, L. (2021). Exploring formative assessment and co-regulation in kindergarten through interviews and direct observation. *Frontiers in Education*, *6*, Article 32373. https://doi.org/10.3389/feduc.2021.732373

Brookhart, S. M., & Chen, F. (2015). The quality and effectiveness of descriptive rubrics. *Educational Review*, *67*(3), 343–368. https://doi.org/10.1080/00131911.2014.929565

Brooks, C., & Thompson, C. (2017). Predictive modelling in teaching and learning. In C. Lang, G. Siemens, A. F. Wise, & D. Gaevic (Eds.), *The handbook of learning analytics* (pp. 61–68). Society for Learning Analytics Research. https://doi.org/10.18608/hla17.005

Bull, S., & Kay, J. (2010). Open learner models. In R. Nkambou, J. Bourdeau, & R. Mizoguchi (Eds.), *Advances in intelligent tutoring systems* (pp. 301–322). Springer Berlin. https://doi.org/10.1007/978-3-642-14363-2_15

Butler, D. L., & Cartier, S. C. (2017). Advancing research and practice about self-regulated learning: The promise of in depth case study methodologies. In D. H. Schunk & J. A. Greene (Eds.), *Handbook of self-regulation of learning and performance* (2nd ed., pp. 352–369). Routledge. https://doi.org/10.4324/9781315697048-23

Butler, D. L., & Schnellert, L. (2015). Success for students with learning disabilities: What does self-regulation have to do with it? In T. Cleary (Ed.), *Self-regulated learning interventions with at-risk youth: Enhancing adaptability, performance, and well-being* (pp. 89–111). American Psychological Association. https://doi.org/10.1037/14641-005

Cade, W. L., Copeland, J. L., Person, N. K., & D'Mello, S. K. (2008). Dialogue modes in expert tutoring. In B. P. Woolf, E. Aïmeur, R. Nkambou, & S. Lajoie (Eds.), *Intelligent tutoring systems* (pp. 470–479). Springer. https://doi.org/10.1007/978-3-540-69132-7_50.

Camacho, V. L., de la Guía, E., Olivares, T., Flores, M. J., & Orozco-Barbosa, L. (2020). Data capture and multimodal learning analytics focused on engagement with a new wearable IoT approach. *IEEE Transactions on Learning Technologies*, *13*(4), 704–717. https://doi.org/10.1109/TLT.2020.2999787

Cerezo, R., Bogarín, A., Esteban, M., & Romero, C. (2020). Process mining for self-regulated learning assessment in e-learning. *Journal of Computing in Higher Education*, *32*, 74–88. https://doi.org/10.1007/s12528-019-09225-y

Chan, C. K. K. (2012). Co-regulation of learning in computer-supported collaborative learning environments: A discussion. *Metacognition and Learning*, *7*, 63–73. https://doi.org/10.1007/s11409-012-9086-z

Chi, M. T. H. (1997). Quantifying qualitative analyses of verbal data: A practical guide. *Journal of the Learning Sciences*, *6*(3), 271–315. https://doi.org/10.1207/s15327809jls0603_1

Chi, M. T. H., Siler, S., Jeong, H., Yamauchi, T., & Hausmann, R. (2001). Learning from human tutoring. *Cognitive Science*, *25*(4), 471–533. https://doi.org/10.1016/S0364-0213(01)00044-1

Chi, M. T. H., & Wylie, R. (2014). The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational Psychologist*, *49*(4), 219–243. https://doi.org/10.1080/00461520.2014.965823

Cho, M.-H., & Yoo, J. S. (2017). Exploring online students' self-regulated learning with self-reported surveys and log files: A data mining approach. *Interactive Learning Environments*, *25*(8), 970–982. https://doi.org/10.1080/10494820.2016.1232278

Cimpian, A., Arce, H.-M. C., Markman, E. M., & Dweck, C. S. (2007). Subtle linguistic cues affect children's motivation. *Psychological Science*, *18*(4), 314–316. https://doi.org/10.1111/j.1467-9280.2007.01896.x

Cohen, P. A., Kulik, J. A., & Kulik, C.-C. (1982). Education outcomes of tutoring: A meta-analysis of findings. *American Educational Research Journal*, *19*(2), 237–248. https://doi.org/10.3102/00028312019002237

Corno, L. (1986). The metacognitive control components of self-regulated learning. *Contemporary Educational Psychology*, *11*(4), 333–346. https://doi.org/10.1016/0361-476X(86)90029-9

Corno, L., & Mandinach, E. B. (1983). The role of cognitive engagement in classroom learning and motivation. *Educational Psychologist*, *18*(2), 88–108. https://doi.org/10.1080/00461528309529266

Cukurova, M., Avramides, K., Spikol, D., Luckin, R., & Mavrikis, M. (2016). An analysis framework for collaborative problem solving in practice-based learning activities: A mixed-method approach. In *Proceedings of the Sixth International Conference on Learning Analytics and Knowledge* (pp. 84–88). Association for Computing Machinery. https://doi.org/10.1145/2883851.2883900.

Cukurova, M., Khan-Galaria, M., Millán, E., & Luckin, R. (2022). A learning analytics approach to monitoring the quality of online one-to-one tutoring. *Journal of Learning Analytics*, *9*(2), 105–120. https://doi.org/10.18608/jla.2022.7411

Cullinane, C., & Montacute, R. (2023). *Tutoring – The new landscape*. Sutton Trust. https://www.suttontrust.com/wp-content/uploads/2023/03/Tutoring-The-New-Landscape.pdf

Darabi, A., & Jin, L. (2013). Improving the quality of online discussion: The effects of strategies designed based on cognitive load theory principles. *Distance Education*, *34*(1), 21–36. https://doi.org/10.1080/01587919.2013.770429

Data Protection Act 2018, c. 12 (2018). https://www.legislation.gov.uk/ukpga/2018/12/contents

De Corte, E., Verschaffel, L., & Masui, C. (2004). The CLIA-model: A framework for designing powerful learning environments for thinking and problem solving. *European Journal of Psychology of Education*, *19*, 365–384. https://doi.org/10.1007/BF03173216

de Medeiros, A. K. A., Guzzo, A., Greco, G., van der Aalst, W. M. P., Weijters, A. J. M., van Dongen, B. F., & Saccà, D. (2008). Process mining based on clustering: A quest for precision. In A. ter Hofstede, B. Benatallah, & H. Y. Paik (Eds.), *Business process management workshops: BPM 2007* (pp. 17–29). Springer. https://doi.org/10.1007/978-3-540-78238-4_4

Design-Based Research Collective. (2003). Design-based research: An emerging paradigm for educational inquiry. *Educational Researcher*, *32*(1), 5–8. https://doi.org/10.3102/0013189X032001005

Dewan, M. A. A., Murshed, M., & Lin, F. (2019). Engagement detection in online learning: A review. *Smart Learning Environments*, *6*, Article 1. https://doi.org/10.1186/s40561-018-0080-z

Di Gesú, M. G. (2021). Building intersubjectivity in online learning: Pupils' and university students' perception of teachers' social presence in technology-mediated teaching and learning processes. *Human Arenas*, *4*(2), 338–349. https://doi.org/10.1007/s42087-021-00226-x

Dignath, C., & Büttner, G. (2018). Teachers' direct and indirect promotion of self-regulated learning in primary and secondary school mathematics classes – Insights from video-based classroom observations and teacher interviews. *Metacognition and Learning, 13*(2), 127–157. https://doi.org/10.1007/s11409-018-9181-x

Dresel, M., Schmitz, B., Schober, B., Spiel, C., Ziegler, A., Engelschalk, T., Jöstl, G., Klug, J., Roth, A., Wimmer, B., & Steuer, G. (2015). Competencies for successful self-regulated learning in higher education: Structural model and indications drawn from expert interviews. *Studies in Higher Education*, *40*(3), 454–470. https://doi.org/10.1080/03075079.2015.1004236

Duffy, M. C., & Azevedo, R. (2015). Motivation matters: Interactions between achievement goals and agent scaffolding for self-regulated learning within an intelligent tutoring system. *Computers in Human Behavior*, *52*, 338–348. https://doi.org/10.1016/j.chb.2015.05.041

Dweck, C. (2012). *Mindset: How you can fulfil your potential*. Constable & Robinson. http://ebookcentral.proquest.com/lib/ucl/detail.action?docID=897458

Edelson, D. C. (2002). Design research: What we learn when we engage in design. *Journal of the Learning Sciences*, *11*(1), 105–121. https://doi.org/10.1207/S15327809JLS1101_4

ElSayed, A. A., Caeiro-Rodriguez, M., Mikic-Fonte, F., & Llamas-Nistal, M. (2019). Research in learning analytics and educational data mining to measure self-regulated learning: A systematic review. In *Proceedings of World Conference on Mobile and Contextual Learning 2019* (pp. 46–53). International Association for Mobile Learning.

Ewijk, C. D., Dickhäuser, O., & Büttner, G. (2013). Assessing how teachers enhance self-regulated learning: A multiperspective approach. *Journal of Cognitive Education and Psychology*, *12*(3), 338–358. https://doi.org/10.1891/1945-8959.12.3.338

Feyereisen, P. (2006). Further investigation on the mnemonic effect of gestures: Their meaning matters. *European Journal of Cognitive Psychology*, *18*(2), 185–205. https://doi.org/10.1080/09541440540000158

Fournier-Viger, P., Lin, J. C.-W., Vo, B., Chi, T. T., Zhang, J., & Le, H. B. (2017). A survey of itemset mining. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *7*(4), Article e1207. https://doi.org/10.1002/widm.1207

Gabadinho, A., Ritschard, G., Müller, N. S., & Studer, M. (2011). Analyzing and visualizing state sequences in R with TraMineR. *Journal of Statistical Software*, *40*(4), 1–37. https://doi.org/10.18637/jss.v040.i04

Gillespie, L. (2015). It takes two: The role of co-regulation in building self-regulation skills. *Young Children*, *70*(3), 94–96. https://www.naeyc.org/resources/pubs/yc/jul2015/rocking-rolling

Graesser, A. C., Conley, M. W., & Olney, A. (2012). Intelligent tutoring systems. In K. R. Harris, S. Graham, T. Urdan, A. G. Bus, S. Major, & H. L. Swanson (Eds.), *APA educational psychology handbook, Vol. 3. Application to learning and teaching* (pp. 451–473). American Psychological Association. https://doi.org/10.1037/13275-018

Graesser, A. C., Person, N. K., & Magliano, J. P. (1995). Collaborative dialogue patterns in naturalistic one-to-one tutoring. *Applied Cognitive Psychology*, *9*(6), 495–522. https://doi.org/10.1002/acp.2350090604

Graesser, A., & McNamara, D. (2010). Self-regulated learning in learning environments with pedagogical agents that interact in natural language. *Educational Psychologist*, *45*(4), 234–244. https://doi.org/10.1080/00461520.2010.515933

Grafsgaard, J., Wiggins, J., Boyer, K., Wiebe, E., & Lester, J. (2013). Automatically recognizing facial indicators of frustration: A learning-centric analysis. In *2013 Humaine Association Conference on Affective Computing and Intelligent Interaction* (pp. 159–165). IEEE. https://doi.org/10.1109/ACII.2013.33

Greene, J. A., & Azevedo, R. (2007). A theoretical review of Winne and Hadwin's model of self-regulated learning: New perspectives and directions. *Review of Educational Research*, *77*(3), 334–372. https://doi.org/10.3102/003465430303953

Greene, J. A., & Azevedo, R. (2010). The measurement of learners' self-regulated cognitive and metacognitive processes while using computer-based learning environments. *Educational Psychologist*, *45*(4), 203–209. https://doi.org/10.1080/00461520.2010.515935

Greene, J. A., Muis, K. R., & Pieschl, S. (2010). The role of epistemic beliefs in students' self-regulated learning with computer-based learning environments: Conceptual and methodological issues. *Educational Psychologist*, *45*(4), 245–257. https://doi.org/10.1080/00461520.2010.515932

Hadwin, A. F., Järvelä, S., & Miller, M. (2011). Self-regulated, co-regulated, and socially shared regulation of learning. In B. J. Zimmerman & D. H. Schunk (Eds.), *Handbook of self-regulation of learning and performance* (pp. 65–84). Routledge.

Hadwin, A., Järvelä, S., & Miller, M. (2017). Self-regulation, co-regulation, and shared regulation in collaborative learning environments. In D. H. Schunk & J. A. Greene (Eds.), *Handbook of self-regulation of learning and performance* (2nd ed., pp. 83–106). Routledge. https://doi.org/10.4324/9781315697048-6

Hadwin, A. F., & Winne, P. H. (2001). CoNoteS2: A software tool for promoting self-regulation. *Educational Research and Evaluation*, *7*(2–3), 313–334. https://doi.org/10.1076/edre.7.2.313.3868

Hattie, J. (2008a). *A synthesis of over 800 meta-analyses relating to achievement*. Routledge. https://doi.org/10.4324/9780203887332

Hattie, J. (2008b). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. Taylor & Francis Group. http://ebookcentral.proquest.com/lib/ucl/detail.action?docID=367685

Heritage, M., & Heritage, J. (2013). Teacher questioning: The epicenter of instruction and assessment. *Applied Measurement in Education*, *26*(3), 176–190. https://doi.org/10.1080/08957347.2013.793190

Hrastinski, S., Stenbom, S., Benjaminsson, S., & Jansson, M. (2021). Identifying and exploring the effects of different types of tutor questions in individual online synchronous tutoring in mathematics. *Interactive Learning Environments*, *29*(3), 510–522. https://doi.org/10.1080/10494820.2019.1583674

Hurme, T.-R., & Järvelä, S. (2005). Students' activity in computer supported collaborative problem solving in mathematics. *International Journal of Computers for Mathematical Learning*, *10*, 49–73. https://doi.org/10.1007/s10758-005-4579-3

Järvelä, S., & Hadwin, A. F. (2013). New frontiers: Regulating learning in CSCL. *Educational Psychologist*, *48*(1), 25–39. https://doi.org/10.1080/00461520.2012.748006

Järvelä, S., Kirschner, P. A., Hadwin, A., Järvenoja, H., Malmberg, J., Miller, M., & Laru, J. (2016). Socially shared regulation of learning in CSCL: Understanding and prompting individual- and group-level shared regulatory activities. *International Journal of Computer-Supported Collaborative Learning*, *11*, 263–280. https://doi.org/10.1007/s11412-016-9238-2

Järvelä, S., Volet, S., & Järvenoja, H. (2010). Research on motivation in collaborative learning: Moving beyond the cognitive-situative divide and combining individual and social processes. *Educational Psychologist*, *45*(1), 15–27. https://doi.org/10.1080/00461520903433539

Jovanović, J., Gasevic, D., Dawson, S., Pardo, A., & Mirriahi, N. (2017). Learning analytics to unveil learning strategies in a flipped classroom. *The Internet and Higher Education*, *33*, 74–85. https://doi.org/10.1016/j.iheduc.2017.02.001

Juhaňák, L., Zounek, J., Záleská, K., Bárta, O., & Vlčková, K. (2019). The relationship between the age at first computer use and students' perceived competence and autonomy in ICT usage: A mediation analysis. *Computers & Education*, *141*, Article 103614. https://doi.org/10.1016/j.compedu.2019.103614

Kinnebrew, J. S. Loretz, K. M., & Biswas, G. (2013). A contextualized, differential sequence mining method to derive students' learning behavior patterns. *Journal of Educational Data Mining*, *5*(1), 190–219.

Kizilcec, R., Pérez-Sanagustín, M., & Maldonado, J. J. (2017). Self-regulated learning strategies predict learner behavior and goal attainment in massive open online courses. *Computers & Education*, *104*, 18–33. https://doi.org/10.1016/j.compedu.2016.10.001

Kuvalja, M., Verma, M., & Whitebread, D. (2014). Patterns of co-occurring non-verbal behavior and self-directed speech; a comparison of three methodological approaches. *Metacognition and Learning*, *9*, 87–111. https://doi.org/10.1007/s11409-013-9106-7

Lim, J., Hall, B. M., Jeong, A. C., & Freed, S. (2017). Intersubjectivity and discussion characteristics in online courses. *Quarterly Review of Distance Education*, *18*(1), 29–45.

Lin, X., & Lehman, J. (1999). Supporting learning of variable control in a computer-based biology environment: Effects of prompting college students to reflect on their own thinking. *Journal of Research in Science Teaching*, *36*, 837–858. https://doi.org/10.1002/(SICI)1098-2736(199909)36:7<837::AID-TEA6>3.0.CO;2-U

Lipnevich, A. A., & Smith, J. K. (2009). “I really need feedback to learn”: Students’ perspectives on the effectiveness of the differential feedback messages. *Educational Assessment, Evaluation and Accountability*, *21*(4), 347–367. https://doi.org/10.1007/s11092-009-9082-2

Mace, F. C., & Kratochwill, T. R. (1985). Theories of reactivity in self-monitoring: A comparison of cognitive-behavioral and operant models. *Behavior Modification*, *9*(3), 322–343 https://doi.org/10.1177/01454455850093003

Mäkitalo, K., Kohnle, C., & Fischer, F. (2011). Computer-supported collaborative inquiry learning and classroom scripts: Effects on help-seeking processes and learning outcomes. *Learning and Instruction*, *21*(2), 257–266. https://doi.org/10.1016/j.learninstruc.2010.07.001

Maldonado-Mahauad, J., Pérez-Sanagustín, M., Kizilec, R. F., Morales, N., & Munoz-Gama, J. (2018). Mining theory-based patterns from big data: Identifying self-regulated learning strategies in massive open online courses. *Computers in Human Behavior*, *80*, 179–196. https://doi.org/10.1016/j.chb.2017.11.011

Malmberg, J., Järvelä, S., & Järvenoja, H. (2017). Capturing temporal and sequential patterns of self-, co-, and socially shared regulation in the context of collaborative learning. *Contemporary Educational Psychology*, *49*, 160–174. https://doi.org/10.1016/j.cedpsych.2017.01.009

Matcha, W., Gašević, D., Ahmad Uzir, N., Jovanović, J., Pardo, A., Maldonado-Mahauad, J., & Pérez-Sanagustín, M. (2019). Detection of learning strategies: A comparison of process, sequence and network analytic approaches. In M. Scheffel, J. Broisin, V. Pammer-Schindler, A. Ioannou, & J. Schneider (Eds.), *Transforming learning with meaningful technologies* (pp. 525–540). Springer. https://doi.org/10.1007/978-3-030-29736-7_39

Matcha, W., Ahmad Uzir, N., Gasevic, D., & Pardo, A. (2020). A systematic review of empirical studies on learning analytics dashboards: A self-regulated learning perspective. *IEEE Transactions on Learning Technologies*, *13*(2), 226–245. https://doi.org/10.1109/TLT.2019.2916802.

McCaslin, M., & Good, T. (1996). *Listening in classrooms*. Harper Collins.

McCaslin, M., & Hickey, D. T. (2001). Self-regulated learning and academic achievement: A Vygotskian view. In B. J. Zimmerman & D. H. Schunk (Eds.), *Self-regulated learning and academic achievement: Theoretical perspectives* (pp. 227–252). Lawrence Erlbaum Associates.

McCombs, B. L. (1986). The role of the self-system in self-regulated learning. *Contemporary Educational Psychology*, *11*(4), 314–332. https://doi.org/10.1016/0361-476X(86)90028-7

McKenney, S., & Reeves, T. C. (2013). Educational design research. In J. Spector, M. Merrill, J. Elen, & M. Bishop (Eds.), *Handbook of research on educational communications and technology* (pp. 131–140). Springer. https://doi.org/10.1007/978-1-4614-3185-5_11

Meyer, D. K., & Turner, J. C. (2002). Using instructional discourse analysis to study the scaffolding of student self-regulation. *Educational Psychologist*, *37*(1), 17–25. https://doi.org/10.1207/S15326985EP3701_3

Miller, M., & Hadwin, A. (2015). Scripting and awareness tools for regulating collaborative learning: Changing the landscape of support in CSCL. *Computers in Human Behavior*, *52*, 573–588. https://doi.org/10.1016/j.chb.2015.01.050

Molenaar, I., & Chiu, M. M. (2014). Dissecting sequences of regulation and cognition: Statistical discourse analysis of primary school children's collaborative learning. *Metacognition and Learning*, *9*(2), 137–160. https://doi.org/10.1007/s11409-013-9105-8

Moos, D. C., & Azevedo, R. (2008). Self-regulated learning with hypermedia: The role of prior domain knowledge. *Contemporary Educational Psychology*, *33*(2), 270–298. https://doi.org/10.1016/j.cedpsych.2007.03.001

O'Neill, G., & McMahon, T. (2005). Student-centred learning: What does it mean for students and lecturers? In G. O'Neill, S. Moore, & B. McMullin (Eds.), *Emerging issues in the practice of university learning and teaching* (pp. 30–39). AISHE.

Paas, F., Tuovinen, J., Tabbers, H., & Van Gerven, P. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist*, *38*(1), 63–71. https://doi.org/10.1207/S15326985EP3801_8

Panadero, E., Klug, J., & Järvelä, S. (2016). Third wave of measurement in the self-regulated learning field: When measurement and intervention come hand in hand. *Scandinavian Journal of Educational Research*, *60*(6), 723–735 https://doi.org/10.1080/00313831.2015.1066436

Paris, S. G., & Paris, A. H. (2001). Classroom applications of research on self-regulated learning. *Educational Psychologist*, *36*(2), 89–101. https://doi.org/10.1207/S15326985EP3602_4

Perry, N. E., VandeKamp, K. O., Mercer, L. K., & Nordby, C. J. (2002). Investigating teacher-student interactions that foster self-regulated learning. *Educational Psychologist*, *37*(1), 5–15. https://doi.org/10.1207/S15326985EP3701_2

Phyu, T. N. (2009). Survey of classification techniques in data mining. In S. I. Ao, O. Castillo, C. Douglas, D. D. Feng, & J.-A. Lee (Eds.), *Proceedings of the 2009 International MultiConference of Engineers and Computer Scientists* (pp. 18–20). Newswood.

Pintrich, P. R. (2000). The role of goal orientation in self-regulated learning. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 451–502). Academic Press. https://doi.org/10.1016/B978-012109890-2/50043-3

Pintrich, P. R., Smith, D. A., Garcia, T., & McKeachie, W. J. (1993). Reliability and predictive validity of the Motivated Strategies for Learning Questionnaire (MSLQ). *Educational and Psychological Measurement*, *53*(3), 801–813. https://doi.org/10.1177/0013164493053003024

Pintrich, P. R., & Zusho, A. (2002). The development of academic self-regulation: The role of cognitive and motivational factors. In A. Wigfield & J. S. Eccles (Eds.), *Development of achievement motivation* (pp 249–284). Academic Press. https://doi.org/10.1016/B978-012750053-9/50012-7

Prieto, L. P., Sharma, K., Dillenbourg, P., & Jesús, M. (2016). Teaching analytics: Towards automatic extraction of orchestration graphs using wearable sensors. *Proceedings of the Sixth International Conference on Learning Analytics and Knowledge* (pp. 148–157). Association for Computing Machinery. https://doi.org/10.1145/2883851.2883927

Prieto, L. P., Sharma, K., Kidzinski, Ł., Rodríguez-Triana, M. J., & Dillenbourg, P. (2018). Multimodal teaching analytics: Automated extraction of orchestration graphs from wearable sensor data. *Journal of Computer Assisted Learning*, *34*(2), 193–203. https://doi.org/10.1111/jcal.12232

Psaltis, A., Apostolakis, K., Dimitropoulos, K., & Daras, P. (2018). Multimodal student engagement recognition in prosocial games. *IEEE Transactions on Games*, *10*(3), 292–303. https://doi.org/10.1109/TCIAIG.2017.2743341

Reeves, T., Herrington, J., & Oliver, R. (2005). Design research: A socially responsible approach to instructional technology research in higher education. *Journal of Computing in Higher Education*, *16*, 96–115. https://doi.org/10.1007/BF02961476

Reimann, P., Frerejean, J., & Thompson, K. (2009). Using process mining to identify models of group decision making in chat data. In C. O'Malley, D. Suthers, P. Reimann, & A. Dimitracopoulou (Eds.), *Proceedings of the 9th International Conference on Computer Supported Collaborative Learning* (pp. 98–107). International Society of the Learning Sciences. https://doi.org/10.3115/1600053.1600068

Riel, J., Lawless, K. A., & Brown, S. W. (2018). Timing matters: Approaches for measuring and visualizing behaviours of timing and spacing of work in self-paced online teacher professional development courses. *Journal of Learning Analytics*, *5*(1), 25–40. https://doi.org/10.18608/jla.2018.51.3

Rogat, T., & Linnenbrink-Garcia, L. (2011). Socially shared regulation in collaborative groups: An analysis of the interplay between quality of social regulation and group processes. *Cognition and Instruction*, *29*(4), 375–415. https://doi.org/10.1080/07370008.2011.607930

Roll, I., & Winne, P. H. (2015). Understanding, evaluating, and supporting self-regulated learning using learning analytics. *Journal of Learning Analytics*, *2*(1), 7–12. https://doi.org/10.18608/jla.2015.21.2

Romero, C., & Ventura, S. (2010). Educational data mining: A review of the state of the art. systems, man, and cybernetics. *IEEE Transactions on Systems, Man, and Cybernetics. Part C, Applications and Reviews*, *40*(6), 601–618. https://doi.org/10.1109/TSMCC.2010.2053532

Roscoe, R., Segedy, J., Sulcer, B., & Biswas, G. (2013). Shallow strategy development in a teachable agent environment designed to support self-regulated learning. *Computers & Education*, *62*, 286–297. https://doi.org/10.1016/j.compedu.2012.11.008

Rosenthal, R., & Jacobson, L. (1968). *Pygmalion in the classroom: Teacher expectation and pupils' intellectual development*. Holt, Rinehart & Winston.

Saariaho, E., Toom, A., Soini, T., Pietarinen, J. (2019). Student teachers' and pupils' co-regulated learning behaviours in authentic classroom situations in teaching practicums. *Teaching and Teacher Education*, *85*, 92–104. https://doi.org/10.1016/j.tate.2019.06.003

Saint, J., Fan, Y., Singh, S., Gasevic, D., & Pardo, A. (2021). Using process mining to analyse regulated learning: A systematic analysis of four algorithms. In *LAK21: 11th International Learning Analytics and Knowledge Conference* (pp. 333–343). Association for Computing Machinery. https://doi.org/10.1145/3448139.3448171

Schellings, G., & Van Hout-Wolters, B. (2011). Measuring strategy use with self-report instruments: Theoretical and empirical considerations. *Metacognition and Learning*, *6*(2), 83–90. https://doi.org/10.1007/s11409-011-9081-9

Schmitz, B., & Perels, F. (2011). Self-monitoring of self-regulation during math homework behaviour using standardized diaries. *Metacognition and Learning*, *6*(3), 255–273. https://doi.org/10.1007/s11409-011-9076-6

Schmitz, B., & Wiese, B. S. (2006). New perspectives for the evaluation of training sessions in self-regulated learning: Time-series analyses of diary data. *Contemporary Educational Psychology*, *31*(1), 64–96. https://doi.org/10.1016/j.cedpsych.2005.02.002

Schneider, M. C., & Andrade, H. (2013). Teachers' and administrators' use of evidence of student learning to take action: Conclusions drawn from a special issue on formative assessment. *Applied Measurement in Education*, *26*(3), 159–162. https://doi.org/10.1080/08957347.2013.793189

Schoor, C., & Bannert, M. (2011). Motivation in a computer-supported collaborative learning scenario and its impact on learning activities and knowledge acquisition. *Learning and Instruction*, *21*(4), 560–573. https://doi.org/10.1016/j.learninstruc.2010.11.002

Schunk, D. H. (1986). Verbalization and children's self-regulated learning. *Contemporary Educational Psychology*, *11*(4), 347–369. https://doi.org/10.1016/0361-476X(86)90030-5

Schwartz, D. L., & Bransford, J. D. (1998). A time for telling. *Cognition and Instruction*, *16*(4), 475–522. https://doi.org/10.1207/s1532690xci1604_4

Sedrakyan, G., Weerdt, J., & Snoeck, M. (2016). Process-mining enabled feedback: "Tell me what I did wrong" vs. "tell me how to do it right." *Computers in Human Behavior*, *57*, 352–376. https://doi.org/10.1016/j.chb.2015.12.040

Snow, R. E. (1991). Aptitude-treatment interaction as a framework for research on individual differences in psychotherapy. *Journal of Consulting and Clinical Psychology*, *59*(2), 205–216. https://doi.org/10.1037/0022-006X.59.2.205

Spates, C. R., & Kanfer, F. H. (1977). Self-monitoring, self-evaluation, and self-reinforcement in children's learning: A test of a multistage self-regulation model. *Behavior Therapy*, *8*(1), 9–16. https://doi.org/10.1016/S0005-7894(77)80115-9

Stahnke, R., Schueler, S., & Roesken-Winter, B. (2016). Teachers' perception, interpretation, and decision-making: A systematic review of empirical mathematics education research. *ZDM – Mathematics Education*, *48*(1), 1–27. https://doi.org/10.1007/s11858-016-0775-y

Steffe, L. P., & Thompson, P. W. (2000). Interaction or intersubjectivity? A reply to Lerman. *Journal for Research in Mathematics Education*, *31*(2), 191–209. https://doi.org/10.2307/749751

Taub, M., & Azevedo, R. (2019). How does prior knowledge influence eye fixations and sequences of cognitive and metacognitive SRL processes during learning with an Intelligent Tutoring System? *International Journal of Artificial Intelligence in Education*, *29*(1), 1–28. https://doi.org/10.1007/s40593-018-0165-4

Taub, M., Azevedo, R., Bradbury, A., Garrett, C., & Millar, J. (2018). Using sequence mining to reveal the efficiency in scientific reasoning during STEM learning with a game-based learning environment. *Learning and Instruction*, *54*, 93–103. https://doi.org/10.1016/j.learninstruc.2017.08.005

Trevors, G., Duffy, M., & Azevedo, R. (2014). Note-taking within MetaTutor: Interactions between an intelligent tutoring system and prior knowledge on note-taking and learning. *Educational Technology Research and Development*, *62*(5), 507–528. https://doi.org/10.1007/s11423-014-9343-8

van der Aalst, W. (2012). Process mining: Overview and opportunities. *ACM Transactions on Management Information Systems*, *3*(2), 1–17. https://doi.org/10.1145/2229156.2229157

van der Aalst, W., Adriansyah, A., de Medeiros, A. K. A., Arcieri, F., Baier, T., Blickle, T., Bose, J. C., van den Brand, P., Brandtjen, R., Buijs, J., Burattin, A., Carmona, J., Castellanos, M., Claes, J., Cook, J., Costantini, N., Curbera, F., Damiani, E., de Leoni, M., … Wynn, M. (2011). Process Mining Manifesto. In F. Daniel, K. Barkaoui, & S. Dustdar (Eds.), *Business process management workshops: BPM 2011 international workshops, Clermont-Ferrand, France, August 29, 2011, revised selected papers, Part I* (pp. 169–194). Springer. https://doi.org/10.1007/978-3-642-28108-2_19

van den Akker, J. (1999). Principles and methods of development research. In J. van den Akker, R. M. Branch, K. Gustafson, N. Nieveen, & T. Plomp (Eds.), *Design approaches and tools in education and training* (pp. 1–14). Kluwer. https://doi.org/10.1007/978-94-011-4255-7_1

van den Bergh, L., Denessen, E., Hornstra, L., Voeten, M., & Holland, R. W. (2010). The implicit prejudiced attitudes of teachers: Relations to teacher expectations and the ethnic achievement gap. *American Educational Research Journal*, *47*(2), 497–527. https://doi.org/10.3102/0002831209353594

Van Laer, S., & Elen, J. (2018). Towards a methodological framework for sequence analysis in the field of self-regulated learning. *Frontline Learning Research*, *16*, 228–249. https://doi.org/10.14786/flr.v6i3.367

VanLehn, K., Graesser, A. C., Jackson, G. T., Jordan, P., Olney, A., & Rosé, C. P. (2007). When are tutorial dialogues more effective than reading? *Cognitive Science*, *31*(1), 3–62. https://doi.org/10.1080/03640210709336984

Veenman, M. V. J. (2007). The assessment and instruction of self-regulation in computer-based environments: A discussion. *Metacognition and Learning*, *2*, 177–183. https://doi.org/10.1007/s11409-007-9017-6

Veenman, M. V. J., Van Hout-Wolters, B. H. A. M., & Afflerbach, P. (2006). Metacognition and learning: Conceptual and methodological considerations. *Metacognition and Learning*, *1*, 3–14. https://doi.org/10.1007/s11409-006-6893-0

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Wang, M. C. (1983). Development and consequences of students' sense of personal control. In J. M. Levine & M. C. Wang (Eds.), *Teacher and student perceptions: Implications for learning* (pp. 213–247). Erlbaum.

Wang, M. C., & Peverly, S. T. (1986). The self-instructive process in classroom learning contexts. *Contemporary Educational Psychology*, *11*(4), 370–404. https://doi.org/10.1016/0361-476X(86)90031-7

Weinstein, C. E., Schulte, A. C., & Palmer, D. P. (1987). *Learning and study strategies inventory*. H & H Publishing.

Winne, P. H. (2010). Improving measurements of self-regulated learning. *Educational Psychologist*, *45*(4), 267–276. https://doi.org/10.1080/00461520.2010.517150

Winne, P. H. (2017a). Cognition and metacognition within self-regulated learning. In D. H. Schunk & J. A. Greene (Eds.), *Handbook of self-regulation of learning and performance* (2nd ed., pp. 36–48). Routledge. https://doi.org/10.4324/9781315697048-3

Winne, P. H. (2017b). Leveraging big data to help each learner and accelerate learning science. Teachers College Record, *119*(3), 1–24. https://doi.org/10.1177/016146811711900305

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

Winne, P. H., & Baker, R. (2013). The potentials of educational data mining for researching metacognition, motivation and self-regulated learning. *Journal of Educational Data Mining*, *5*, 1–8. https://doi.org/10.5281/zenodo.3554619

Winne, P. H., & Hadwin, A. F. (1998). Studying as self-regulated learning. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 277–304). Routledge. https://doi.org/10.4324/9781410602350-19

Winne, P. & Hadwin, A. F. (2013). nStudy: Tracing and supporting self-regulated learning in the internet. In R. Azevedo & V. Aleven (Eds.), *International handbook of metacognition and learning technologies* (pp. 293–308). Springer. https://doi.org/10.1007/978-1-4419-5546-3_20

Winne, P. H., Hadwin, A. F., & Perry, N. E. (2013). Metacognition and computer-supported collaborative learning. In C. Hmelo-Silver, C. Chinn, C. Chan, & A. O'Donnell (Eds.), *The international handbook of collaborative learning* (pp. 462–479). Routledge. https://doi.org/10.4324/9780203837290.ch26

Xing, W., & Du, D. (2018). Dropout prediction in MOOCs: Using deep learning for personalized intervention. *Journal of Educational Computing Research*, *57*(3), 547–570. https://doi.org/10.1177/0735633118757015

Zhou, M., & Winne, P. H. (2012). Modeling academic achievement by self-reported versus traced goal orientation. *Learning and Instruction*, *22*(6), 413–419. https://doi.org/10.1016/j.learninstruc.2012.03.004

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

Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13–39). https://doi.org/10.1016/B978-012109890-2/50031-7

Zimmerman, B. J. (2008). Investigating self-regulation and motivation: Historical background, methodological developments, and future prospects. *American Educational Research Journal*, *45*(1), 166–183. https://doi.org/10.3102/0002831207312909

Zimmerman, B. J., & Kitsantas, A. (1996). Self-regulated learning of a motoric skill: The role of goal setting and self-monitoring. *Journal of Applied Sport Psychology*, *8*(1), 60–75. https://doi.org/10.1080/10413209608406308

Zimmerman, B. J., & Martinez-Pons, M. (1988). Construct validation of a strategy model of student self-regulated learning. *Journal of Educational Psychology*, *80*(3), 284–290. https://doi.org/10.1037/0022-0663.80.3.284