



THE UNIVERSITY *of* EDINBURGH

This thesis has been submitted in fulfilment of the requirements for a postgraduate degree (e.g. PhD, MPhil, DClinPsychol) at the University of Edinburgh. Please note the following terms and conditions of use:

This work is protected by copyright and other intellectual property rights, which are retained by the thesis author, unless otherwise stated.

A copy can be downloaded for personal non-commercial research or study, without prior permission or charge.

This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author.

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author.

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given.

ANALYTICS OF STUDENT INTERACTIONS: TOWARDS
THEORY-DRIVEN, ACTIONABLE INSIGHTS

O. E. D. Fincham



Doctor of Philosophy
Institute for Adaptive Neural Computation
School of Informatics
University of Edinburgh
2020

O. E. D. Fincham

Analytics of Student Interactions: Towards Theory-driven, Actionable Insights

Doctor of Philosophy, 2020

PRINCIPAL SUPERVISOR:

Professor Dragan Gašević, Ph.D., *Monash University, Australia.*

ASSISTANT SUPERVISORS:

Professor Jelena Jovanović, Ph.D., *University of Belgrade, Serbia.*

Professor Shane Dawson, Ph.D., *University of South Australia, Australia.*

Yi-Shan Tsai, Ph.D., *University of Edinburgh, United Kingdom.*

EXAMINATION COMMITTEE:

Christopher Lucas, Ph.D., *University of Edinburgh, United Kingdom.*

Professor Bart Rienties, Ph.D., *Open University, United Kingdom*

*“So moving for the one who writes,
yet so boring for the one who reads.”*

Abstract

The field of learning analytics arose as a response to the vast quantities of data that are increasingly generated about students, their engagement with learning resources, and their learning and future career outcomes. While the field began as a collage, adopting methods and theories from a variety of disciplines, it has now become a major area of research, and has had a substantial impact on practice, policy, and decision-making.

Although the field supports the collection and analysis of a wide array of data, existing work has predominantly focused on the digital traces generated through interactions with technology, learning content, and other students. Yet for any analyses to support students and teachers, the measures derived from these data must (1) offer practical and actionable insight into learning processes and outcomes, and (2) be theoretically grounded. As the field has matured, a number of challenges related to these criteria have become apparent. For instance, concerns have been raised that the literature prioritises predictive modeling over ensuring that these models are capable of informing constructive actions. Furthermore, the methodological validity of much of this work has been challenged, as a swathe of recent research has found many of these models fail to replicate to novel contexts.

The work presented in this thesis addresses both of these concerns. In doing so, our research is pervaded by three key concerns: firstly, ensuring that any measures developed are both structurally valid and generalise across contexts; secondly, providing actionable insight with regards to student engagement; and finally, providing representations of student interactions that are predictive of student outcomes, namely, grades and students' persistence in their studies. This research programme is heavily indebted to the work of Vincent Tinto, who conceptually distinguishes between the interactions students have with the academic and social domains present within their educational institution. This model has been subjected to extensive empirical validation, using a range of methods and data. For instance, while some studies have relied upon survey responses, others have used social network metrics, demographic variables, and students' time spent in class together to evaluate Tinto's claims. This model provides a foundation for the thesis, and the work presented may be categorised into two distinct veins aligning with the academic and social aspects of integration that Tinto proposes. These two domains, Tinto argues, continually modify a student's

goals and commitments, resulting in persistence or eventual disengagement and dropout.

In the former, academic domain, we present a series of novel methodologies developed for modeling student engagement with academic resources. In doing so, we assessed how an individual student's behaviour may be modeled using hidden Markov models (HMMs) to provide representations that enable actionable insight. However, in the face of considerable individual differences and cross-course variation, the validity of such methods may be called into question. Accordingly, ensuring that any measurements of student engagement are both structurally valid, and generalise across course contexts and disciplines became a central concern. To address this, we developed our model of student engagement using sticky-HMMs, emphasised the more interpretable insight such an approach provides compared to competing models, demonstrated its cross-course generality, and assessed its structural validity through the successful prediction of student dropout.

In the social domain, a critical concern was to ensure any analyses conducted were valid. Accordingly, we assessed how the diversity of social tie definitions may undermine the validity of subsequent modeling practices. We then modeled students' social integration using graph embedding techniques, and found that not only are student embeddings predictive of their final grades, but also of their persistence in their educational institution.

In keeping with Tinto's model, our research has focused on academic and social interactions separately, but both avenues of investigation have led to the question of student disengagement and dropout, and how this may be represented and remedied through the provision of actionable insight.

Lay Summary

This thesis presents novel methods that address a number of the key challenges currently facing the field of learning analytics, namely: the validity, theoretical foundations, and interpretability of any analysis. In doing so, our research is framed by the work of Vincent Tinto, who posits that the social and academic interactions that students have with their educational institution influence not only their academic performance, but also their persistence in their studies. Taking inspiration from this model, the work presented in this thesis not only provides valid, theoretically-grounded analyses, but also scales to the vast quantities of digital data that educational institutions generate. Finally, the methods we introduce are capable of distilling complex student data into interpretable insight, such as the identification of at-risk students.

Acknowledgment

“There was no need to praise anyone for writing a book, since it was always done by somebody else.”

Dragan Gašević

Jelena Jovanović

Shane Dawson

Yi-Shan Tsai

Vitomir Kovanović

Srećko Joksimović

Nick Hoernle

Alexander Whitelock-Wainwright

Benedek Rózemberczki

Paul Fincham

Alison Fincham

Kenneth Fincham

Henry Fincham

Declaration of authorship

I declare that this thesis has been composed solely by myself and that it has not been submitted, either in whole or in part, in any previous application for a degree. Except where otherwise acknowledged, the work presented is entirely my own. This thesis also includes five peer-reviewed publications produced under the joint authorship:

- (1) Fincham, E., Gašević, D., & Pardo, A. (2018). From social ties to network processes: Do tie definitions matter? *Journal of Learning Analytics*, 5(2), 9–28. doi:10.18608/jla.2018.52.2
- (2) Fincham, E., Rózemberczki, B., Kovanović, V., Joksimović, S., & Gašević, D. (2019). Persistence and performance in co-enrolment network embeddings. *Manuscript submitted for publication*
- (3) Fincham, E., Gašević, D., Jovanović, J., & Pardo, A. (2019). From study tactics to learning strategies: An analytical method for extracting interpretable representations. *IEEE Transactions on Learning Technologies*, 12(1), 59–72. doi:10.1109/TLT.2018.2823317
- (4) Fincham, E., Whitelock-Wainwright, A., Kovanović, V., Joksimović, S., van Staalduinen, J.-P., & Gašević, D. (2019). Counting clicks is not enough: Validating a theorized model of engagement in learning analytics. In *Proceedings of the 9th international conference on learning analytics & knowledge* (pp. 501–510). LAK19. Tempe, AZ, USA: ACM. doi:10.1145/3303772.3303775
- (5) Fincham, E., Hoernle, N., Gal, K., & Gašević, D. (2019). The road not taken: Preempting dropout in moocs. *Manuscript submitted for publication*

I declare that I substantially contributed to all five publications (i.e., over 50% of the work done) and was involved in all phases of the research process, including study conceptualisation, data collection, data analysis and interpretation, as well as the writing of the final publication.

O. E. D. Fincham

Table of Contents

Abstract	iv
Lay Summary	vi
Acknowledgement	vii
Declaration	ix
1 Introduction	1
1.1 Research goals and questions	4
1.2 Methodology	5
1.3 Thesis structure and overview	6
1.4 Ethics approval	11
2 The Validity of Social Ties	13
2.1 Introduction	13
2.2 Publication: From Social Ties to Network Processes: Do Tie Definitions Matter? . . .	14
2.3 Summary	41
3 Performance and Persistence in Co-enrolment Networks	42
3.1 Introduction	42
3.2 Publication: Persistence and Performance in Co-enrolment Network Embeddings . . .	44
3.3 Summary	62
4 Interpretable Representations of Student Behaviour	63
4.1 Introduction	63
4.2 Publication: From Study Tactics to Learning Strategies: An Analytical Method for Extracting Interpretable Representations	65
4.3 Summary	80

5	Validating a Theoretical Model of Student Engagement	81
5.1	Introduction	81
5.2	Publication: Counting Clicks is Not Enough: Validating a Theorized Model of Engagement in Learning Analytics	82
5.3	Summary	93
6	Towards Interpretable Insight	94
6.1	Introduction	94
6.2	Publication: The Road Not Taken: Preempting Dropout in MOOCs	95
6.3	Summary	131
7	Conclusions	132
7.1	Impact of the Present Work	132
7.2	Implications for Research and Practice	135
7.3	Conclusions	138
	Bibliography	139

1

Introduction

OVER the last two decades, the number of students enrolled in higher education has more than doubled to over 200 million (Calderon, 2018). This growth is set to continue and, by 2040, enrolments are anticipated to rise to over 600 million (Calderon, 2018). This growing student body has placed increased demand on limited academic and administrative resources such as teaching hours. At the same time, tuition fees across a number of higher education institutions have been steadily increasing (Altbach & Reisberg, 2018), which risks placing these institutions in the unenviable position of charging more while offering less. This is at odds with the expectations of students who, since the introduction of fees, have increasingly viewed themselves as consumers (Rolfe, 2001; Kandiko & Mawer, 2013; Tomlinson, 2017), prompting institutions to operate under the forces of marketisation which demand competitiveness, efficiency, and consumer satisfaction (Lesnik-Oberstein, 2015; Gunn, 2018). Accordingly, higher education institutions are under pressure not only to promote student outcomes such as employability and life-long learning (Moore & Morton, 2017; McCowan, 2017), but also improve the student experience through personalised support and learning (McCowan, 2017).

Given their limited resources, higher education institutions have shown considerable interest in learning analytics as a means to alleviate the pressures of increased student numbers and expectations (Siemens, 2013). Such approaches rely on “the unreasonable effectiveness of data” (Halevy, Norvig, & Pereira, 2009), and have prompted institutions to collect and analyse the thousands of digital traces that students generate through their programme enrolments and interactions with learning management systems (Siemens, 2013). While this vast collection of data promises much, it would be misguided to consider it a panacea; digital traces are inherently ambiguous and requires additional exploration and theorising in order to understand what, for instance, an extended pause of reading means (Siemens, 2013). The challenge, then, is to leverage this information to provide students and instructors with practical insight into learning processes and outcomes, whilst ensuring that any insights are both valid and theoretically-grounded (Gašević, Dawson, & Siemens, 2015).

In reviewing the educational research literature, there are a number of theoretical approaches which emphasise the importance of students’ interactions. For instance, a large number of studies within Computer-Supported Collaborative Learning (CSCL) have assessed the relation between

student interactions, learning processes, and student outcomes in virtual environments (Rienties, Tempelaar, den Bossche, Gijsselaers, & Segers, 2009). This research has indicated that students' interactions with learning materials and their peers are mediated by their motivations (Ryan & Deci, 2000; Fairchild, Horst, Finney, & Barron, 2005). In particular, the quality and content of students' discourse is associated with the degree of their motivation (Rienties et al., 2009). Alternatively, the literature surrounding shared mental models posits that students' interactions within groups is the source of knowledge construction (den Bossche, Gijsselaers, Segers, & Kirschner, 2006). This collaboration, however, does not occur without intent but requires a conscious, continued effort on the part of students (Roschelle & Teasley, 1995). In addition, the communities of inquiry model emphasises the importance of students' interactions in the development of students' critical thinking skills (Shea & Bidjerano, 2010). The model decomposes online learning into three "presences" which include *social presence*, or students' interactions and the social climate of the course (Rourke, Anderson, Garrison, & Archer, 1999). These models, however, emphasise learning and engagement, but pay limited attention to student retention. Within the context of educational institutions, the promotion of both student engagement and student retention are central challenges (Siemens, 2013).

While there has been ample theoretical and empirical discussion on the causes of dropout (Thomas, 2000; Strayhorn, 2018), one of the most widely cited is Tinto's (1975) model of student integration. This theory has its roots in Durkheim's theory of suicide (Durkheim, 1961), and postulates that students are more likely to dropout when they are insufficiently integrated into the fabric of a university system (Tinto, 1975). However, universities are not solely comprised of an academic system, but also serve an social purpose. Distinguishing between these two structures is important, not only because there is a direct relationship between a student's participation in the academic domain and their academic outcomes (Spady, 1970; Strayhorn, 2018), but also because this distinction suggests that a student may be capable of achieving integration in one area without doing so in the other (Tinto, 1975, 1993, 2012). For instance, students may withdraw due to poor academic performance, or else may succeed academically but withdraw due to their inability to integrate into the social life of the institution. Furthermore, one may expect a reciprocal relationship to exist between these two domains such that excessive integration in one would come at a cost to integration in the other; for example, excessive social activities at the expense of academic studies (Tinto, 1975, 1993, 2012).

Tinto's (1975) model of student integration, displayed in Figure 1, posits that the student integration may be viewed as a longitudinal process of interactions between the individual and the academic and social domains of their university. This process, however, leads different students to varying degrees of persistence. Accordingly, Tinto (1975) argues that we must not only account for background characteristics of students (such as sex, ability, and demographics), but also motivational attributes (such as career expectations and academic achievement). This contextual infor-

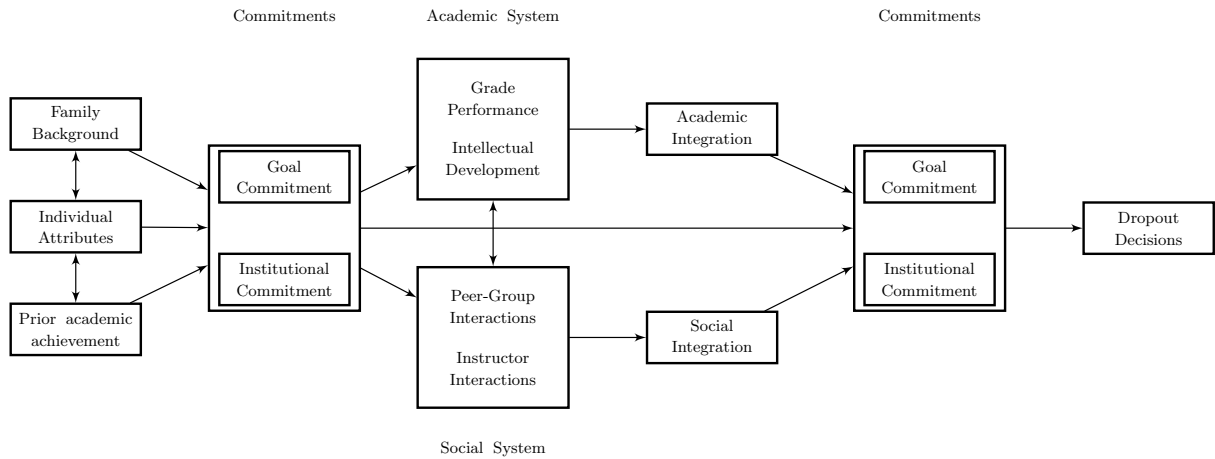


Figure 1. Student integration model (adapted from Tinto, 1975)

mation influences students' initial commitments to the institution and their studies, which are then modified over time as a result of interactions with the academic and social structures within which students find themselves (Tinto, 1975). This model has been subjected to extensive empirical validation (Braxton, Sullivan, & Johnson, 1997; Thomas, 2000) using a variety of methods and data. For instance, while some studies have focused on survey responses (Nora & Rendon, 1990), others have used the time that students spent in class together, various categories of social network variables, and a number of demographic variables (Eckles & Stradley, 2012). These results have important implications for higher education institutions seeking to leverage their student data to understand student engagement, learning outcomes, and student retention. In particular, the model provides a theoretical framework for analysing students' academic and social interactions separately, whilst acknowledging the interplay of these domains as part of a broader network of associated forces that influence students' academic achievement and persistence in their studies.

In this thesis, we take inspiration from Tinto's (1975) model and present novel methods for analysing and understanding students' academic and social domains, using data collected by higher education institutions. In doing so, our research is pervaded by three key concerns: firstly, ensuring that any measures developed are both structurally valid and generalise across contexts; secondly, providing interpretable insight with regards to student engagement; and finally, providing representations of students' behaviour that are predictive of their academic outcomes and their persistence in their studies. Taken together, these analyses provide theoretically-grounded representations of students' interactions, and lay the groundwork for interpretable insight to redress student disengagement and dropout. Importantly, while this work is inspired by Tinto's (1975) model, particularly with regards to dropout, it is not an empirical validation and at times relies on alternative theoretical approaches to frame and inform our research.

1.1 Research goals and questions

Although the work presented in this thesis separately analysed the academic and social systems, identified within Tinto's (1975) model, both avenues of investigation were guided by four research questions. In particular, our work has emphasised validity. This is in part due to criticisms calling into question the methodological validity of some learning analytics research (Caulfield, 2013); a result supported by recent replication studies (Dawson, Jovanovic, Gašević, & Pardo, 2017). Furthermore, the accuracy of predictive models has often been found to decline when applied in different contexts (Jayaprakash, Moody, Lauría, Regan, & Baron, 2014), indicating the need for a renewed focus on generality within the field. Thus, our first research question is:

RESEARCH QUESTION 1:

In the course of analysing and assessing students' interactions within the academic and social systems of their educational institution, how can we ensure that the models we develop not only provide faithful representations of the construct they are intended to measure, but also measure the same construct across populations and contexts?

However, models and data alone are not sufficient for informative research into students' behaviour and decision-making (Gašević et al., 2015). Rather, the role of theory is increasingly recognised as essential for informing the choice of questions asked and the hypotheses tested (Rogers, Dawson, & Gašević, 2016; Wise & Shaffer, 2015). In advocating for theoretically-grounded analytics, some researchers have gone so far as to state that purely data-driven approaches are a misconception of the scientific method and commit "the logical (and ethical) error of using descriptions of the past as prescriptions for the future" (Reimann, 2016, p. 136). However, these two approaches are not mutually exclusive; data-driven analyses may be used to validate theoretical claims and, in doing so, can potentially inform changes in theory. This emphasis on theory informs our second research question:

RESEARCH QUESTION 2:

To what extent can we ground our computational analyses within existing educational theory? Furthermore, can the use of theoretical frameworks not only inform the choice of hypotheses we investigate but also identify which associations identified are meaningful with respect to students' academic outcomes?

The insights derived from any analytics, however, are only useful if they can be communicated to the relevant stakeholders. This implies that the output of any modeling must offer intuitive, interpretable insight for both instructors and students. To date, however, research has found that learners experiencing difficulty with learning analytics tools, such as dashboards, is a common occurrence (Corrin & de Barba, 2014; Matcha, Ahmad Uzir, Gašević, & Pardo, 2019). The challenge, then, lies in bridging the knowledge gap between how data is collected, processed, and modeled, and

the subsequent interpretation and decision to deploy pedagogical actions and interventions (Pardo et al., 2016). Reconciling the expert knowledge required to interpret these models with the pedagogical decision to intervene motivates our third research question:

RESEARCH QUESTION 3:

In analysing students' trace data, can we develop models which provide not only a prediction but are also interpretable? Furthermore, can this interpretation be provided without requiring the recipient to understand how the data that produced the prediction was collected, processed, and modeled?

Finally, while each of the preceding research questions touch on key issues facing research and practice within learning analytics, they are, in isolation, insufficient. As our understanding of the complexity of learning processes matures (Jacobson, Kapur, & Reimann, 2016), it is essential that we also develop predictive models that can cope with this complexity. General regression models, for instance, may be inadequate to address the structure of such complex systems (Reimann, 2016). It is thus essential that learning analytics co-opts data science to foster the development of novel methodologies that are consistent with the requirements of the preceding research questions. Developing novel methodologies which are valid, grounded in theory, and interpretable motivates our final research question:

RESEARCH QUESTION 4:

Can novel methodologies from the data science and machine learning literature be applied to learning analytics to provide valid, theoretically-grounded analyses? In addition to greater predictive power, can these more complex models also provide more simple explanations to students and instructors?

1.2 Methodology

The work presented in this thesis is conceptually indebted to Tinto's (1975) model of student integration. Accordingly, our analysis follows the conceptual distinction between the social and academic domains with which students interact. In investigating the former, social, domain we relied upon data drawn from MOOCs as well as students' university enrolment data. In our first study, we used statistical methods from social network analysis (SNA) to assess the validity of a number of different social tie definitions (RQ1). Concretely, we investigated whether the choice of tie definition influenced the structural and statistical properties of the derived network, as well as any associations between centrality metrics and student grades. Next, we conducted an empirical investigation of Tinto's (1975) central claims (RQ2). In doing so, we created a co-enrolment network from over three decades of students' university course enrolments and used graph-embedding techniques to generate a latent representation of each student (RQ4). We found that these representations were

not only predictive of students' academic performance and dropout decisions, but were also robust to changes in the embeddings' parameterisation and corruption of the underlying network (RQ1).

In investigating the latter, academic, domain we relied upon data drawn from MOOCs and blended learning environments in higher education. In our first study, we sought to condense the highly dimensional data that student interactions generate into interpretable representations (RQ3). In identifying these representations, our methodology took heed of the theoretical literature (RQ2), and found that students' sequences of these representations were associated with their academic outcomes. Next, we empirically validated a theoretical model of student engagement (RQ2), and evaluated its generality across three disparate MOOCs (RQ1). Finally, we developed a novel model of student interactions (RQ4) that not only produces interpretable insights through the identification of students at risk of disengaging (RQ3), but also generalises across courses and contexts (RQ1).

While MOOC data, particularly trace logs, can tell us much about students' *actual* behaviour (P Winne & Jamieson-Noel, 2002; Hadwin, Nesbit, Jamieson-Noel, Code, & Winne, 2007), they are limited in their ability to tell us *why* a student behaved in a particular way. Additional data, such as questionnaires and surveys, is required to provide information about students' dispositions (Tempelaar, Rienties, & Nguyen, 2017). While the former data can help instructors identify *where* to intervene, the latter data can help them identify *how* to intervene. Due to the data that was available, the work presented in this thesis focuses on the former question but provides a foundation that can be extended to include the latter.

Tinto's (1975) model of student integration originated within the context of university education. MOOCs, however, comprise a different educational context, and it is by no means evident that the mechanisms of student integration that Tinto (1975) posits transfer to this novel context. Accordingly, one may question whether Tinto (1975) is a suitable framework for understanding dropout within MOOCs. Unfortunately, while there is nascent theoretical work into MOOC engagement, there is limited research that discusses MOOC retention from a theoretical perspective. Given that this thesis has focused on the development of computational models, we have relied on what theoretical work exists; namely, Tinto (1975). Nevertheless, we acknowledge that there is considerable potential for further theoretical work regarding student dropout in MOOCs.

1.3 Thesis structure and overview

To address our four research questions, we focused our efforts on two of the key theoretical predictions of Tinto's (1975) model, namely, that students' interactions with the academic and social systems present in their educational institution are associated with their academic outcomes and their persistence in their studies. This resulted in five chapters, where each focuses on one or more research questions (Table 1), and includes one peer-reviewed publication, constituting the core of the chapter. In addition, we provide a short preface and summary to each publication to explain

Table 1. Overview of how the research questions are addressed within each chapter.

Chapter	Research questions			
	Validity	Theoretical Grounding	Interpretability	Novel Methodology
Chapter 2	✓			
Chapter 3	✓	✓		✓
Chapter 4		✓	✓	
Chapter 5	✓	✓		
Chapter 6	✓		✓	✓

how they fit into the overall structure of the thesis. In the remainder of this section, we provide a brief overview of each chapter and indicate how they address our research questions, and contribute to the development of the thesis.

1.3.1 Overview of Chapter 2: “The Validity of Social Ties” (RQ 1)

Before analysing how students’ academic performance and dropout decisions are associated with their local social network, it was first necessary to investigate how these networks are formed. SNA has been one of the most commonly applied techniques within learning analytics, and while the literature employs a wide variety of social tie definitions, the impact this choice may have on the results of these analyses has been largely overlooked. Accordingly, the spectre of validity looms large, necessitating a study of how this seemingly minor methodological decision can have far-reaching implications for the results of any analysis.

Research contributions:

- Using the discussion forums from one MOOC and one blended learning course, we derived social networks using a range of social tie definitions found in the literature.
- We then investigated the extent to which such social tie extraction methods influenced the structural and statistical properties of the networks.
- Not only does the choice of social tie definition influence the structural and statistical properties of the derived network, but also the association between centrality measures and academic performance.

Research output:

1. Fincham, E., Gašević, D., and Pardo, A. (2018): “From Social Ties to Network Processes: Do Tie Definitions Matter?”: A journal article describing the variation in the literature as to what constitutes a social tie, and how these variations play an important role in shaping the results of any analysis, published in the *Journal of Learning Analytics*.

1.3.2 Overview of Chapter 3: “Performance and Persistence in Co-enrolment Networks” (RQs 1 & 2 & 4)

In this chapter, we focused on the social domain identified in Tinto’s (1975) student integration model, and empirically investigated two key theoretical predictions: namely, whether students’ social networks are associated with their academic performance and their persistence in their studies. In doing so, we conducted this analysis at a scale unprecedented in the literature, necessitating the use of novel graph-embedding techniques. These models provide a latent representation of each student, capturing the features of their local neighbourhood, and the structural role that they play within it. We found that not only do these representations predict students’ final GPA, but are also able successfully to classify students who dropout. To ascertain the validity of these results, we evaluated the predictive performance of the latent representation over a range of hyper-parameter settings, and conducted a corruption procedure that randomly pruned edges. We found that the predictive performance of the latent representation is robust to both these changes, indicating that hyper-parameters may be selected to reduce the computational demands of this method without loss of predictive power.

Research contributions:

- We empirically investigated a number of Tinto’s (1975) theoretical predictions, within the context of a university’s social domain.
- By using graph-embedding techniques, we created latent representations of students within a large co-enrolment network. These latent representations were found to be predictive of students’ final GPA and their dropout decisions.
- This predictive power was robust to both changes in the model’s parameterisation and corruption of the underlying networks (prior to their embedding).
- With the inclusion of select covariates, the regression model achieved an R^2 of 0.24, while the classifier achieved F_1 -scores of up to 0.83, significantly outperforming the existing literature.

Research output:

1. Fincham, E., Rózemberczki, B., Kovanović, V., Joksimović, S., and Gašević, D. (2019): “*Persistence and Performance in Co-enrolment Network Embeddings*”: A journal article presenting our novel methodology, currently under review.

1.3.3 Overview of Chapter 4: “Interpretable Representations of Student Behaviour” (RQs 2 & 3)

To investigate the extent to which students interact with the academic domain described in Tinto’s (1975) model, we began by examining the trace data generated by students within a blended learning environment. In doing so, we grounded our computational analysis within educational theory, and sought to identify a set of “learning strategies” (Weinstein, Husman, & Dierking, 2012): that is, any behaviours that facilitate the acquisition, understanding, or later transfer of new knowledge.

Following the theoretical literature, we postulated that these strategies were themselves composed of shorter-term “tactics”, indicating cognitive routines for performing specified tasks (Alexander, Graham, & Harris, 1998; Kirby, 1988). Students’ study tactics were modeled using a hidden Markov model (HMM), the parameters of which offered interpretable descriptions of students’ behaviour. Sequences of study tactics were then generated for each student and were clustered to identify overarching learning strategies. Differences in students’ choice of learning strategy were found to be significantly associated with their final GPA. By analysing three consecutive runs of the course, over which a feedback intervention was incrementally implemented, we provided a qualitative discussion of how this feedback could be associated with students’ choice of strategy.

Research contributions:

- We developed a method for identifying different study tactics and learning strategies with respect to students’ use of resources available within a blended learning environment.
- Using the data from several offers of this course, we identified different learning strategies and examined their association with learning outcomes, as measured by final grade.
- Taking inspiration from the theoretical literature, our method provides interpretable descriptions of students’ behaviour within the learning environment.

Research output:

1. Fincham, E., Gašević, D., Jovanović, J., and Pardo, A. (2019): “From Study Tactics to Learning Strategies: An Analytical Method for Extracting Interpretable Representations”: A journal article describing our methodology for identifying learning strategies, and their association with course outcomes, published in the *IEEE Transactions on Learning Technologies* journal.

1.3.4 Overview of Chapter 5: “Validating a Theoretical Model of Student Engagement” (RQs 1 & 2)

After our initial foray into the academic domain, described in Chapter 4, we sought to address some of the limitations of this methodology. In particular, it was clear that course design would have an outsized influence upon the types of study tactics and learning strategies identified. This not only calls into question the generality of this approach, but also the validity of our learning strategy construct. To resolve this, we turned to the theoretical literature, in particular the research into student engagement, and sought to operationalise the model of engagement proposed by Reschly and Christenson (2012) and updated for online learning environments by Joksimović et al. (2018). This model posits that student engagement may be decomposed into four facets, each of which may be captured by a series of trace-based metrics common to the learning analytics literature. We calculated these metrics using three MOOCs from diverse disciplines and which represent a range of pedagogical approaches. Using exploratory and confirmatory factor analysis on random subsamples of this data, we identified a latent variable structure and compared it to the predictions of the theoretical model. We found that while some of these latent variables were associated with

each other, only the “behavioural” facet was associated with students’ final course grade. Finally, we evaluated how our model generalised across the different courses in our dataset.

Research contributions:

- We conducted an empirical validation of the theoretical model of student engagement provided by Joksimović et al. (2018).
- We used EFA and CFA to identify a latent variable model structure and compared this result with the predictions of Joksimović et al.’s (2018) model. We then augmented this latent variable model with a path analysis to form a structural equation model (SEM). This enabled us to not only evaluate the associations between our latent variables, but also between our latent variables and student outcomes.
- We then assessed the measurement invariance of our model to assess the extent to which our SEM generalised across different subject domains.

Research output:

1. Fincham, E., Whitelock-Wainwright, A., Kovanović, V., Joksimović, S., van Staaldouin, J.-P., and Gašević, D. (2019): “*Counting Clicks is Not Enough: Validating a Theorized Model of Engagement in Learning Analytics*”: A full conference paper presenting our theoretically-grounded methodology for measuring student engagement in online learning environments. The paper was presented at the Ninth International Conference on Learning Analytics and Knowledge (LAK’19).

1.3.5 Overview of Chapter 6: “Towards Interpretable Insight” (RQs 1 & 3 & 4)

In our final investigation of students’ interactions with the academic domain (Tinto, 1975), we sought to address both the validity and generality constraints of Chapter 4 and the lack of interpretable insights of Chapter 5. In doing so, we extricated ourselves from theoretical debates surrounding the nature of student engagement and instead focused on disengagement: that is, the process through which students’ interactions with academic materials are curtailed, ultimately resulting in dropout. In contrast to our previous work in this area, we posited that student behaviour is temporally situated, and that understanding the dynamics of this activity can be used to identify at-risk individuals. We operationalised this by aggregating student behaviour at a weekly granularity. These weekly student representations were then modeled with an HMM, as this not only captured a set of typical interaction patterns – states – but also modeled the transitions between them. To validate that this representation was capable of identifying student disengagement, students’ state representations were used on an incremental, weekly basis to predict whether or not students would disengage from the course in the following week. To demonstrate that this approach generalised across contexts, we repeated this procedure, but tested models trained on separate courses. Finally, we qualitatively compared the representation of student behaviour provided by our model with that of comparable approaches in the learning analytics literature. We argued that our representa-

tion provides a more explicit accounting of time whilst requiring less expert knowledge to identify students at-risk.

Research contributions:

- We developed a model of student interactions with academic resources that accounts for temporal changes in activity. In doing so, we were able to predict when students were about to disengage from the course resources.
- Our representation of student behaviour was sufficiently general that the model could be trained on one dataset and evaluated on the task of predicting dropout on another without significant performance loss.
- We also compared the representation of student activity generated by our model to the current state-of-the-art, and argued that our model provides more interpretable insight.

Research output:

1. Fincham, E., Hoernle, N., Gal, K., and Gašević, D. (2019): *“The Road Not Taken: Preempting Dropout in MOOCs”*: A journal article presenting our novel methodology, currently under review.

1.3.6 Overview of Chapter 7: “Conclusions”

Finally, in Chapter 7 we examine how the work presented in this thesis has impacted upon the four research questions identified in Chapter 1. In addition, we discuss the practical applications of our research, as well as the interesting avenues of potential future research that our work has generated. Finally, we conclude with a short overview of the thesis and a summary of its key contributions.

1.4 Ethics approval

The data used in this thesis originates from two sources: students’ interactions with resources and their peers within higher education blended learning environments and MOOCs; and students’ university enrolment data. Both of these raise ethical concerns relating to participant consent, and the potential risk of de-anonymisation. Prior to any data access, ethical approval was granted via the appropriate channels at the relevant institutions.

In Chapter 2, we relied on flipped classroom data from the University of Sydney and MOOC data from the University of Edinburgh. Access to the former was approved by the relevant panel at the University of Sydney, while access to the latter followed the ethics procedure at the School of Informatics (as of 2017). This approval was obtained by my supervisor, Professor Dragan Gašević. In Chapter 3, our analysis involved student enrolment data from the University of South Australia, for which approval was granted via the relevant ethical panel. In Chapter 4, we relied upon the same flipped classroom data as Chapter 2, for which approval was already granted. Chapter 5 relied upon MOOC data sourced from TU Delft, for which approval was obtained from the relevant panel. Finally, in Chapter 6 we used the same University of Edinburgh MOOC data as Chapter 2

1. INTRODUCTION

as well as MOOC data from Columbia University, for which approval was granted via the relevant panel.

2

The Validity of Social Ties

2.1 Introduction

AT the core of Tinto's (1975) model is the prediction that students' academic outcomes and dropout decisions may be partially understood in terms of the social systems that students inhabit. Within learning analytics, questions about the types of relationships and interactions that occur between individuals, groups, and communities have long been the purview of SNA. The validity of these methods, however, is a vital question that has often been overlooked in the literature. Therefore, before conducting any investigation into the claims of Tinto's (1975) model, it is first necessary to ensure that the constructs used in such an investigation are valid. In the present study, social ties refer not to real social relationships, but rather the links created between individuals as they interact in learning environments, in particular, discussion forums.

Following Messick (1995), validity is a multi-faceted construct that may be divided into three core types: namely structural validity, or the extent to which a metric or measurement actually describes the construct it is intended to capture; generalisability, or the extent to which a measurement captures the same construct across populations; and external validity, such as supportive or dissuasive evidence arising from related constructs.

As the basis of any network analysis, tie definition is a crucial methodological decision, as each definition carries with it a set of assumptions about the nature of social interactions. Compare, for instance, a "direct-reply" structure whereby social ties are created on the basis of dyadic interactions to a "co-presence" structure where all participants within a domain are connected. While the former may be suited for, say, analysing email chains within a corporation, the latter may more suited to analysing communities of preference. However, the theoretical subtext of this decision is often overlooked in the learning analytics literature, and studies often establish a tie definition without providing an explanation or a rationale. Furthermore, researchers' choice of tie definition are not always examined when the results of two competing studies are compared.

To demonstrate that this raises grave concerns with regards to validity, we conduct a computational analysis. In doing so, we address each aspect of validity: structural validity is assessed by comparing tie definitions on the basis of the structural and statistical properties of the derived

networks; external validity is evaluated by investigating how measures of network centrality are associated with academic performance; and generalisability is assessed by pursuing the foregoing analysis in two distinct MOOC settings.

2.2 Publication: From Social Ties to Network Processes: Do Tie Definitions Matter?

The following section includes an extended copy of the following publication:

Fincham, E., Gašević, D., & Pardo, A. (2018). From social ties to network processes: Do tie definitions matter? *Journal of Learning Analytics*, 5(2), 9–28. doi:10.18608/jla.2018.52.2

From Social Ties to Network Processes: Do Tie Definitions Matter?

Ed Fincham

School of Informatics
The University of Edinburgh
ed.fincham@ed.ac.uk

Dragan Gašević

Faculty of Education
Monash University
dragan.gasevic@monash.edu

Abelardo Pardo

Division of Information Technology, Engineering and the Environment
University of South Australia
abelardo.pardo@unisa.edu.au

ABSTRACT: The widespread adoption of digital e-learning environments and other learning technology has provided researchers with ready access to large quantities of data. Much of this data comes from discussion forums and has been studied with analytical methods drawn from social network analysis. However, within this large body of research there exists considerable variation in the definition of what constitutes a social tie, and the consequences of this choice are rarely described or examined. This paper presents findings from two distinct learning environments regarding different social tie extraction methods and their influence on the structural and statistical properties of the induced networks, and the association between measures of centrality and academic performance. Our findings indicate that social tie definitions play an important role in shaping the results of our analyses. The primary purpose of this paper is to raise awareness of the consequences that such methodological choices may have, and to promote transparency in future research.

NOTES FOR PRACTICE

- Social network analysis has been one of the most commonly applied methods within learning analytics. However, many of the common constructs and tools these methodologies employ have not been subjected to robust validation. Such concerns pertain to construct validity: namely, does a metric actually measure what it purports to measure?
- In this study, we find that different social tie extraction methods influence the structural and statistical properties of the induced networks, as well as the associations between centrality measures and academic performance.
- Our results emphasise not only the importance of transparency in the choice of tie definition, but also the importance of providing a justification for that choice. Given the impact that tie definitions may have, we advise that practitioners investigate a number of options to ascertain the extent to which such methodological choices can bias their results.

Keywords: Social network analysis, Discussion forum, Tie extraction, Construct validity, ERGM, MOOC, Academic achievement

1 INTRODUCTION

Research into learning analytics has garnered much attention for its potential impact upon a number of central issues in education. For instance, the identification of learning strategies (Jovanović et al., 2017), the prediction of academic success (Gašević et al., 2016), and the provision of personalized feedback at scale (Pardo et al., *in press*), to name a few. While this field of research promises much, the spectre of validity looms large, and many of the most frequently relied upon measures have not been subjected to robust validation. A pertinent example of this is time-on-task, the estimation of which was rarely discussed in the literature. Instead, researchers would often opt for a heuristic approach such as limiting session activity to a defined time period (Ba-Omar et al., 2007; Munk & Drlík, 2011). However, there was little consideration of the consequences such estimation heuristics had on the results of the final predictive model. To address this oversight, Kovanović et al. (2015) investigated how different time-on-task estimation methods affected predictive models of learner outcomes. Across diverse learning contexts, the authors found that estimation methods play an important role in shaping the final study results.

Concerns regarding validity ultimately relate to the extent to which a metric actually measures what it purports to measure. This is known as construct validity (Messick, 1995) and is highly relevant to learning analytics methods, particularly, in the context of this study on social network analysis (SNA).

SNA has been one of the most commonly applied methods within learning analytics (Joksimović et al., 2016; Dawson et al., 2014). While SNA can offer insight into the types of relationships and interactions that occur between individuals, groups, and communities, little research has considered the validity of findings derived from common SNA methods. For instance, although Batool & Niazi (2014) assessed the construct validity of centrality metrics in complex networks, studies such as this are the exception rather than the rule. Existing research has paid little attention to the validity of a number of common SNA constructs and, in particular, the impact of various tie definitions on these constructs remains largely overlooked.

Following Messick (1995), validity may be deconstructed into a number of different aspects, including structural validity, or the fidelity of the scoring structure to the structure of the construct itself (Loevinger, 1957); generalisability, or extent to which score properties generalize to and across populations and settings (Cook & Campbell, 1979; Shulman, 1970); and external validity, such as supportive or dissuasive evidence arising from related constructs.

The present study investigates the construct validity of a number of social tie definitions in the context of online discussion forums. These definitions seek to represent the relationships formed between individuals on the basis of interactions and mutual participation within threads. That is, we assess how variations in tie definition result in different characterizations of these relationships. In doing so, we investigate three aspects of construct validity. Structural validity is assessed by comparing tie definitions on the basis of the structural features of the derived networks, and by using statistical models to compare the statistical properties of these networks. External validity is evaluated by investigating how

measures of network centrality are associated with academic performance, and generalisability is assessed by pursuing the foregoing analysis in two distinct learning settings.

2 LITERATURE REVIEW

2.1 SNA and Discussion Forums

The analysis of discussion forums, particularly Massive Open Online Course (MOOC) discussion forums, has received considerable attention in recent years. In this body of research, SNA has proven to be a powerful tool in extracting patterns of connections between learners, exploring their relationship with learning, and generating understanding about the underlying relational structure of a community across a variety of contexts. In particular, the emergence of MOOCs has provided ample opportunity for the application of SNA methods (Gašević et al., 2014; De Laat & Prinsen, 2014). Given the increasing number of students enrolling in MOOCs (Jordan, 2015), SNA has become an increasingly adopted tool for visualising and extracting interaction patterns from social learning activities (Dowell et al., 2015; Jiang et al., 2014), as well as for investigating the association of network centrality with constructs such as academic performance (Joksimović et al., 2016; Schreurs et al., 2013; Skrypnik et al., 2015), sense of community (Dawson, 2008), social presence (Kovanović et al., 2014), and creative potential (Dawson et al., 2011).

There is considerable heterogeneity in how learners interact with the discussion forum. Gillani and colleagues (2014, 2014), for example, analysed forum users on the basis of co-participation in the same threads, and found that the coherence of the network mainly depends on a small set of central users. Rather than a close-knit community, forums users may be more accurately characterized as a loosely connected crowd. Poquet and Dawson (2016) explicitly analysed different user groups, and found that regular users form a denser, more centralized network as they have more opportunities to establish connections. Further work by Boroujeni et al. (2017) confirmed that membership of these groups remains stable over time. However, there is more to discussion forums than structure alone; Wise et al. (2018) distinguished between discussions that were related to course material and those that were not. They found that students who made both content and non-content related posts had a higher passing rate than those who only contributed to one type. Furthermore, those who contributed to content-related threads performed slightly better than those who did not.

The results of these analyses, however, have not always been consistent. For instance, Joksimović et al. (2016) investigated the factors that influence social connections in two instances of a MOOC, offered in English and Spanish, that taught students how to programme. Ties were extracted on the basis of direct reply from an online discussion forum. In examining the association between centrality degree, closeness, betweenness, and academic performance, the authors found weighted degree was significantly associated only in the English offering, while the effects of betweenness and closeness were only found in the Spanish offering. Furthermore, the authors found evidence of performance-based homophily, indicating that learners tend to talk to those in the same performance group as themselves. Another study, by Jiang et al. (2014), also investigated the associations between social centrality and academic performance. Their study was conducted on two MOOCs in algebra and finance, and ties were

extracted via co-presence in a thread, that is, on the basis of shared activity and participation in the discussion. While degree and betweenness were positively correlated with academic performance in the algebra course, no significant correlation was found between any centrality measure and academic performance in the finance course. In further contradiction of Joksimović's findings, the authors found that students tend to talk to those in difference performance groups than themselves.

The findings of these two studies are largely inconsistent; a discrepancy which may in part be attributed to methodological differences. For instance, Joksimović et al. (2016), hypothesise that the association between academic assessment and network centrality measures was moderated by the presence of Simmelian ties, that is, three-cliques with reciprocal ties (Krackhardt, 1999). In lieu of such a hypothesis, Jiang and colleagues' (2014) methodology did not consider the presence of a Simmelian influence. In this case and others, researchers have used different methods to extract social ties yet the effects on those extractions are rarely studied. While in the case of these two studies, the effects of tie extraction methods are studied in connection with the association between network centrality and academic performance, the same methodological oversight may be found regarding other constructs and hypotheses. Moreover, not all research into networked learning has relied upon MOOCs (Cho et al., 2007; Dado et al., 2017) and, in investigating how social tie extraction methods impact upon the structures and statistical properties of networks, there is scope for a comparison between networks extracted from MOOCs and other, more formal, educational contexts.

2.2 Network Processes and Exponential Random Graph Models

Studies that apply SNA methods rely upon mathematical models to describe relationships between variables, to reveal important characteristics, and to identify processes within the social network (Carrington et al., 2005; Goodreau et al., 2009). For instance, descriptive models enable us to identify whether or not reciprocity exists within a given network. However, to understand whether or not such processes occur more often than expected if ties were generated randomly, we must rely on statistical models (Goodreau et al., 2009). One commonly proposed method are Exponential Random Graph Models (ERGMs) (Joksimović et al., 2016; Morris et al., 2008; DuBois et al., 2013).

Introduced by Frank and Strauss (1986) and Wasserman and Pattison (1996), ERGMs belong to a family of probability models that allow for generalisable inferences over the structural foundations of social behavioural patterns within networks (Morris et al., 2008; Robins et al., 2007). ERGMs treat network ties as random variables, and model the overall network structure through a set of local network processes, such as triadic closure, mutuality, or transitivity (Robins et al., 2007). The model assumes each tie within these processes is conditionally dependent, indicating that "empirical network ties do not form at random, but they self-organize into various patterns arising from underlying social processes" (Wang et al. 2013, p. 3).

Though ERGMs have long been successfully applied in other fields, their application to the structural analysis of forum networks is relatively novel (e.g. Poquet & Dawson, 2016; Kellogg et al., 2014; Joksimović et al., 2016; Zhu et al., 2016). In general, these results have revealed a reciprocal tie effect within networks, and a lack of network centralisation beyond a few influential nodes. For instance,

Kellogg et al. (2014) used ERGMs to provide a more comprehensive understanding of the dynamic processes underpinning peer support learning in MOOCs. The authors used both descriptive and statistical methods and found a strong and significant reciprocity effect, indicating that students are more likely to aid their peers when there is prior evidence of reciprocity. In a more recent study, Joksimović et al. (2016) utilized ERGMs to determine whether network social dynamics, such as Simmelian ties, have an impact on the predictive power of network positions. The study found that incorporating both descriptive and statistical models allowed for more nuanced and contextually salient inferences about learning within a network. Poquet et al. (2017) found that different facilitation, or pedagogical, approaches mediated the extent of reciprocity. That is, while direct reciprocal ties were characteristic of non-facilitated forums, triadic reciprocal ties were more prominent in forums with a high degree of facilitation (that is, instructor involvement).

While statistical models such as ERGMs have facilitated valuable research and provided considerable insight into network processes, the learning analytics literature has neglected research into whether and, if so, to what extent, network processes and statistical properties are influenced by variations in the tie definitions that underpin them.

2.3 Social Tie Definitions

Research into SNA and, in particular, SNA studies of MOOCs have relied upon a variety of definitions to construct social ties. While some authors (e.g. Joksimović et al., 2016; Kellogg et al., 2014) defined ties on the basis of direct replies, others (e.g. Gillani & Eynon, 2014; and Jiang et al., 2014) have relied on co-presence. As the basis of any SNA analysis, tie definition is crucial and each definition carries with it a set of assumptions about the nature of social interactions. In the literature, this theoretical oversight remains largely unaddressed and studies often establish a tie definition with no explanation nor rationale. Even when one is provided, each decision often carries its own shortcomings. For instance, Gruzd & Haythornthwaite (2008) consider three potential tie definitions and note that each makes specific assumptions about the nature of social interactions that may not hold.

In a study of MOOC forums, Wise et al. (2017) investigated the impact of different tie definitions on social network structures and the resultant characteristics at the network, community, and individual node level. While their study found that network properties were characterized by a limited sensitivity to differences in tie definitions, their analysis was limited to descriptive statistics and did not consider the statistical properties of networks, such as the propensity for reciprocity or homophilic ties. Accordingly, there is scope for an investigation into how different tie definitions relate to differences in the statistical properties of networks and the interpretation of such social networks.

Social tie definitions can typically be classified into two distinct types: those which interpret a tie as created when an individual speaks to another, and those which extract ties on the basis of co-participation within a discussion. Perhaps the most prevalent and straightforward of the former is Direct Reply. Under this rubric, a tie is constructed when there is a reply relationship between two nodes in the same thread. For instance, between the starter of a thread and the author of a reply post addressed to it, or between the author of a reply post and the author of a reply to that reply. While this definition has

been widely used (see Joksimović et al., 2016; Kellogg et al., 2014) there is no guarantee that users will opt for the correct location and level of post, nor that the platform itself will support a sufficiently rich reply structure. For instance, in one popular platform for online discussions, Piazza, only three levels of post are supported: post, reply, and reply to a reply. If a poster wishes to reply to a “reply to a reply” post, it is logged as another “reply to a reply” post. In building a network, a Direct Reply tie definition would link this new post to the “reply” post rather than the “reply to a reply” post to which it was originally intended to address (see Figure 1).. Accordingly, the extent to which the reply structure thus derived reflects the actual relations among learners is open to question.

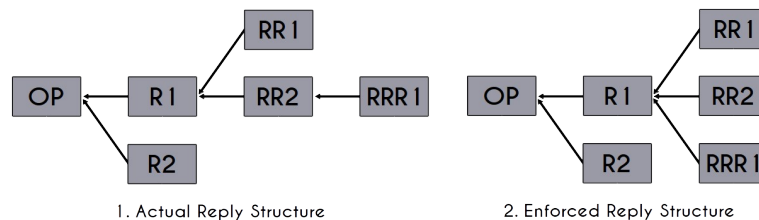


Figure 1: Limitations of Platform Supported Reply Structures

To address such concerns, Zhu and colleagues (Zhu et al., 2016) proposed the Star Reply tie definition. While Direct Reply considers multiple levels of replies and defines ties as connections across levels, Star Reply does away with reply structures and considers all posts within a thread as tied to the thread starter. The justification for this is that even if a reply post does not directly address the thread starter, it was made in the context of the discussion originated by the thread starter. While Star Reply emphasizes the thread starter, it fails to distinguish between different levels of replies and does not consider connections formed between posters within the same thread. To address this, Direct Star (Gruzd & Haythornthwaite, 2008) amalgamates the two and defines ties across different levels within a reply structure on the basis of Direct Reply, while also linking posts within a discussion back to the thread starter. However, the methods thus far identified strictly emphasize the act of speaking to another individual. Within a thread, a prospective poster may read much of the existing discussion before penning their own reply. Accordingly, defining ties on the basis of speaking contact alone overlooks the interactions between individuals who do not speak directly, but share an interest and an awareness of each other within the same thread.

Tie definitions of the second type – those based on co-participation within a discussion – seek to address this issue by creating a network of co-presence across nodes. Within such a network, a tie is defined as being present in the same part of the discussion; there is no necessity for direct interaction. Ties are thus created without regard for the reply structure present in a discussion: connections are formed both between a parent node (post) and its children (replies to it), and between the children themselves. Accordingly, this type of tie definition represents online discussions not as strict hierarchies but as collective conversations.

Within this type of tie definition, a common approach is that of total co-presence, where any two nodes in the same thread are connected, regardless of post type (Gruzd & Haythornthwaite, 2008). While this is often used to map interaction, in the case of large threads it can prove problematic. In the case of small threads, linking all individuals as part of a collective discussion might be reasonable, however, this assumption becomes implausible when the number of replies is very large. One way to address this problem, proposed by Wise and colleagues (Wise et al., 2017) is to set a cap on the reasonable number of posts in the same thread to create a measure of limited co-presence. Beyond this threshold, all posts within subthreads are connected to all other posts within that subthread, and the thread starter.

An alternative method of assessing co-participation which has been largely unexplored is that of viewing ties as contained within a moving window. Within a large thread, an prospective poster may only attend to recent posts in framing their reply and so the collective conversation of which they are part is defined as some subset of the overall thread. A moving window, defined as some number of posts, moves sequentially over a thread and, at every step, all posts within the window are connected. This approach ignores the hierarchical structure of a discussion and instead emphasizes each post as being part of a temporally defined, collective discussion. However, there is no *a priori* rationale for choosing one window size over another, and different sizes may lead to a variety of different conclusions.

While each tie definition carries with it a set of assumptions about the nature of social interaction, in the literature little heed to paid to this, and extraction methods may instead be chosen on the basis of expediency, such as whichever structure is readily permitted by the discussion platform. Furthermore, there has been little research into the impact that variations in tie definitions have, both on the statistical properties of networks and on the association between network centrality metrics and academic performance.

2.4 Study Framing

In this study, we examine the effects that different social tie definitions have on the structural and statistical properties of the derived networks. These range from network-level properties such as reciprocity to individual properties, such as the association between metrics of centrality and academic performance. To validate our results, our analysis is applied in two separate contexts: a blended learning environment and a MOOC. The importance of doing so is two-fold: on the one hand, it provides a glimpse into how differences in learning contexts may impact upon social interaction. On the other, the two contexts allow us to assess the construct validity of tie definitions by measuring their impact on the structural and statistical properties of the resultant networks. The paucity of existing research into whether network construction choices determine network properties motivates our two research questions:

RQ1: Do differences in tie formation mechanisms determine the statistical properties of networks across different learning contexts?

RQ2: Do differences in tie formation mechanisms affect the association between social centrality and academic performance across different learning contexts?

3 METHOD

3.1 Data Sources

This study analysed forum discussions from two separate courses. The first dataset comes from a flipped classroom, first-year engineering course at an Australian higher education institution offered in 2016. The course, called Introduction to Computer Systems (ICS), lasted 13 weeks and, of a total enrollment of 477 students, 227 students participated in the discussion forum. The flipped classroom design was composed of two elements: a set of online resources intended to be completed in preparation for the plenary session (the lecture), and the re-framing of the plenary session to embrace an active learning design requiring students' preparation and participation in collaborative problem solving tasks (Lage et al., 2000; Pardo & Mirriahi, 2017).

The second dataset comes from a course called Code Yourself! (CY), which was delivered on the Coursera platform in 2015. The MOOC was designed to introduce teenagers to computer programming, while covering basic topics in software engineering and computational thinking. The course lasted 5 weeks and, of a total enrollment of 59,900 students, discussion forum data was available for 1,421 students. The content consisted of lecture videos, quizzes and peer-assessed programming projects. If students scored at least 50% in their coursework, they were deemed to have passed, while a distinction was awarded to students receiving a score of 75% or more.

Participation in the discussion forum was not required in either course. Forum activity in ICS consisted of 536 threads, comprised of 1,115 posts. Activity in CY, by contrast, consisted of 774 threads, comprised of 5,950 posts. Summary statistics of the two forums are provided in Table 1.

Table 1: Forum Summary Statistics

Statistic	ICS	CY
Thread Count	536	774
Post Count	1,115	5,950
Average Thread Length	2.08	7.69
Average Sub-thread Length	1.47	1.94
Average Posts Per Student	4.91	4.01

The courses were selected as they provide two disparate learning contexts for assessing the construct validity of typical SNA methods. In particular, the two courses exhibit drastic differences regarding structure: while ICS is a blended learning environment where the students are likely to have offline connections not captured by the discussion forum, CY is a MOOC where students are likely to interact solely through the discussion forum. This difference is particularly salient since ICS involves offline, collaborative problem solving. Furthermore, ICS is considerably longer, lasting 13 weeks compared to just 5 for CY. Pedagogy also differs: in ICS instructors mediate and interact with students in the forum, with the intention of prompting in-depth discussion of the relevant concepts. By contrast, no such

mediation exists in CY. These differences and others frame two different contexts and are essential for understanding and interpreting the different social relationships that arise within them.

3.2 Tie Extraction

Ties were extracted using the six tie definitions. Self-ties were excluded in all cases.

Direct Reply Ties (Figure 2.1): The author of each post was connected with the author of its parent post. Concretely, for each thread in the discussion forum and each post within each thread, if a post was either an instructor answer, a student answer, or a level-two post (that is, a reply to a thread starter), a tie was created from the level-two poster to the thread starter. However, if a post was classified as a level-three post (a reply to a level-two post), a tie was created from the level-three poster to the author of the parent level-two post (to which the post was directed).

Star Reply Ties (Figure 2.2): The author of each level-two and level-three post was connected with the author of the thread starting post. To be more concrete, for each thread in the discussion forum, the thread starter was identified and, for each post in the thread, a tie was created from the poster to the thread starter.

Direct Reply & Star Ties (Figure 2.3): Ties defined in both Direct Reply and Star Reply were included but the same tie was never counted more than once. Specifically, for each thread in the discussion forum and each post within each thread, if the post was an instructor answer, a student answer, or a level-two post, a tie was created from the poster to the thread starter. In the case of level-three post, a tie was created from the level-three poster to the level-two poster and, if they were not one and the same person, the thread starter too.

Total Co-presence (Figure 2.4): All authors in the same thread were connected with each other. In this case, ties are considered to be undirected.

Limited Co-presence (Figure 2.5): All users in small threads (<5 replies) were connected to each other with undirected ties; in larger threads users were connected to all other users in their sub-thread and the thread starter only. For threads of five or more posts, if a post was a level-two post (a reply to a thread starter) the level-two poster was connected to the thread starter. However, if level-three posts exist (posts replying to a level-two post), then each level-three post was linked to each other, the level-two post, and the thread starter.

Moving Window (Figure 2.6): All posts within a moving window of size N were connected to each other. Concretely, for each thread in the discussion forum, if the number of posts within a thread was less than N , then a tie was created between each post. Otherwise, starting from the first post, the first N posts are selected, and an undirected tie was created between each post, then the moving window moved to the second post and the next N posts were selected and ties created. This process repeated until the window reached the end of the thread. In this study we investigated windows of sizes two through five.

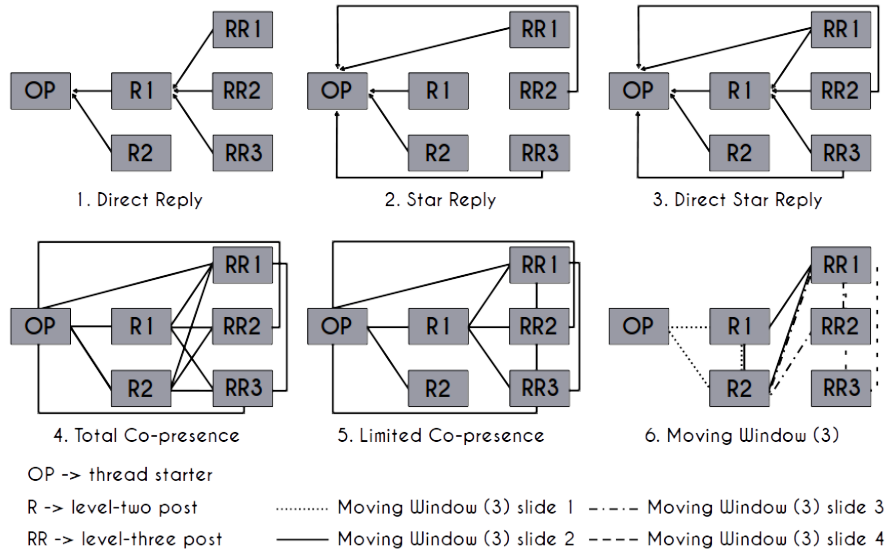


Figure 2: Six Social Tie Definitions

Given our definitions, each type of social tie resulted in the number of connections as show in Table 2:

Table 2: ICS and CY Tie Counts

Tie Definition	ICS	CY
Direct Reply	567	3,386
Star Reply	564	2,641
Direct & Star Reply	588	3,761
Total Co-presence	3,757	126,225
Limited Co-presence	589	5,960
Moving Window (5)	969	9,884
Moving Window (4)	863	8,134
Moving Window (3)	727	6,152
Moving Window (2)	520	3,630

3.3 Analysis

3.3.1 Social Network Analysis

To address our research questions, networks were extracted for the two courses in accordance with the six tie definitions. Social network analysis was then conducted across all networks in two complementary phases: structural and statistical network analysis.

Our analysis of the networks' structural features relied on some of the most commonly used node-level SNA metrics to characterize centrality, including degree, closeness, betweenness, and eigenvector

centrality. Degree centrality captures the local structure surrounding the node, and indicates the number of connections (in- and out-going for reply-based networks) a node has (Freeman, 1979). For this reason, degree is often considered a measure of popularity (Carrington et al., 2005). Closeness centrality measures the distance of a given node to all other nodes in a network (Freeman, 1979) and so can be viewed as a measure of each node's potential to connect with other nodes. Betweenness measures the number of shortest paths between all other nodes that a given node lies on, and so can be viewed as a metric of brokerage or the extent to which a node bridges distinct parts of the network. Finally, eigenvector centrality gives greater prominence to a node the more it is connected to other highly prominent nodes. Accordingly, it can be viewed as a ranked metric of influence.

Additionally, we investigated structural features at the network-level, including density, diameter, and average path length. Within a network, density measures the proportion of actual connections between nodes to all possible connections and so can be viewed as a measure of the extent to which all members of a network are connected to each other (Wasserman, 1994). Diameter measures the maximum eccentricity of any node in a network, that is, the maximum distance between any two nodes. Finally, average path length measures the average number of steps along the shortest paths for all possible pairs of network nodes.

For the statistical analysis of networks, ERGMs were used to reveal a variety of network statistics and investigate network formation processes. In particular, we aimed to investigate the effects of reciprocity, popularity, and transitivity. As a network statistic, reciprocity represents the tendency of students to form mutual ties and group together (Morris et al., 2008). In the context of our datasets and tie definitions, this would indicate whether or not students tend to continue interaction with their peers who replied to their posts. As this metric represents directed loops of length two, it only applies in the case of reply-based networks (that is, networks with directed edges). Popularity was modeled by the *geometrically weighted degree distribution* (*gwdegree* and *gwodegree* in reply-based networks; *gwdegree* in co-participation-based networks, which have undirected edges). *Gwdegree* is a statistic that geometrically discounts the value of incoming ties when the indegrees are summed in the statistic or, more intuitively, captures a popularity effect. *Gwodegree* considers the number of ties an individual sends in the network, and captures structures that result from highly active nodes. Transitivity refers to the extent to which the relation that ties two connected nodes in a network is transitive: that is, the extent to which the friend of my friend is also my friend. This statistic is represented by the *geometrically weighted edgewise shared partner distribution* (*gwesp*).

For each of the networks we consider a variety of models. In the case of reply-based networks (Direct Reply, Star Reply, Direct Star Reply), we examined a model for each of our statistics of interest (*reciprocity*, *gwdegree*, *gwodegree*, and *gwesp*). Similarly, for co-participation-based networks (Total Co-presence, Limited Co-presence, Moving Window 5, 4, 3, and 2), we examined a model for each of our statistics of interest (*gwdegree* and *gwesp*). Models were then analysed on the basis of goodness-of-fit statistics. Networks were extracted using the *ergm* 3.8.0 (Hunter et al., 2008) R package.

3.3.2 Regression Analysis

To examine the association between academic performance and our node-level measures of centrality, and so answer our second research question, we conducted a regression analysis. In the case of reply-based networks, we examined six metrics and, in the case of co-participation-based networks, which do not distinguish between in- and out-going ties, we examined four. Since our dependent variable, the course outcome, was measured differently in our two datasets, two distinct approaches were required. In the case of ICS, the dependent variable was continuous, so a linear regression model was fitted. By contrast, the CY course outcome was categorical (obtained certificate). Accordingly, multinomial logistic regression, a method which explains the association between a categorical dependent variable and one or more continuous independent variables (Cramer, 2003), was adopted. To investigate this association, four models were fitted for each dataset. Each model included the dependent variable (course result), one of the centrality measures, and, in order to control for an activity effect, a variable representing an individual's forum post count. Multinomial logistic regressions were performed using the *mlogit* 0.2-4 R package (Croissant, 2013). In the case of ICS, the dependent variable was heavily skewed, and all independent variables across both courses appeared to following a power law distribution; they were therefore log transformed.

4 RESULTS

4.1 Networks Formed by Six Tie Definitions

Descriptive statistics for both datasets reveal clear distinctions between reply-based tie definitions and those based on co-participation. Network centrality metrics were calculated by averaging over the node-level values for each network and centrality type, except for eigenvector centrality. Being more akin to a ranking measure, this network metric was calculated as the sum of differences between each node's eigenvector centrality and the maximum eigenvector centrality, divided by the maximum possible such value.

Reply-based definitions (Direct Reply, Star Reply and Direct Star Reply) produced networks of striking similarity regarding centrality metrics at both node and network level (Tables 3 and 4). Regarding co-participation-based definitions, the range of Moving Windows investigated exhibit clear trends across all centrality and network level metrics. As the window size decreased, the number of connections to each was, on average, attenuated, as was the overall density of the network. Furthermore, the distribution of influence across the networks, as measured by eigenvector (network) centrality increased. Compared to the other co-participation-based tie definitions, Total Co-Presence produced dramatically distinct networks in both datasets. By contrast, Limited Co-presence resulted in very similar networks to the Moving Window 3 definition across both courses.

Table 3: ICS Network Descriptive Statistics

Descriptives	DR	SR	DSR	TC	LC	MW(5)	MW(4)	MW(3)	MW(2)
Degree	2.66	2.65	2.76	35.28	5.53	9.10	8.10	6.83	4.88
Closeness	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Betweenness	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
Eigenvector	207.31	207.53	207.20	131.59	190.95	188.66	190.49	192.33	195.14
Density	0.013	0.012	0.013	0.166	0.026	0.043	0.038	0.032	0.023
Diameter	13	13	11	6	6	6	6	6	6
Average Path Length	6.66	5.92	6.04	2.04	2.46	2.37	2.41	2.48	2.72

Table 4: CY Network Descriptive Statistics

Descriptives	DR	SR	DSR	TC	LC	MW(5)	MW(4)	MW(3)	MW(2)
Degree	2.38	1.86	2.65	117.66	8.39	13.91	11.45	8.66	5.11
Closeness	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Betweenness	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Eigenvector	1388.19	1380.01	1386.76	1036.18	1368.68	1347.94	1359.07	1370.76	1385.39
Density	0.00	0.00	0.00	0.13	0.01	0.01	0.01	0.01	0.0
Diameter	10	11	9	5	8	6	7	8	13
Average Path Length	4.31	3.98	4.18	2.11	2.923	2.93	3.06	3.29	3.90

4.2 Statistical Networks Properties

In both courses under investigation, the three reply-based tie definitions produced networks with largely consistent statistical properties. Estimated coefficients are presented in Table 5. Across all reply-based definitions and datasets, *gwesp* was insignificant indicating an absence of transitive ties. This consistency across two disparate learning environments is surprising: in the case of ICS, students interact in both the discussion forum and the face-to-face plenary sessions, and accordingly one might expect the derived networks to resemble those emerging from social media, where transitive ties are a *sine qua non* (Nick et al., 2013). Regarding reply-based networks, the effect of reciprocity was significant in all networks across both courses, except for Star Reply in ICS. Across all tie definitions and courses, the effects of popularity and activity, as measured by *gwidegree* and *gwodegree*, were strong, negative, and highly significant, indicating an absence network structures characterized by highly popular or active agents.

In the case of co-participation-based networks, results across both courses were consistent in that all of the investigated network processes were predominantly absent, and the baseline model provided the best fit. The only exception to this pattern was for Moving Window 5 and Moving Window 4 in the CY course. Here the derived networks exhibited evidence of transitive ties. In future research, it is worth investigating to what extent this transitivity is induced by selective mixing, as increasing the likelihood of within category ties provides opportunities for completed triangles within categories, especially when groups are small, as the low density in Table 4 indicates (Goodreau et al., 2009).

Table 5: Reply-based Network Properties

Properties	ICS						CY					
	DR		SR		DSR		DR		SR		DSR	
	Est.	S.E.	Est.	S.E.	Est.	S.E.	Est.	S.E.	Est.	S.E.	Est.	S.E.
<i>Baseline (Edges)</i>	-4.36*	0.04	-4.37*	0.04	-4.33*	0.04	-6.39*	0.02	-6.64*	0.02	-6.64*	0.02
<i>Reciprocity</i>	1.18*	0.31	0.72	0.39	1.10*	0.31	3.26*	0.15	2.73*	0.21	2.73*	0.21
<i>Gwdegree</i>	-3.19*	0.15	-3.21*	0.15	-3.26*	0.14	-2.06*	0.06	-1.38*	0.08	-1.38*	0.07
<i>Gwdegree</i>	-2.96*	0.05	-3.07*	0.14	-3.00*	0.15	-4.91*	0.07	-8.66*	0.09	-8.66*	0.09
<i>Gwesp</i>	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

Note: * $p < .001$

Table 6a: ICS Co-participation-based Network Properties

Properties	TC		LC		MW(5)		MW(4)		MW(3)		MW(2)	
	Est.	S.E.	Est.	S.E.	Est.	S.E.	Est.	S.E.	Est.	S.E.	Est.	S.E.
<i>Baseline (Edges)</i>	-1.61*	0.02	-3.62*	0.04	-3.10*	0.03	-3.23*	0.03	-3.40*	0.04	-3.75*	0.04
<i>Gwdegree</i>	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
<i>Gwesp</i>	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

Note: * $p < .001$

Table 6b: CY Co-participation-based Network Properties

Properties	TC		LC		MW(5)		MW(4)		MW(3)		MW(2)	
	Est.	S.E.	Est.	S.E.	Est.	S.E.	Est.	S.E.	Est.	S.E.	Est.	S.E.
<i>Baseline (Edges)</i>	-1.94*	0.00	-5.13*	0.01	-4.62*	0.01	-4.81*	0.01	-5.09*	0.01	-5.62*	0.01
<i>Gwdegree</i>	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
<i>Gwesp</i>	NA	NA	NA	NA	8.55*	0.16	3.30*	0.06	NA	NA	NA	NA

Note: * $p < .001$

4.3 Centrality and Academic Achievement

Assessing the association between node-level centrality measures, forum activity, and academic outcomes revealed further differences between the tie definitions. In the case of reply-based networks in the ICS dataset (Table 7a), no centrality metrics exhibited any significant association with course performance. However, for all centrality metrics except for in-degree, activity was significantly and positively associated with course performance (albeit with a small coefficient).

In the case of co-participation-based networks in the ICS dataset (Tables 7b & 7c), no centrality metrics except closeness (Total Co-presence) and eigenvector (Moving Window 2) were significant. However, for all tie definitions except for Limited Co-presence (degree), and Moving Windows 4, 3, and 2 (eigenvector), activity was significantly and positively associated with course performance. In the case of Total Co-presence, closeness centrality was significantly and negatively associated with course performance, indicating that as the mean distance between nodes decreased, academic outcomes

suffered. However, it should also be noted that these assessments across centrality measures effectively constitute multiple comparisons which have not been controlled for. Given the number of estimated parameters, the occurrence of some significant parameters at the 5% level is likely even under the null hypothesis.

Table 7a: ICS Reply-based Linear Regression Results

	DR		SR		DSR	
	Est.	S.E.	Est.	S.E.	Est.	S.E.
<i>In Degree</i>						
Degree	0.07	0.07	0.06	0.07	0.07	0.07
Activity	0.12	0.06	0.12	0.06	0.12	0.06
<i>Out Degree</i>						
Degree	0.01	0.06	-0.01	0.05	0.01	0.05
Activity	0.16**	0.06	0.17**	0.05	0.16**	0.06
<i>In Closeness</i>						
Closeness	507.25	1719.78	3929.00	5649.00	470.30	1702.00
Activity	0.16**	0.05	0.15**	0.05	0.16**	0.05
<i>Out Closeness</i>						
Closeness	1038.00	2168.00	1190.00	2360.00	1038.00	2150.00
Activity	0.16***	0.05	0.15**	0.05	0.16***	0.05
<i>Betweenness</i>						
Betweenness	0.23	0.84	0.80	1.91	0.34	0.90
Activity	0.16**	0.05	0.15**	0.05	0.16**	0.05
<i>Eigenvector</i>						
Eigenvector	0.98	0.71	0.88	0.75	1.01	0.72
Activity	0.14**	0.05	0.15**	0.05	0.14**	0.05

Note: * $p < .05$, ** $p < 0.01$, *** $p < .001$

Table 7b: ICS Co-participation-based Linear Regression Results

	TC		LC		MW(5)	
	Est.	S.E.	Est.	S.E.	Est.	S.E.
<i>Degree</i>						
Degree	-0.03	0.02	0.06	0.09	-0.03	0.05
Activity	0.19***	0.05	0.12	0.08	0.19**	0.06
<i>Closeness</i>						
Closeness	-257.61*	100.93	-154.02	162.27	-86.19	163.60
Activity	0.21***	0.05	0.19***	0.05	0.18**	0.05
<i>Betweenness</i>						
Betweenness	1.49	8.46	-1.37	1.97	2.48	7.89
Activity	0.16**	0.05	0.17***	0.05	0.15**	0.06

<i>Eigenvector</i>						
Eigenvector	-0.18	0.12	-0.15	0.77	0.13	0.84
Activity	0.18***	0.05	0.18*	0.07	0.16*	0.07

Note: * $p < .05$, ** $p < 0.01$, *** $p < .001$

Table 7c: ICS Co-participation-based Linear Regression Results

	MW(4)		MW(3)		MW(2)	
	Est.	S.E.	Est.	S.E.	Est.	S.E.
<i>Degree</i>						
Degree	-0.04	0.06	-0.03	0.08	0.21	0.11
Activity	0.20**	0.06	0.19*	0.07	0.01	0.09
<i>Closeness</i>						
Closeness	-47.85	159.67	-7.23	154.14	179.98	140.54
Activity	0.17**	0.05	0.17**	0.05	0.13*	0.05
<i>Betweenness</i>						
Betweenness	5.26	8.11	5.77	7.17	4.45	4.96
Activity	0.14*	0.06	0.14*	0.06	0.14**	0.05
<i>Eigenvector</i>						
Eigenvector	0.44	1.03	0.94	1.06	2.18*	0.99
Activity	0.13	0.09	0.10	0.09	0.02	0.07

Note: * $p < .05$, ** $p < 0.01$, *** $p < .001$

For the CY dataset, in the case of reply-based networks (Table 8a), in-degree centrality was significantly associated with obtaining a certificate of distinction (Direct - $\chi^2(60) = 179.18$, $p < 0.001$; Star - $\chi^2(54) = 213.43$, $p < 0.001$; Direct Star - $\chi^2(70) = 178.77$, $p < 0.001$), but it did not have a significant impact upon the likelihood of obtaining a normal certificate. By contrast, for Direct Reply and Direct Star Reply, out-degree centrality increased the likelihood of obtaining a normal certificate (Direct - $\chi^2(64) = 168.85$, $p < 0.001$; Direct Star - $\chi^2(72) = 192.82$, $p < 0.001$) but not a certificate of distinction. Betweenness centrality was significantly and negatively associated with course performance across all reply-based networks (Direct - $\chi^2(756) = 1098.20$, $p < 0.001$; Star - $\chi^2(360) = 584.01$, $p < 0.001$; Direct Star - $\chi^2(696) = 1010.50$, $p < 0.001$). Specifically, increases in betweenness significantly reduced the likelihood of obtaining a certificate with distinction.

In the case of co-participation-based networks in the CY dataset (Tables 8b & 8c), nodes ranked higher by eigenvector centrality were significantly less likely to obtain either certificate in the cases of Total and Limited Co-presence (Total Co-presence - $\chi^2(1384) = 1953.20$, $p < 0.001$; Limited Co-presence - $\chi^2(1840) = 2265$, $p < 0.001$) but were significantly more likely to obtain either certificate across all other definitions (Moving Window 5 - $\chi^2(2634) = 2737.70$, $p = 0.078$; Moving Window 4 - $\chi^2(2654) = 2738.00$, $p = 0.125$; Moving Window 3 - $\chi^2(2682) = 2750.80$, $p = 0.173$; Moving Window 2 - $\chi^2(2610) = 2679.30$, $p = 0.167$). Increases in activity significantly increased the likelihood of obtaining both a distinction and a normal certificate for all metrics except for eigenvector centrality where influence was only significantly associated with a normal certificate for Moving Window 4 and 3.

Table 8a: CY Reply-based Multinomial Regression Results

	DR		SR		DSR	
	Est.	S.E.	Est.	S.E.	Est.	S.E.
<i>In Degree</i>						<i>D</i>
Distinct (deg)	0.29*	0.14	0.76***	0.17	0.31*	0.13
Normal (deg)	-0.10	0.21	0.30	0.24	-0.07	0.19
Distinct (act)	0.91***	0.12	0.79***	0.11	0.91***	0.11
Normal (act)	0.87***	0.17	0.74***	0.16	0.85***	0.17
<i>Out Degree</i>						
Distinct (deg)	0.07	0.11	0.12	0.10	0.11	0.10
Normal (deg)	0.33*	0.15	0.23	0.14	0.31*	0.15
Distinct (act)	1.02***	0.12	1.01***	0.11	0.98***	0.12
Normal (act)	0.53**	0.18	0.65***	0.16	0.54**	0.18
<i>In Closeness</i>						
Distinct (close)	934950.00	497760.00	2273700.00	1302200.00	93170.00	497690.00
Normal (close)	316290.00	758370.00	513790.00	1981100.00	316610.00	758270.00
Distinct (act)	1.04***	0.10	1.05***	0.09	1.04***	0.09
Normal (act)	7.86***	0.14	0.79***	0.14	0.79**	0.14
<i>Out Closeness</i>						
Distinct (close)	109090.00	89692.00	235150.00*	11355.00	109160.00	89657.00
Normal (close)	316240.00*	139270.00	528180.00**	168350.00	316060.00*	139220.00
Distinct (act)	1.00***	0.11	1.00***	0.10	1.00***	0.11
Normal (act)	0.61***	0.16	0.55***	0.16	0.61***	0.16
<i>Betweenness</i>						
Distinct (bet)	-45.15**	14.28	-46.12*	23.52	-45.44**	14.65
Normal (bet)	-113.42	74.36	-105.08	107.66	-146.91	93.71
Distinct (act)	1.18***	0.10	1.12***	0.10	1.18***	0.10
Normal (act)	0.98***	0.16	0.87***	0.15	1.01***	0.16
<i>Eigenvector</i>						
Distinct (eigen)	6.73**	2.46	3.57	2.19	6.41**	2.40
Normal (eigen)	-0.68	3.80	0.18	3.21	-0.20	3.66
Distinct (act)	0.87***	0.11	0.98***	0.11	0.89***	0.11
Normal (act)	0.87***	0.17	0.82***	0.16	0.85***	0.17

Note: * $p < .05$, ** $p < 0.01$, *** $p < .001$; Reference levels for each analysis was "None" - i.e. student did not obtain a certificate

Table 8b: CY Co-participation-based Multinomial Regression Results

	TC		LC		MW(5)	
	Est.	S.E.	Est.	S.E.	Est.	S.E.
<i>Degree</i>						
Distinct (deg)	-0.14***	0.04	0.00	0.09	0.05	0.10

Normal (deg)	-0.05	0.06	-0.13	0.14	0.13	0.16
Distinct (act)	1.14***	0.09	1.07***	0.12	1.03***	0.12
Normal (act)	0.81***	0.14	0.91***	0.18	0.70***	0.18
<i>Closeness</i>						
Distinct (close)	35347.39	27880.28	32963.69	28063.95	45736.00	29815.00
Normal (close)	85267.53	65480.71	78347.90	62951.64	91833.00	68420.00
Distinct (act)	1.06***	0.09	1.06***	0.09	1.06***	0.09
Normal (act)	0.77***	0.14	0.77***	0.14	0.77***	0.14
<i>Betweenness</i>						
Distinct (bet)	-34.36*	17.06	-21.25	12.37	-25.68**	8.62
Normal (bet)	-143.51	83.05	-52.66	53.16	-137.71*	60.41
Distinct (act)	1.13***	0.10	1.13***	0.10	1.16***	0.10
Normal (act)	0.99***	0.17	0.90***	0.16	1.13***	0.18
<i>Eigenvector</i>						
Distinct (eigen)	-1.67***	0.30	-3.72***	1.06	13.08***	2.26
Normal (eigen)	-1.25**	0.46	-6.26**	2.39	6.62*	3.10
Distinct (act)	1.18***	0.10	1.16***	0.10	0.31*	0.16
Normal (act)	0.87***	0.14	0.94***	0.14	0.52*	0.23

Note: * $p < .05$, ** $p < 0.01$, *** $p < .001$; Reference levels for each analysis was "None" - i.e. student did not obtain a certificate

Table 8c: CY Co-participation-based Multinomial Regression Results

	MW(4)		MW(3)		MW(2)	
	Est.	S.E.	Est.	S.E.	Est.	S.E.
<i>Degree</i>						
Distinct (deg)	0.09	0.11	0.19	0.14	0.25	0.18
Normal (deg)	0.13	0.18	0.14	0.21	0.07	0.28
Distinct (act)	1.01***	0.13	0.92***	0.14	0.88***	0.17
Normal (act)	0.70***	0.20	0.69**	0.21	0.75**	0.25
<i>Closeness</i>						
Distinct (close)	46786.00	30046.00	48880.00	30462.00	55103.00	31639.00
Normal (close)	90892.00	67927.00	93322.00	69012.00	104740.00	74237.00
Distinct (act)	1.05***	0.10	1.05***	0.09	1.05***	0.09
Normal (act)	0.77***	0.14	0.76***	0.13	0.76***	0.14
<i>Betweenness</i>						
Distinct (bet)	-23.60**	7.78	-20.05**	7.01	-16.36**	5.89
Normal (bet)	-147.05*	60.07	-154.93**	59.19	-129.18**	49.11
Distinct (act)	1.16***	0.10	1.15***	0.10	1.14***	0.10
Normal (act)	1.16***	0.18	1.19***	0.18	1.17***	0.18
<i>Eigenvector</i>						
Distinct (eigen)	17.53***	2.70	22.04***	3.22	19.96***	3.43

Normal (eigen)	8.65*	3.64	10.33*	4.34	10.57*	4.84
Distinct (act)	0.17	0.16	0.10	0.16	0.38*	0.15
Normal (act)	0.49*	0.23	0.49*	0.24	0.53*	0.22

Note: * $p < .05$, ** $p < 0.01$, *** $p < .001$; Reference levels for each analysis was "None" - i.e. student did not obtain a certificate

5 DISCUSSION

5.1 Structural Network Properties

Descriptive statistics (Section 4.1) for both datasets clearly partitioned social ties according to reply-based and co-participation-based definitions. While these two tie types produced distinct network structures, there remained some notable intra-type variations. In particular, Total Co-presence produced dramatically different networks with high values of degree centrality. This finding, in keeping with Wise et al. (2017), suggests that Total Co-presence should be used with caution due to the disproportionate influence it assigns to large threads. By contrast, in both datasets Limited Co-presence produced networks comparable to Moving Window, in particular Moving Window 3. There are a number of possible explanations for this similarity, such as the predominance of short threads and associated sub-threads in both datasets (see Table 1).

5.2 Statistical Network Properties

In addressing our first research question, there is evidence that across different learning contexts and pedagogies, variations in social tie formation mechanisms may produce different statistical properties in the derived networks. In the case of reply-based networks, we investigated the propensity of networks to form directed loops of length two (that is, reciprocal ties), a popularity and activity effect (whether or not the degree distribution, affected by popular or active agents, contributes to the likelihood of the observed network, captured by *gwidegree* and *gwodegree*), and the extent to which the friend of my friend is also my friend (that is, transitive ties, here captured by *gwesp*). Across all tie definitions, the results were broadly consistent for our statistics of interest, except for Star Reply in the ICS course, where reciprocal ties were notably absent.

In the case of co-participation-based networks, we investigated the propensity of networks to exhibit a popularity effect (whether or not the degree distribution, affected by popular agents, contributes to the likelihood of the observed network, captured by *gwdegree*), and the propensity for transitive ties to form (in the case of undirected edges, this represents the average probability that two neighbors of a vertex are themselves nearest neighbors). The co-participation-based tie definitions we investigated produce a number of transitive (closed) triangles within each thread and sub-thread. However, a transitive effect is only identified in the case of Moving Window 5 and 4. The predominance of "NA" values in Tables 6a and 6b is the result of model degeneracy.

The absence of reciprocal ties in Star Reply (ICS) may in part be explained by both the definition itself, whereby all ties are from responders to a thread starter, and the relatively low student (on an absolute basis) and thread count of the ICS dataset compared to the CY dataset; since reciprocal ties are not

created within Star Reply threads, a larger student and thread count provides greater opportunity for reciprocity to manifest itself. Beyond this exception, the estimated coefficients are broadly consistent on an intra-course basis. While the estimates in CY are quite high, these are in line with those found in existing studies (e.g. Joksimović et al., 2016; Kellogg et al., 2014; Lusher et al., 2012). As Lusher et al. (2012) have argued, it appears that a strong reciprocal effect may be seen as a defining characteristic of interaction in online social networks in general. Furthermore, Lusher et al. (2012) identified such networks as “self-disclosing”, characterized by strong relations between nodes. In these networks, students self-disclose themselves to establish *social presence* (Cobb, 2009; Richardson et al., 2017), creating comfortable learning and knowledge sharing environments. However, the low network level cohesion, evidenced by low density, indicates that students commonly interact with smaller groups (Scott, 2017).

Across all reply-based tie definitions and courses, the effects of popularity and activity (as measured by *gwidegree* and *gwodegree*, respectively) were strong, negative, and significant. Such an effect could indicate that within the network, the distribution of popularity and activity were largely homogeneous, rather than being centralized on in- or out-degree. Regarding popularity, this result is consistent with existing studies (e.g. Joksimović et al., 2016; Kellogg et al., 2014; Lusher et al., 2012). In the case of reply-based CY definitions, where reciprocity and a negative popularity effect were particularly strong, the interpretation is quite intuitive: rather than be concentrated in a few individuals, the high propensity of students to engage with each other on a reciprocal basis distributes the effects of popularity over the population.

5.3 Centrality and Academic Achievement

Regarding our second research question, our results indicate that the choice of tie definition can affect the observed association between centrality and academic performance. For instance, for all reply-based networks in the CY dataset, in-degree centrality significantly improved the likelihood of obtaining a certificate of distinction. However, for the very same networks, betweenness centrality significantly decreased the likelihood of obtaining a certificate of distinction.

Given these findings, it is important to assess the underlying assumptions that give rise to such inconsistencies. For instance, while in-degree centrality was significantly and positively associated with obtaining a distinction in all CY reply-based networks, the relation was reverted in the case of Total Co-presence. This may be because the construct being measured differs between the two definitions. In the case of reply-based networks, in-degree centrality indicates the extent of social prominence. However, in the case of Total Co-presence, degree centrality measures the extent of shared interest. While shared interest increases with thread size, social prominence is diluted, which could account for the contrary associations. Similarly, in the CY dataset, eigenvector centrality rank across all tie definitions (except Star Reply) was significantly associated with obtaining a distinction, but the direction of the association depends upon the tie definition. While this association was positive for most definitions, it was negative for Total and Limited Co-presence. This may be because reply-based and, to a limited extent, Moving Window definitions represent the purposeful, direct exchange of information. Total and Limited Co-presence, by contrast, dilute this effect, place inflated importance on large threads, and so provide

limited information for assessing influence. These findings give cause to reiterate the warning provided by Wise et al. (2017) that Total Co-presence and, to a lesser extent, Limited Co-presence should be used with caution.

While our results indicate that the choice of tie definitions can affect observed associations, it is important to emphasize that these comparisons have been made at the overall network level, not for specific individuals. It remains unclear to what extent individual centrality metrics are consistent across definitions. This is an interesting avenue of future research, and is an essential consideration when seeking to identify individuals with certain social status.

5.4 Learning and Pedagogical Context

While inconsistent associations between centrality metrics and performance may be in part explained by tie definition, they may also be attributable to differences in learning and pedagogical context. Regarding such contextual factors, the two courses analysed exhibit important differences. For instance, while ICS is a blended learning environment where students interact both inside and outside the discussion forum, CY is a MOOC where the forum is students' only point of contact. This could result in MOOC interactions being characterized more by Q&A than in-depth discussions. While a content analysis would have to be conducted to ascertain this in our dataset, the literature provides some evidence in favour of such a hypothesis: for instance, Gillani & Eynon (2014) found that forums harbour crowds, not communities of learners; networks were fragmented and became increasingly so over the duration of the course. Furthermore, although instructor mediation existed in ICS, no such mediation occurred in CY. Poquet et al. (2017) found that different facilitation strategies mediated a reciprocal effect whereby non-facilitated forums were characterized by direct reciprocal ties. Our findings replicate this result: a reciprocal effect was present across all reply-based tie definitions but was almost three times stronger in the case of CY.

Pedagogy may also have impacted on student behaviour: for instance, ICS involved a collaborative problem solving exercise in the plenary session, which could have led learners to participate differently in the discussion forum. Participation may also be affected by contextual factors. For instance, while students created far more posts in the CY forum, students, on average, posted a similar number of times in each course. However, a small proportion of MOOC users even participated in the forum, let alone consistently; not only did a far greater proportion of ICS participate, but they also had offline connections formed over a far longer course period (13 weeks compared to just 5).

While our research questions did not directly address the impact of pedagogy and learning context, such factors likely played an important role and should be explicitly addressed before any conclusions or comparisons can be made between courses. However, our results also point to the importance of selecting an appropriate tie definition for a given research goal. For instance, reply-based tie definitions emphasize the purposeful, directed exchange of information between individuals and the derived networks may be useful for identifying roles or influence within a group. Co-participation-based definitions, such as Total and Limited Co-presence, instead treat all ties within a thread as homogenous and focus on identifying shared interest. Our more novel tie definition, Moving Window, has a number

of appropriate applications, depending on the course context and window size. For instance, it may be useful within a collaborative learning context where posts within a thread are strongly related to recent posts within the same thread. However, such a structure is highly idealized: not only do learners not always read discussions chronologically (Hewitt, 2003), but in an asynchronous, many-to-many discussion forum, messages may refer to several others appearing far earlier in the chain (Gruzd & Haythornthwaite, 2008). That said, the Moving Window definition emphasises the temporal structure of threads, a potentially important aspect which is overlooked by more conventional tie definitions.

6 IMPLICATIONS

This study investigated the construct validity of a number of social tie definitions. Such ties purport to characterize the relationships formed between individuals on the basis of interactions within an online discussion forum. Structural validity was assessed by comparing the definitions on the basis of the structural and statistical properties of the networks they induced from our datasets. Our tie definitions could be categorized into two types (reply-based relationships and co-participation-based relationships, respectively) and while we found broadly consistent structural and statistical properties within these two categories, across category comparisons revealed striking differences.

External validity was assessed by investigating how measures of network centrality were associated with academic performance. While we found that increased social centrality was predominantly associated with opportunities and improved academic outcomes for students, there were some notable exceptions including significant, negative associations. This would indicate that external validity cannot be assured, and that the choice of tie definition does matter.

We also assessed the extent to which our findings generalized by conducting our analyses in two distinct learning settings. Regarding the structural properties of networks, we found reply-based tie definitions produced strikingly similar node-level centrality measures, even in spite of the considerable differences in course context and scale. This was not reflected in the case of co-participation-based networks although, in both contexts, Total Co-presence produced vastly inflated figures. Cross-context similarities were also found within the networks' statistical properties: for reply-based tie definitions, both courses exhibited a significant, negative popularity effect counter-balanced by a significant, positive reciprocal effect. Regarding the association between centrality and academic performance, most metrics enjoyed consistent associations. However, there were deviations, particularly Total Co-presence which, compared to other definitions, in some cases exhibited the opposite association. Though, in many cases, our results generalized across the definitions under study, the departures from this consistency indicate that the validity of SNA methods cannot be assured, and researchers should proceed with caution.

Our results lend support to the argument that researchers should be transparent in their choice of tie definition and, moreover, provide justification for their choice. Given the impact that tie definitions can have, it is advised that researchers try a number of different methods to ascertain the extent to which such methodological choices can bias their results. On the basis of this study, we recommend future SNA researchers pursue an exploratory comparison of Total Co-presence with a reply-based definition, as this could produce contrasting results and provide clarity on the internal validity of their chosen methods.

REFERENCES

- Ba-Omar, H., Petrounias, I., & Anwar, F. (2007). A framework for using web usage mining to personalise e-learning. *Proceedings of the 7th IEEE International Conference on Advanced Learning Technologies (ICALT '07)*, 937-938. <http://dx.doi.org/10.1109/ICALT.2007.13>
- Batool, K., & Niazi, M. A. (2014). Towards a Methodology for Validation of Centrality Measures in Complex Networks. *PloS ONE*, 9(4), e90283. <http://dx.doi.org/10.1371/journal.pone.0090283>
- Boroujeni, M. S., Hecking, T., Hoppe, H. U., & Dillenbourg, P. (2017). Dynamics of MOOC Discussion Forums. *Proceedings of the Seventh International Learning Analytics & Knowledge Conference (LAK '17)*, 128-137. <http://doi.acm.org/10.1145/3027385.3027391>
- Carrington, P. J., Scott, J., & Wasserman, S. (2005). *Models and methods in social network analysis*, 28. Cambridge, UK: Cambridge University Press.
- Cho, H., Gay, G., Davidson, B., & Ingraffea, A. (2007). Social networks, communication styles, and learning performance in a CSCL community. *Computers & Education*, 49(2), 309-329. <http://dx.doi.org/10.1016/j.compedu.2005.07.003>
- Cobb, S. C. (2009). Social Presence and Online Learning: A Current View from a Research Perspective. *Journal of Interactive Online Learning*, 8(3), 241-254.
- Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Chicago: Rand McNally.
- Cramer, J. S. (2003). The standard multinomial logit model, in *Logit Models from Economics and Other Fields*, Cambridge University Press, 104-125. <https://doi.org/10.1017/CBO9780511615412.008>
- Croissant, Y., (2013). *mlogit: multinomial logit model*.
- Dado, M., & Bodemer, D. (2017). A review of methodological applications of social network analysis in computer-supported collaborative learning. *Educational Research Review*, 22, 159-180. <http://dx.doi.org/10.1016/j.edurev.2017.08.005>
- Dawson, S., Tan, J. P. L., & McWilliam, E. (2011). Measuring creative potential: Using social network analysis to monitor a learners' creative capacity. *Australasian Journal of Educational Technology*, 27(6), 924-942. <http://dx.doi.org/10.14742/ajet.921>
- Dawson, S., Gašević, D., Siemens, G., & Joksimović S. (2014). Current state and future trends: A citation network analysis of the learning analytics field. *Proceedings of the Fourth International Conference on Learning Analytics And Knowledge (LAK '14)*, 231-240. <https://doi.org/10.1145/2567574.2567585>
- Dawson, S. (2008). A Study of the Relationship between Student Social Networks and Sense of Community. *Educational Technology & Society*, 11(3), 224-238.
- De Laat, M., & Prinsen, F. (2014), Social learning analytics: Navigating the changing settings of higher education. *J. Res. Pract. Assess.*, 9(4), 51-60.
- N. Dowell, N., Skrypnik, O., Joksimović, S., Graesser, A. C., Dawson, S., Gašević, D., de Vries, P., Hennis, T., & Kovanović, V. (2015). Modeling Learners' Social Centrality and Performance through Language and Discourse. *Proceedings of the 8th International Conference on Educational Data Mining*, 250-257.

- DuBois, C., Butts, C., & Smyth, P. (2013). Stochastic blockmodeling of relational event dynamics. *Proceedings of the Sixteenth International Conference on Artificial Intelligence and Statistics*, 238–246.
- Frank, O. D., & Strauss, D. (1986). Markov Graphs. *Journal of the American Statistical Association*, 81(395), 832–842. <http://dx.doi.org/10.1080/01621459.1986.10478342>
- Freeman, L. C. (1979). Centrality in social networks conceptual clarification. *Soc. Netw.*, 1(3), 215–239. [http://dx.doi.org/10.1016/0378-8733\(78\)90021-7](http://dx.doi.org/10.1016/0378-8733(78)90021-7)
- Gašević, D., Dawson, S., Rogers, T., & Gasevic, D. (2016). Learning analytics should not promote one size fits all: The effects of instructional conditions in predicting learning success. *The Internet and Higher Education*, 28, 68–84. <http://dx.doi.org/10.1016/j.iheduc.2015.10.002>
- Gašević, D., Kovanović, V., Joksimović, S., & Siemens, G. (2014). Where is research on massive open online courses headed? A data analysis of the MOOC Research Initiative. *Int. Rev. Res. Open Distrib. Learn.*, 15(5). <http://dx.doi.org/10.19173/irrodl.v15i5.1954>
- Gillani, N., & Eynon, R. (2014). Communication patterns in massively open online courses. *The Internet and Higher Education*, 23, 18–26. <http://dx.doi.org/10.1016/j.iheduc.2014.05.004>
- Gillani, N., Yasserli, T., Eynon R., & Hjorth, I. (2014). Structural limitations of learning in a crowd: communication vulnerability and information diffusion in MOOCs. *Scientific Reports*, 4:6447. <http://dx.doi.org/10.1038/srep06447>
- Goodreau, S., Kitts, J., & Morris, M. (2009). Birds of a Feather, or Friend of a Friend? Using Exponential Random Graph Models to Investigate Adolescent Social Networks*. *Demography*, 46(1), 103–125. <http://dx.doi.org/10.1353/dem.0.0045>
- Gruzd, A., & Haythornthwaite, C. (2008). Automated discovery and analysis of social networks from threaded discussions. *International Sunbelt Social Network Conference*.
- Hewitt, J. (2003). How habitual online practices affect the development of asynchronous discussion. *Journal of Educational Computing Research*, 28(1), 31–45. <http://dx.doi.org/10.2190/PMG8-A05J-CUH1-DK14>
- Hunter, D. R., Handcock, M. S., Butts, C. T., Goodreau, S. M., & Morris M. (2008). ergm: A Package to Fit, Simulate and Diagnose Exponential-Family Models for Networks. *J. Stat. Softw.*, 24(3), 1–29. <http://dx.doi.org/10.18637/jss.v024.i03>
- Jiang, S., Fitzhugh, S. M., & Warschauer, M. (2014). Social Positioning and Performance in MOOCs. *Proceedings of the Workshops held at Educational Data Mining, 7th International Conference on Educational Data Mining (EDM 2014)*, 1183, 14.
- Joksimović, S., Dawson, S., Manataki, A., Kovanović, V., Gašević, D., & Friss de Kereki, I. (2016). Translating network position into performance: Importance of centrality in different network configurations. *Proceedings of the Sixth International Conference on Learning Analytics And Knowledge (LAK '16)*, 314–323. <http://dx.doi.org/10.1145/2883851.2883928>
- Jovanović, J., Gašević, 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. <http://dx.doi.org/10.1016/j.iheduc.2017.02.001>
- Jordan, K. (2015). Synthesising MOOC completion rates | MoocMoocher. [Online]. Available: <https://moocmoocher.wordpress.com/2013/02/13/synthesising-mooc-completion-rates/>.

- Kellogg, S., Booth, S., & Oliver, K. (2014). A social network perspective on peer supported learning in MOOCs for educators. *The International Review of Research in Open and Distributed Learning*, 15(5). <http://dx.doi.org/10.19173/irrodl.v15i5.1852>
- Kovanović, V., Joksimović, S., Gašević, D., & Hatala, M. (2014). What is the source of social capital? The association between social network position and social presence in communities of inquiry. *Proceedings of the Workshops held at Educational Data Mining 2014, co-located with the 7th International Conference on Educational Data Mining (EDM'14)*.
- Kovanović, V., Gašević, D., Dawson, S., Joksimović, S., Baker, R. S., & Hatala, M. (2015). Does Time-on-task Estimation Matter? Implications for the Validity of Learning Analytics Findings. *Journal of Learning Analytics*, 2(3), 81-110. <http://dx.doi.org/10.18608/jla.2015.23.6>
- Krackhardt, D. (1998). Super Strong and Sticky. *Power and Influence in Organisations*, 21. <http://dx.doi.org/10.4135/9781483345291.n2>
- Krackhardt, D. (1999). The Ties that Torture: Simmelian Tie Analysis in Organizations. *Research in the Sociology of Organizations*, 16, 183-210.
- Lage, M. J., Platt, G. J., & Tregua, M. (2000). Inverting the classroom: A gateway to creating an inclusive learning environment. *The Journal of Economic Education*, 31(1), 30-43. <http://dx.doi.org/10.2307/1183338>
- Loevinger, J. (1957). Objective tests as instruments of psychological theory. *Psychological Reports*, 3, 635-694. <http://dx.doi.org/10.2466/pr0.1957.3.3.635>
- Lusher, D., Koskinen, J., & Robins G. (2012). *Exponential Random Graph Models for Social Networks: Theory, Methods, and Applications*. Cambridge, UK: Cambridge University Press.
- Messick, S. (1995). Validity of Psychological Assessment: Validation of Inferences from Persons' Responses and Performances as Scientific Inquiry into Score Meaning. *American Psychologist*, 50(9), 741-749. <http://dx.doi.org/10.1037/0003-066X.50.9.741>
- Morris, M., Handcock, M. S., & Hunter D. R. (2008). Specification of Exponential-Family Random Graph Models: Terms and Computational Aspects. *Journal of Statistical Software*, 24(4). <http://dx.doi.org/10.18637/jss.v024.i04>
- Munk, M., & Drlík, M. (2011). Impact of different pre-processing tasks on effective identification of users' behavioral patterns in web-based educational system. *Procedia Computer Science*, 4, 1640-1649. <http://dx.doi.org/10.1016/j.procs.2011.04.177>
- Nick, B., Lee, C.-K., Cunningham, P., & Brandes, U. (2013). Simmelian backbones: amplifying hidden homophily in facebook networks. *Advances in Social Networks Analysis and Mining (ASONAM), 2013 IEEE/ACM International Conference*, 525-532. <http://dx.doi.org/10.1145/2492517.2492569>
- Pardo, A., Jovanović, J., Dawson, S., & Gašević, D. (in press). Using Learning Analytics to Scale the Provision of Personalised Feedback. *British Journal of Educational Technology*. <https://doi.org/10.1111/bjet.12592>
- Pardo, A., & Mirriahi, N. (2017). Design, Deployment and Evaluation of a Flipped Learning First-Year Engineering Course in: Reidsema C., Kavanagh L., Hadgraft R., & Smith N. (eds) *The Flipped Classroom*. Springer, Singapore. 177-191. http://dx.doi.org/10.1007/978-981-10-3413-8_11
- Poquet, O., & Dawson, S. (2017). Untangling MOOC learner networks. *Proceedings of the 6th International Conference on Learning Analytics & Knowledge (LAK '16)*, 208-212.

- Richardson, J. C., Maeda, Y., Lv, J., & Caskurlu, S. (2017). Social presence in relation to students' satisfaction and learning in the online environment: A meta-analysis. *Computers in Human Behavior*, 71, 402-417. <http://dx.doi.org/10.1016/j.chb.2017.02.001>
- Robins G., Pattison P., Kalish Y., & Lusher D. (2007). An introduction to exponential random graph (p*) models for social networks. *Social Networks*, 29(2), 173-191. <http://dx.doi.org/10.1016/j.socnet.2006.08.002>
- Schreurs, B., Teplovs, C., Ferguson, R., de Laat, M., & Buckingham Shum, S. (2013). Visualizing Social Learning Ties by Type and Topic: Rationale and Concept Demonstrator. *Proceedings of the Third International Conference on Learning Analytics and Knowledge (LAK '13)*, 33-37. <http://dx.doi.org/10.1145/2460296.2460305>
- Scott, J. (2017). *Social Network Analysis, Fourth Edition*. London: SAGE Publications.
- Shulman, L. S. (1970). Reconstruction of educational research. *Review of Educational Research*, 40, 371-396.
- Skrypnik, O., Joksimović, S., Kovanović, V., Gašević, D., & Dawson, S. (2015). Roles of course facilitators, learners, and technology in the flow of information of a cMOOC. *International Review of Research in Online and Distance Learning*, 16(3), 188-217.
- Wang P., Robins G., Pattison P., & Lazega E. (2013). Exponential random graph models for multilevel networks. *Social Networks*, 35(1), 96-115. <http://dx.doi.org/10.1016/j.socnet.2013.01.004>
- Wasserman S., & Pattison P. (1996). Logit models and logistic regressions for social networks: An Introduction to Markov graphs and p. *Psychometrika*, 61(3), 401-425. <http://dx.doi.org/10.1007/BF02294547>
- Wasserman S. (1994). *Social network analysis: Methods and applications*. Cambridge University Press.
- Wise, A. F., Cui, Y., & Jin, W. Q. (2017). Honing in on social learning networks in MOOC forums: examining critical network definition decisions. *Proceedings of the Seventh International Conference on Learning Analytics And Knowledge (LAK '17)*, 383-392. <http://dx.doi.org/10.1145/3027385.3027446>
- Zhang, J., Skryabin, M., & Song, X. (2016). Understanding the dynamics of MOOC discussion forums with simulation investigation for empirical network analysis. *Distance Education*, 37(3), 270-286.
- Zhu, M., Bergner, Y., Zhang, Y., Baker, R., Wang, Y., & Paquette, L. (2016). Longitudinal engagement, performance, and social connectivity: a MOOC case study using exponential random graph models. *Proceedings of the Sixth International Conference on Learning Analytics and Knowledge (LAK '16)*, 223-230. <http://dx.doi.org/10.1145/2883851.2883934>

2.3 Summary

In discussion forums, social ties purport to characterise the relationships formed between individuals on the basis of interactions. How best to define these ties, however, is an open question in the research literature, and the decision may depend on a range of contextual factors. Perhaps because of this difficulty, the impact this decision may have on the results of any analyses has received limited attention in the learning analytics literature and, if mentioned at all, studies often provide little justification for their choice. The present study has demonstrated how problematic this reticence can be: the choice of tie definition substantially alters the structural and statistical properties of a network and, in extreme cases, may lead to the opposite conclusions being drawn. For instance, while we found that across most tie definitions increased social centrality was associated with improved academic outcomes for students, other definitions resulted in significant, negative associations.

This result is particularly relevant in the light of Tinto's (1975) model. This model makes two key theoretical predictions: namely, that students' interactions with the academic and social systems present in their educational institution are associated with their academic outcomes and their persistence in their studies. In examining students' social system, the use of methods drawn from SNA is an obvious choice. However, the results presented in this chapter suggest that our choice of tie definition may influence the outcome of this analysis. While this is a question that we leave for the introduction to Chapter 3, the present study has made an important contribution to the literature in highlighting the concerns regarding validity that this seemingly minor methodological decision can create. As the field of learning analytics matures, validity is of paramount importance for building an empirically sound and robust body of knowledge.

3

Performance and Persistence in Co-enrolment Networks

3.1 Introduction

HAVING emphasised, in the previous chapter, the importance of ensuring validity within the study of social networks, we now turn our attention to the analysis of the social domain described in Tinto's (1975) model. According to this model, students' initial commitments to their academic institution are continually modified, in part, by their integration into the social networks within which they find themselves. Tinto (1975) posits that successful integration enhances not only students' commitment to their studies, but also their academic outcomes. While this claim has been the subject of numerous empirical validations (Munro, 1981; Pascarella & Terenzini, 1980; Amaury Nora, 1990; Nora & Rendon, 1990; Terenzini & Pascarella, 1977), none of these have examined the model from the perspective of students' local neighbourhood and their role within that neighbourhood. Accordingly, in this chapter, we investigate the extent to which students' local network and position within it are not only associated with their academic performance, but are also predictive of their dropout decisions.

In investigating students' co-enrolment networks, we do so at a scale not previously examined, and analyse over three decades of undergraduate and postgraduate student enrolment data from a research-intensive public university in Australia. However, many of the oft-used methods from social network analysis (SNA) are constrained by their high computational complexity, rendering them impractical for a dataset of such size. To circumvent these limitations, substantial research has been committed to developing novel network embedding techniques to generate low-dimensional vector representations of nodes that offer potentially richer representations of network structure than the single, scalar values afforded by common centrality metrics (Cui, Wang, Pei, & Zhu, 2017).

3.1.1 Graph-Embeddings

As networks grow in size and complexity, the computation of traditional network measures becomes prohibitively complex. Accordingly, the central problem in machine learning on graphs is to encode information about graph structure into a lower dimensional space. In recent years, the field has received considerable attention and a large number of methods have been developed (Cui et al.,

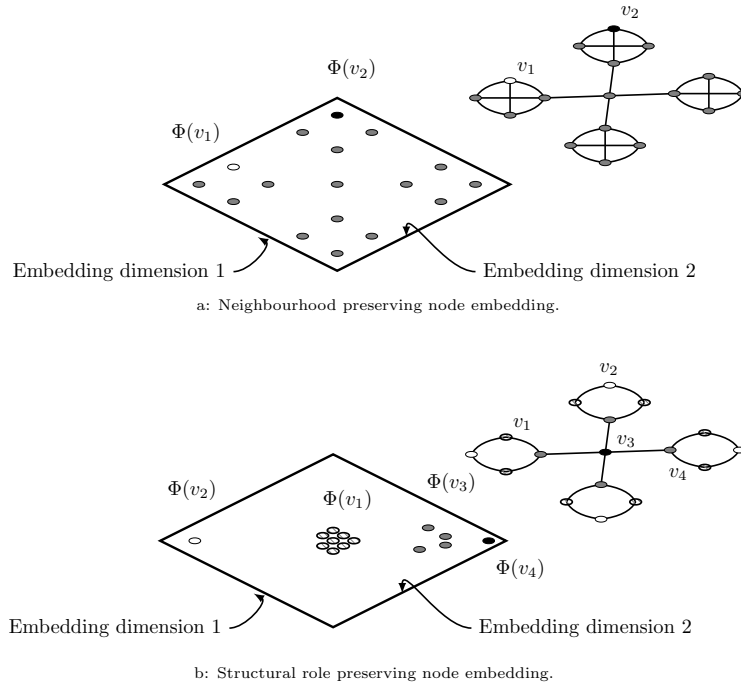


Figure 2. Graph-embedding procedures map nodes into a latent space where a notion of similarity between pairs of nodes is preserved. This similarity may relate to how the local neighbourhood is constituted (DeepWalk), or the structural role that a given node plays in the graph (Role2Vec).

2017). Much of this recent work has sought to provide approximate factorisations of the adjacency matrix whilst retaining a time complexity that is linear in the size of the input.

In the current chapter, we utilise two such methods – *DeepWalk* (Perozzi, Al-Rfou, & Skiena, 2014) and *Role2Vec* (Ahmed et al., 2018) – which learn mapping functions to capture the local neighbourhood and structural role of students, respectively. For instance, in the case of DeepWalk, the trained model captures similarities between neighbourhoods such that students with similar neighbourhoods will be situated in close proximity within the embedding space (see Figure 2). Importantly, this allows us to use the embedding for downstream tasks, namely, the prediction of students’ final grades and their dropout decisions. Accordingly, we use latent representations of students’ local neighbourhood and their structural role to evaluate the claims of Tinto’s (1975) model.

3.1.2 Validity

While the graph-embedding methodology that we propose is certainly novel within the learning analytics literature, it is also essential to ensure that our analysis is valid. We evaluate the validity of our method using two distinct approaches. On the one hand, we assess how performance on the grade and dropout prediction tasks is influenced by changes in the parameterisation of DeepWalk and Role2Vec; on the other, our analysis is underpinned by the assumption that enrolment implies attendance, an assumption which is readily violated (Kelly, 2012). To assess the impact such violation may have, we perform a corruption procedure randomly to prune our networks, be-

fore evaluating how this corruption influences the performance of the embeddings on the grade and dropout prediction tasks.

3.2 Publication: Persistence and Performance in Co-enrolment Network Embeddings

The following section includes the verbatim copy of the following publication:

Fincham, E., Rózemberczki, B., Kovanović, V, Joksimović, S., & Gašević, D. (2019). Persistence and performance in co-enrolment network embeddings. *Manuscript submitted for publication*

Persistence and Performance in Co-enrolment Network Embeddings

Ed Fincham, Benedek Rózemberczki, Vitomir Kovanović, Srećko Joksimović,
Jelena Jovanović, and Dragan Gašević

Abstract—In this paper, we empirically validate Tinto’s Student Integration model, in particular, the predictions the model makes regarding both students’ academic outcomes and their dropout decisions. In doing so, we analyse three decades’ worth of student enrolments at an Australian university and present a novel methodological approach using graph embedding techniques to capture both structural and neighbourhood based features of the co-enrolment network. In keeping with Tinto’s model, we find that not only do these embedded representations of students’ social network predict their final GPA, but also are able to successfully classify students who dropout. Our results show that these embedded representations of a student’s social network can achieve F_1 -scores of up to 0.79 when classifying dropout, and explain up to 10% of the variance in student’s final GPA. When controlling for a small set of covariates and variables common to the literature, this performance increases to 0.83 and 24%, respectively. Furthermore, the performance of these methods are robust to both changes in their parameterisation and to corruption of the underlying social networks. Importantly, this implies that hyper-parameters may be selected to reduce the computational demands of this method without loss of predictive power. The novelty of this method, and its ability to identify student dropout, merits further investigation to preemptively identify at-risk students.

I. INTRODUCTION

SINCE the turn of the new millennium, the number of students enrolled in higher education globally has more than doubled to over 200 million [1]. By 2040, this number is anticipated to grow to nearly 600 million, with much of this growth coming from East Asia and the Pacific [1]. While this trend offers improved employment prospects and lifelong learning opportunities to an unprecedented number of students, it also poses a number of challenges to higher education institutions. In particular, how can existing resources and infrastructure be leveraged to assist this growing student body? In addressing these concerns, the field of educational data science holds considerable promise, and the existing literature offers an array of methods for predicting key outcomes such as student success [2], [3], [4], [5], [6] and student retention [7], [8], [9]. This research provides higher education institutions

with an opportunity to not only develop early warning systems for at-risk students, but also offer personalised interventions.

There is also a large body of theoretical literature examining the processes and factors that lead to student success as well as student dropout. One of the most widely cited, Tinto’s Student Integration Model [10], [11], [12], posits that persistence in higher education is a temporal process, where students’ initial commitments to their programme and institution are mediated by a wide array of contextual factors. Once initiated, these commitments are continually modified by the students’ interactions with the academic and social systems of the institution.

While Tinto’s model has been the subject of numerous empirical validations [13], [14], [15], [16], [17], these have typically relied upon survey responses [16], or a range of data such as the time that students spent in class together, social network variables characterising the attrition and retention of the local network, and a demographic variables [18]. However, limited research has examined the model from the sole perspective of students’ social ties and their position in the social networks that those ties form. In particular, the use of methods from social network analysis (SNA) has been largely overlooked. Within educational research, SNA is a well-established field [19], [20], and provides a myriad of techniques to analyze the relationships that occur between individuals, groups, and communities [2], [19], [21]. While SNA provides a broad array of tools to help unpack the influence that students exert over each-other, and the impact that this has on learning [2], [22], the majority of studies utilizing SNA in education are primarily focused on examining structural regularities and structural properties, such as centrality and density [3], [4], [21], [23], [24]. For example, examining different centrality metrics can reveal the most influential actors, those individuals that enjoy more advantageous positions, and how patterns of interaction can affect learning [2], [21], [25].

However, one of the shortcomings of centrality measures is their limited applicability. As the size of networks grows, these traditional metrics pose several challenges to network processing and analysis due to their high computational complexity [26]. This limitations, however, may be circumvented with the use of methods drawn from the nascent machine learning literature on network embeddings [26], [27], [28]. These techniques embed a given network within continuous vector space, whilst retaining structural properties of the network, such that nodes with similar degree centrality and clustering coefficient value and distribution are located close to each other in the embedding space [29], [30], [31], [32].

Ed Fincham and Benedek Rózemberczki are with the School of Informatics, University of Edinburgh, United Kingdom.

Vitomir Kovanović and Srećko Joksimović are with the University of South Australia, 160 Currie St, Adelaide, SA 5000, Australia.

Jelena Jovanović is with the Faculty of Organizational Sciences, University of Belgrade, Serbia.

Dragan Gašević is with the Faculty of Information Technology, Monash University, Australia.

Manuscript received September 26, 2019. Corresponding author: E. Fincham (email: ed.fincham@ed.ac.uk).

Importantly, these approaches have been shown to retain the salient properties of graphs across a range of domains [27], [28], and allow social interactions to be analyzed using machine learning methods unsuited to graphical formats.

In investigating Tinto’s model [10], [11], [12], the present study makes a number of contributions. In particular, we evaluate two key predictions of the model – that students’ persistence and their academic outcomes are associated with their social networks – using methods and techniques drawn from the SNA literature. Within educational data mining, the ability to predict outcomes such as student dropout using some representation of social ties is a vital question that has received limited attention [33]. In investigating student dropout, we make use of co-enrolment networks, which are defined as student-student graphs where edges represent two students being enrolled in the same course; the strength of the tie is proportion to the number of courses the two share. We evaluate the association of these co-enrolment ties with student outcomes at a scale unprecedented in the educational research literature and present a novel methodological approach based on graph embedding techniques. We demonstrate that not only can these methods outperform more conventional network analysis techniques [34], but do so at a scale that would otherwise be prohibitive. We find that not only are students’ interactions predictive of their academic outcomes, but also of their dropout decisions. In addition to finding additional validation of Tinto’s model, the ability to predict student persistence on the basis of peer interactions opens up an interesting avenue of research into how higher education institutions might utilise this association.

II. THEORETICAL BACKGROUND

A. Student Integration

Decades of theoretical and empirical research into students’ experiences on university campuses has demonstrated that social interactions and peer culture play a central contextual role in understanding a variety of student outcomes [35], [36], [37]. One of the major, comprehensive conceptual models born of this research is Tinto’s Student Integration Model [10], [11], [12]. Drawing on Durkheim’s seminal work on suicide, Tinto’s model specifies a longitudinal process whereby a number of contextual factors (such as race, previous academic performance, family encouragement, etc.) interact to define students’ initial commitment to their studies. These commitments are then modified over time as a result of students’ interactions with the university community and the social structures within which they find themselves [12]. The constructs of students’ integration into the social and academic systems of an institution are at the model’s conceptual core, and [12] posits that successful integration enhances students’ commitment and positively influences not only their intended persistence in their studies, but also their academic outcomes.

Tinto’s [12] model has been widely cited and, with regards to student dropout, a number of studies have produced evidence supporting its construct and predictive validity [13], [14], [15], [16], [17]. Although the model was initially conceived for a university setting, it has since been used to

interpret attrition in distance education [38], [39], as well as online learning environments such as MOOCs [33].

Within this extensive body of literature, however, the use of SNA has been largely overlooked; a surprising lacuna, given the predominant role social integration plays in Tinto’s [12] model. A notable exception to this trend is the work of [18], who used the attrition and retention of first-, second-, and third-order neighbours to predict student dropout. The authors found that the persistence of students’ local network had a greater impact on their retention than any background or performance variables. Similarly, in the context of MOOCs, [33], used a range of common centrality measures, clustering coefficients, authority, and hub scores to predict student dropout. In a survival analysis, the authors found that only the authority and hub scores were associated with dropout. While this finding calls into question the predictive utility of centrality metrics with regards to student dropout, this may not hold in alternative educational contexts such as a university setting. Nevertheless, given the limited attention it has received in the literature, investigating student attrition through the lens of SNA merits consideration.

In addition to dropout, Tinto’s model predicts that students’ peer groups are related to their intellectual development and academic performance [12]. By contrast, this claim has been extensively studied and validated across a variety of educational contexts. In this case, such analyses have often relied upon methods drawn from SNA, and the influence of students’ position and structural role are often gauged using well-established centrality metrics such as degree, betweenness, and closeness [3]. Within this body of literature, significant associations have been identified between academic performance and, at varying times, degree centrality [23], closeness centrality [23], [40], [41], and betweenness centrality [23], [42].

While SNA holds much promise for empirically investigating Tinto’s model, such an approach is not without challenges. In particular, the academic social networks analysed to date have been limited in scope and have utilised modest datasets [34]. For large networks, however, the high computational complexity of these traditional network measures pose several challenges to network processing and analysis. For instance, of the standard metrics, only degree centrality has a linear time complexity, whereas betweenness centrality has at least quadratic time complexity, which becomes prohibitive as the number of vertices becomes large [43]. To circumvent such difficulties, substantial research has been committed to developing novel network embedding techniques to generate low-dimensional vector representations of nodes that offer potentially richer representations of network structure than the single, scalar values afforded by common centrality metrics [26]. Importantly, depending on the technique used, these representations can be generated in linear or quasi-linear time [26].

B. Graph Embeddings

As the number of vertices within a social network grows, not only does the computation of traditional network measures

become prohibitively complex, but also the application of common machine learning models becomes infeasible [26]. This is because such models assume that data can be represented by the adjacency matrix of a network and, as a network becomes larger, the high dimensionality renders any computation difficult [26]. Accordingly, the central problem in machine learning on graphs is finding a method to encode information about graph structure into a lower dimensional space.

Within this field, a number of approaches have been developed that learn representations which encode structural information about a given network [28], [44], [45]. The guiding principle behind such methods is to learn a mapping that embeds nodes as points in a low-dimensional (typically Euclidean) vector space [46]. These mappings are then optimised such that geometric relationships in the embedding space reflect the structure of the original graph [46]. Finally, these embeddings can then be used as feature inputs for downstream tasks; a property that we exploit in this study.

Early research into graph embedding techniques emphasised the utility of these methods for dimensionality reduction [47]. In a typical example, a similarity graph would be constructed for a set of nodes based on the pairwise neighbourhood overlaps. This graph would then be embedded in the lower dimensional vector space with the intention of keeping connected nodes in close proximity [47]. Examples of these techniques include Laplacian Eigenmaps [48] and Locally Linear Embedding (LLE) [49]. In these early days, scalability was a major concern, as methods often had a quadratic or cubic time complexity in the size of the graph.

In response to this limitation, recent research has focused on developing scalable techniques which leverage the sparsity of real-world datasets. Early work in this area, such as the Natural Graph Factorization algorithm [50], generate node embeddings from an approximate factorization of the adjacency matrix, and achieves a time complexity that is linear in the size of the input. This approach occurs over all observed edges rather than all possible edges and, being an approximation, may introduce noise into the solution. As an alternative, a number of graph embedding techniques have arisen which use random walks across nodes to approximate many properties of the graph. Given an appropriately large graph, such node-level embeddings may be calculated and used to estimate structural [31] or even neighbourhood properties [28], [45] of nodes within the network.

C. Co-enrolment Networks

Within higher education research, the influence that students exert over their peers is referred to as the “peer-effect”, and an extensive body of literature has explored the role such influence plays across a variety of network types, for instance, friendship networks [51], demographic networks [52], and course or class year networks [53].

In the case of co-enrolment networks, however, understanding how students influence their peers through the lens of network analysis has garnered limited attention [34], [54]. These networks are defined as a student-student graph where

edges represent two students being enrolled in the same course. Although node-level attributes are lacking, these edges may approximate the weak ties found within social networks [55]. However, the literature has found that students typically only maintain a limited set of social ties from their courses [56], [57], [58]. Accordingly, co-enrolment networks may induce considerable noise. Nevertheless, they provide ample opportunity for investigating social interactions, student dropout, and academic outcomes.

While co-enrolment studies have typically relied upon modest datasets [40], [59], a recent study by [34] analyzed the complete undergraduate co-enrolment network from a decade of students at a large American public university. By collecting 116 features including student demographics, course subject, credit hours, meeting days and times, the authors aggregated four feature sets (representing *mean*, *count*, *binary*, and *proportion* of first-order neighbour attributes). Using a simple OLS regression, these combined feature sets were found to explain approximately 8% of the variance in student performance across semesters [34]. While this result suggests that network-based features are associated with student performance, in large co-enrolment networks such rich node-level attributes may be unavailable. In these cases, graph embedding techniques offer a promising yet largely unexplored avenue for investigating the influence of student peer effects on academic performance and persistence.

III. RESEARCH QUESTIONS

While Tinto’s [12] model has been subjected to a number of empirical validations [13], [14], [15], [16], [17], none of these have investigated two of the model’s key claims – that students’ local social structure and their positioning within the larger social network are associated with their persistence in their studies and their academic outcomes – using methods and techniques drawn from the SNA literature. As to the former claim, investigating student dropout using such approaches has received limited attention beyond MOOC settings [33]. By contrast, the latter claim has been widely investigated, and a number of studies have found significant associations between social networks and academic performance [23], [34], [40], [42], [54], [59], [60], [61]. These studies have often been conducted on datasets of restricted scale. However, limited research has investigated the use of unsupervised graph embedding techniques to capture salient properties of large, co-enrolment networks as features for understanding not only how students’ social integration is associated with their academic performance, but also their persistence in their academic studies.

The present study seeks to address these lacunae by empirically validating Tinto’s model of student integration from a social network perspective. In doing so, we analyse student co-enrolment networks at a scale unprecedented in the educational research literature. While we are able to compute degree centrality at this scale – due to its linear time complexity – other centrality metrics are not computationally feasible. To circumvent these computational limitations, we investigate the ability of graph embedding techniques to capture salient

information about students’ local network structure. Specifically, we explore two distinct graph embedding techniques that represent the proximity of nodes within a neighbourhood [47] (DeepWalk [28]) and the structural similarity of neighbourhoods within the overall graph [62] (Role2Vec [27]). Using these methods, we empirically evaluate the central predictions of Tinto’s [12] model, and examine the extent to which embedded representations of students’ local network are predictive of their persistence and academic performance.

Research Question 1a: *To what extent does the analysis of large co-enrolment networks provide empirical validation for the theoretical predictions of Tinto’s model? Specifically, are embedded representations of these networks predictive of students’ persistence in their studies and final GPA? And how does this compare with the predictive power of more traditional methods of analyzing social ties, particularly degree centrality?*

It should be noted, however, that both embedding methods employed in this study are influenced by a range of hyper-parameters. Consequently, it is important to assess whether the predictive power of these embeddings is robust to changes in their parameterization.

Research Question 1b: *Is the predictive power of the network embeddings, with regards to students’ persistence and final GPA, sensitive to changes in the parameterization of the two models?*

An implicit assumption of the foregoing analysis is that enrolment implies attendance, which necessitates some degree of interaction between two students. But this assumption is readily violated: for instance, research into university lecture attendance suggests that student turnout can range from 28% to 78% [63]. Furthermore, in the present study, this assumption is known to be violated: the datasets analysed contain both university attendants and distance learners without distinction. Accordingly, a number of edges constitute false ties between distance and campus students co-enrolled in the same course, which is problematic as it calls into question the validity of any predictive modeling conducted on such a network. However, these concerns regarding validity may be addressed by ensuring that the results of any modeling are robust to the violation of our attendance assumption. This violation may be simulated via a corruption procedure that randomly removes edges from the networks prior to their embedding. Such perturbation studies are common practice within the graph representation literature, and are used to evaluate the robustness of a given model [44]. Predictive models are then trained on these corrupted embeddings to evaluate how robust

the predictive power of the embeddings are to the violation of the attendance assumption.

Research Question 2: *To what extent does the violation of our attendance assumption, operationalised through random corruption of the networks, influence the predictive power of our embeddings?*

IV. METHODS

A. Data Sources

The data used in this study consists of three decades of undergraduate and postgraduate student enrolment data (1989–2018) across three colleges within a research-intensive public university in Australia. These colleges include the Business college (BU), the Health Sciences college (HS), and the Information Technology and Engineering college (IT). For each student, the data consists of an identifying code and, for each course that they studied, the corresponding course code, the semester, the year the course was taken, and the student’s final grade for that course. To calculate a student’s final GPA score, their grades were weighted by course credits and then summed. The number of students enrolled at these schools ranged from 49,943 within IT (817,787 course enrolments), to 123,155 within BU (1,262,736 course enrolments).

B. Graph Embeddings

Prior to any analysis, the enrolment data was first transformed into an appropriate graphical format. Since each student attends a course in a given semester, we treat attendance as a bipartite graph where the two distinct sets of node represent students and courses taken. From this bipartite graph, a weighted projection was calculated forming a student-student graph where edges represent students who attended courses together, and weights represent the number of courses they shared. Tables I & II describe the properties of these graphs, along with the number of unique courses offered within each school, and the total number of course offerings over the entirety of each data-set. To investigate the predictive capacity of the co-enrolment networks using standard statistical techniques, it was first necessary to encode these relations within continuous vector space. This was achieved using two distinct graph embedding methods to capture neighbourhood-based and structural relations between students, respectively.

1) *Neighbourhood*: The first embedding procedure we employ, DeepWalk [28], is an unsupervised feature learning technique that uses information obtained from truncated random walks across a graph to learn a latent representation. Social relations are encoded within a continuous vector space such

Network	Students	Courses	Course Offerings	Total Enrolments	Edges	Projection Density
Business	123,155	4,189	23,899	1,262,736	91,287,333	0.0194
Health Sciences	64,437	4,498	22,772	882,235	23,999,775	0.0146
Information Technology and Engineering	49,943	6,629	36,690	817,787	11,201,346	0.0090

TABLE I: Descriptive Statistics of the Networks by School

Network	Grade	Course Count	Fail Count	Withdrawal Count	Degree
Business	4.219 (± 1.347)	13.657 (± 9.684)	0.610 (± 2.234)	0.440 (± 1.262)	1791.478 (± 1763.235)
Health Sciences	4.732 (± 1.228)	15.864 (± 11.312)	0.239 (± 1.252)	0.252 (± 0.915)	809.260 (± 691.247)
Information Technology and Engineering	4.404 (± 1.492)	16.401 (± 14.012)	0.701 (± 2.739)	0.701 (± 2.739)	449.393 (± 456.487)

TABLE II: Mean and Standard Deviation of Student Enrolment and Performance Statistics by School

that the latent features characterise the neighbourhood similarity and community membership of the vertices. Accordingly, students from the same cohort who took a large number of courses together should be proximal in the embedding space. Beyond first-order connections, high-order proximity is also preserved and students with high neighbourhood overlaps should also be close together (see Fig. 1a).

DeepWalk arose as a generalization of models such as *Word2Vec* [64], [65], which serve to estimate the likelihood of a given sequence of words appearing within a corpus. In this case, a graph is explored through a series of fixed-length, uniformly sampled random walks per node, with the intention of estimating the likelihood of observing a particular node, given the previous nodes visited thus far in the walk. However, the aim of DeepWalk is not only to estimate the probability distribution of node co-occurrences, but also to learn a latent representation [28]. Accordingly, the model includes a mapping function that characterises the latent representation for each node. As DeepWalk is trained, this mapping function is updated to maximise the probability of a given node co-occurring with its context, where this context is the local graph as characterised by a series of random walks originating from the node. Accordingly, the trained model generates representations that, for each node v , capture similarities in local graph structure between vertices, such that vertices with similar neighbourhoods will be situated in close proximity within the embedding space. In contrast to more traditional network centrality metrics, DeepWalk generates such representations in linear time [26]. For a comprehensive technical description of the model, refer to Appendix A.

2) *Structural*: In contrast to neighbourhood-based node embeddings, structural role embeddings create representations in a latent space such that the structural similarity of nodes is preserved within the embedding space [29], [30], [31], [32]. That is, nodes which have a similar degree, centrality, and clustering coefficient value and distribution in their neighbourhood are located close to each other in the embedding space. This would imply that students who bridge disparate neighbourhoods would be close to each other within the latent space. The same would hold for students located in the core of highly interconnected communities. Importantly, this method implies that proximity within the graph does not necessarily translate to proximity within the embedding (see Fig. 1b).

The structural node embedding procedure that we adopted, Role2Vec [27], is a generalization of the node and document embedding models DeepWalk [28] and *Paragraph2Vec* [66]. Since a structural role labelling mechanism is required for Role2Vec, we opted for the *Weisfeiler-Lehman* procedure [62]. Given an initial node labelling, this procedure iteratively

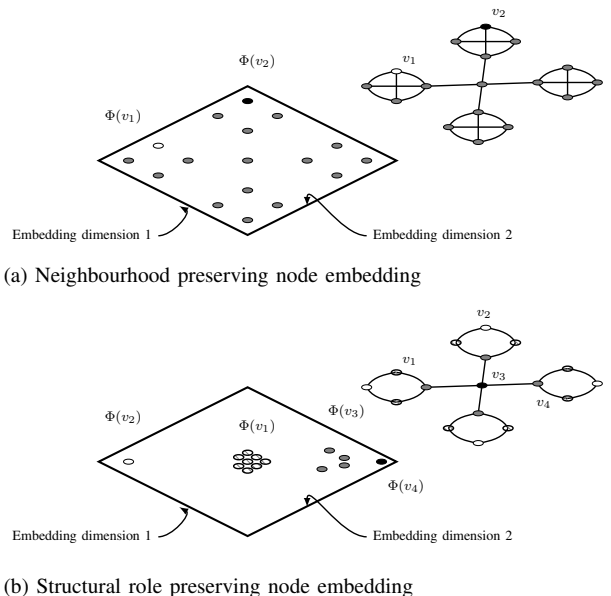


Fig. 1: Node embedding procedures map nodes into a latent space where a notion of similarity between pairs of nodes is preserved. In Fig. 1a, neighborhood node embedding procedures preserve proximity in the graph between pairs of nodes in the embedding space. The distance between nodes v_1 and v_2 is preserved approximately as they are represented by the coordinates $\Phi(v_1)$ and $\Phi(v_2)$. In Fig. 1b, structural node embedding procedures preserve structural similarity between pairs of nodes in the embedding space. The structural properties that are encoded when such an embedding is created include degree, centrality, local clustering coefficient, and the distribution of these quantities in the respective node's neighbourhood. The nodes v_1 , v_2 , v_3 and v_4 each represent different structural roles in the graph, which are distinguished by distinct colouring. Nodes that have similar or the same structural roles are clustered in the embedding space.

relabels the vertices. Concretely: for each node, labels of the node's neighbours are taken and sorted, and the sorted label sets are then used as the new labels. Nodes that have the same labelled neighbourhoods receive the same new label. Note that if the labelling is based upon some structural property, this procedure can capture the structural similarity of neighbourhoods. In order to capture this neighbourhood structure, we log-binned the degree of each node and treated the bin membership as an initial label for each node within

the graph.

To create this structural embedding, we first performed *Weisfeiler-Lehman* relabelling k times to generate node labels. In this way, each node had $k + 1$ labels which described the structural role of the node itself, and the structure of node neighbourhoods up to k hops distant. After this labelling procedure, we applied the Role2Vec embedding method, which performs a series of attributed, fixed-length random walks from every source node. From these attributed random walks we extracted features within a fixed window and trained a structural role embedding. This embedding is generated in linear time, with respect to the number of nodes and edges. For a comprehensive technical description of Role2Vec, refer to Appendix A.

Notably, both DeepWalk and Role2Vec were designed for unweighted graphs. However, both methods may be viewed as corner cases of *Node2Vec*, which generalises to weighted graphs [44]. This allows us to conduct our analysis on weighted graphs, as required by co-enrolment networks.

C. Research Question 1

To answer our first research question and investigate the extent to which the analysis of large co-enrolment networks validates the predictions of Tinto’s [12] model, we began by embedding these networks using the DeepWalk and Role2Vec methods outlined above. We then evaluated the derived embeddings, separately and combined, on the classification task of identifying students who dropped out, and on the regression task of grade prediction. These analyses were conducted three times: (1) once with just the embeddings as independent variables in our regression and classification models, (2) once with covariate features and variables common to the literature added to the model, and (3) once with just the covariate features. These covariates include the number of courses studied, the number of withdrawals, and the number of failed courses. We also calculated the degree centrality of each student. All variables were normalised with respect to their means to aid model convergence. To ascertain whether the performance of the combined model (2) was significantly different from the simple covariate model (3), we used the Wilcoxon signed-rank test [67].

For this initial investigation, default parameter values were used in both DeepWalk and Role2Vec. Specifically, we set $d = 32$ (number of dimensions), $r = 10$ (number of random walks), $l = 80$ (random walk length), $w = 6$ (context window size), and the optimization was run for a single epoch. For Role2Vec, we set $d = 32$, $r = 10$, and $l = 80$. These default parameter values were in keeping with those used in the studies where these models were introduced [27], [28]. For the classification task of identifying dropout, a standard logistic model was trained using 10-fold cross-validation on 80% labeled data, from which F_1 -scores were calculated. For the regression task of grade prediction, performance was assessed with a Lasso regression model trained using 10-fold cross-validation on 80% of the data with a search over $\alpha \in \{100, 10, 1, 0.1, 0.01, 0.001\}$. The choice of a Lasso model was motivated by its usage of L1 regularization, which encourages simple, sparse models and facilitates model selection [68]. All models were repeated for 30 random seed initializations, from which mean and standard deviation statistics were calculated.

To investigate the extent to which the predictive power of our two embedding methods was influenced by their respective parameterizations, we analyzed their sensitivity across several hyper-parameters. For each setting of each hyper-parameter, a new embedding was derived and evaluated on the dropout classification task and the grade prediction regression task (both with and without the inclusion of covariate and degree features). As before, each model was assessed over 30 runs with 10-fold cross-validation.

D. Research Question 2

To answer our second research question and investigate the extent to which the violation of our attendance assumption influences the predictive power of the network embeddings, we conducted a simulated corruption procedure. In this study, performance on the dropout classification and grade prediction tasks were measured as a function of the fraction of removed edges (relative to the original network). Specifically, edges were removed in increments of 10% up to 50%. Edges were chosen randomly, subject to the constraint that the number of connected components remained fixed. The rationale for this

Network	Embedding Method	Graph Embeddings Only				With Covariates & Degree				Only Covariates & Degree			
		F_1	$Prec.$	$Recall$	AUC	F_1	$Prec.$	$Recall$	AUC	F_1	$Prec.$	$Recall$	AUC
BU	DeepWalk	0.744	0.669	0.838	0.625	0.819	0.785	0.856	0.762	0.793	0.748	0.845	0.720
	Role2Vec	0.729	0.650	0.841	0.599	0.794	0.769	0.828	0.738				
	Combined	0.748	0.680	0.833	0.638	0.824	0.794	0.855	0.770				
HS	DeepWalk	0.789	0.667	0.966	0.525	0.818	0.753	0.895	0.670	0.811	0.709	0.947	0.606
	Role2Vec	0.789	0.662	0.977	0.516	0.814	0.745	0.898	0.657				
	Combined	0.789	0.670	0.960	0.533	0.827	0.764	0.901	0.687				
IT	DeepWalk	0.721	0.659	0.794	0.612	0.780	0.738	0.827	0.710	0.778	0.685	0.899	0.663
	Role2Vec	0.709	0.633	0.804	0.579	0.777	0.727	0.836	0.700				
	Combined	0.719	0.663	0.784	0.616	0.789	0.751	0.832	0.725				

TABLE III: Dropout Classification Performance of Embeddings with Default Parameter Settings

constraint was to prevent the creation of isolated communities which would present a fundamental change in the co-enrolment network structure. These corrupted networks were then embedded and performance was evaluated on the two tasks. Default parameter values were used for both models, as described in Section IV-C.

V. RESULTS

A. Research Question 1

Table III presents the results of the dropout classification task for the default parameter settings. Across all embeddings and datasets, the models exhibited minimal variance. While DeepWalk matched or outperformed Role2Vec across all three datasets, when the embeddings were combined the performance was either additive (BU), unchanged (HS), or inferior (IT). Comparing the combined model with the covariates only model, the Wilcoxon signed-rank test indicated that the performance of all combined models was significantly different to their covariate counterpart, with the exception of Role2Vec for the IT network ($Z=166, p=0.17$).

Table IV presents the results of the same task using only the covariate and degree feature set. In this case, the odds ratios are quite intuitive: while a unit increase in the number of courses taken vastly increases the chances of persisting in a degree of study, a unit increase in the number of failed courses decreases the chance of students in BU persisting by 16%. Interestingly, the odds ratio for degree centrality is approximately 1, implying that a unit increase in degree centrality results in a 1–2% change in classification outcome. Given this, the significant p-values for degree centrality are surprising, but may be attributed to the scale of the dataset.

Table V presents the results of the grade prediction task for the default parameter settings. Across all embeddings and datasets, the models exhibited minimal variance. All three datasets presented a similar pattern such that while DeepWalk alone outperformed Role2Vec, their performance was marginally additive when the two were combined, especially in the case of HS and IT. By contrast, Table VI presents the results of the same task using only the covariate and degree feature set. Notably, while degree centrality was a significant parameter across all three datasets, the standardised coefficients were extremely small, particularly compared to other variables such as course count. When the two feature sets were combined, however, the performance improvement was additive (see Table V), indicating that the embeddings, covariates, and degree were not highly correlated. Moreover, in the case of BU, the performance gain of combining the embedded features with the covariate and degree features was super additive; the variance explained by the combined feature set was greater than the sum of its parts. In the best model, the total feature set explained 24% of the variance in student grades.

To investigate the extent to which the predictive power of our two embedding methods is influenced by their respective parameterizations, we analyzed their sensitivity across several hyper-parameters. For each setting of each hyper-parameter, a new embedding was derived and evaluated on the same

dropout classification and grade prediction tasks (both with and without the inclusion of covariate and degree features). Performance was assessed with 10 runs of each model, each trained using 10-fold cross-validation. Due to the remarkable similarity of results across datasets and, in the interests of brevity, the results of the parameter sensitivity assessment are only displayed for the IT dataset.

Fig. 2 shows the effects of increasing l , the length of the random walks used to explore the graphs; r , the number of random walks; and w , the size of the context window (an additional parameter only applicable to the DeepWalk model). Across all datasets, embedding techniques, and prediction tasks, the results were very consistent, indicating that changes in these parameters had limited impact on the predictive capacity of the derived embeddings. Furthermore, in the case of the grade prediction task, the combination of the embeddings demonstrated a degree of additive performance.

B. Research Question 2

In addition to analyzing the parameter sensitivity of our approach, we investigated the extent to which the violation of our attendance assumption influenced the predictive power of the derived embeddings. This was conducted via a perturbation analysis involving the random removal of edges from each of our networks. The effects of random corruption on the predictive performance of the two embedding methods is shown in Fig. 3.

In the case of dropout classification, as shown in Fig. 3a, the models were highly robust to incremental corruption in the underlying data. Similarly, in the case of grade prediction, shown in Fig. 3b, the models were reasonably robust, exhibited minimal variation over the fraction of removed edges, and demonstrated additive performance. As with Section V-A, these results were remarkably similar across datasets so, in the interests of brevity, we only present the results of the corruption procedure for the IT dataset.

VI. DISCUSSION

A. RQ 1: Tinto’s Student Integration model

The purpose of the present study was to empirically investigate and attempt to validate the associations posited by Tinto’s model of student integration [12]. To briefly recap, Tinto’s model is a conceptual schema for university dropout, which views persistence as a longitudinal process of interactions between the student and the academic and social systems of the university [10]. The extent to which a student integrates into these two systems modifies their commitment to study and to the institution, ultimately informing their dropout decisions. Within this schema, peer-group interactions are an essential part of social integration, determining dropout decisions, but also modify students’ academic outcomes: namely, their grade performance.

In evaluating the predictions of Tinto’s model, we also leveraged the data collected by university institutions at an unprecedented scale. Due to the computational limitations of traditional SNA measures, this motivated exploring the ability of graph embedding techniques to represent salient

Network	Course Count			Withdrawal Count			Fail Count			Degree		
	OR	SE	p	OR	SE	p	OR	SE	p	OR	SE	p
BU	3.449	0.006	0.000	0.845	0.001	0.000	0.877	0.002	0.000	1.021	0.002	0.000
HS	2.705	0.006	0.000	0.874	0.001	0.000	0.940	0.001	0.000	1.023	0.001	0.000
IT	2.085	0.006	0.000	0.847	0.002	0.000	0.900	0.002	0.000	1.039	0.002	0.000

TABLE IV: Dropout Classification Performance of Covariate & Degree Features with Standardized Coefficients

information about students’ social networks; a novel use-case for such methods. In addition to these embedded representations, we also considered the impact that covariates and predictor variables common to the literature had on students’ dropout decisions and final GPA. Our results indicate that these two sets of features are complementary as their performance, shown in Tables III and V, is mostly additive. This is especially so for grade prediction where, in the case of BU, the combined performance of the embeddings, covariates, and degree is super-additive. While this is difficult to interpret due to the unsupervised nature of the embedded features, it could be partially addressed by investigating the correlation of these embeddings with ground truth labels, such as programme membership. In recent years, a variety of neural methods have also been developed for such a task [69].

By contrast, the predictive performance of the two embedding techniques, when used in tandem, is of negligible impact for dropout classification (see Table III), and is only weakly additive for grade prediction (see Table V). This implies that the neighbourhood features captured by DeepWalk are highly correlated with the structural features captured by Role2Vec. This is likely due to the dense structure of the co-enrolment networks (on average, 0.0143). While the density of our networks, reported in Table I, may seem low, within the context of social networks, they are unusually high [70]. Furthermore, as social networks become larger, they typically become sparser, and the density decreases [70]. By contrast, density in our networks increases as the networks grow. Given this dense connectivity structure, the high correlation between structural and neighbourhood features is not entirely surprising: since every community is fully connected, each is likely to be assigned a unique structural role label by the Weisfeiler-Lehman procedure, leading to a proliferation of

roles. As the number of structural role labels generated by the Weisfeiler-Lehman procedure tends towards the number of vertices, Role2Vec converges to baseline random walk methods such as DeepWalk [27].

Accordingly, the prevalence of fully connected communities within co-enrolment networks can lead to neighbourhood and structural features becoming equivalent. The additive, albeit weakly so, performance of the two embedding techniques implies that this is not entirely the case and indicates that certain nodes within communities have distinct labels, for instance, students that bridge otherwise disparate clusters within the network. Exploring what prompts these differences is an important question that future work should endeavour to answer.

Across all datasets, the Role2Vec embedding does not outperform the DeepWalk embedding in either the classification or the regression task. But considering the techniques these algorithms use, this result is unsurprising: while DeepWalk conducts truncated random walks at a node-level, Role2Vec first aggregates nodes into labels before randomly exploring across these labels; a necessarily less granular approximation. Although the fully-connected community structure of our co-enrolment networks may have resulted in the partial convergence of our two embedding procedures, *our results indicate that neighbourhood features are more important predictors of student dropout and their final GPA than structural features*. In other words, *the individuals that a student associates with have greater influence on their dropout decisions and final GPA than the number of students with whom they associate*. While this finding is in keeping with the existing literature [40], the results of the covariate and degree classification and regression models provide further evidence of this claim. In particular, the results for degree centrality, where the odds ratios being

Network	Embedding Method	Graph Embeddings Only		With Covariates & Degree		Covariates & Degree Only
		R^2	σ	R^2	σ	R^2
BU	DeepWalk	0.059	0.003	0.241	0.004	0.164
	Role2Vec	0.039	0.003	0.217	0.003	
	Combined	0.060	0.003	0.242	0.004	
HS	DeepWalk	0.089	0.004	0.215	0.005	0.131
	Role2Vec	0.052	0.004	0.174	0.005	
	Combined	0.095	0.005	0.227	0.004	
IT	DeepWalk	0.073	0.003	0.190	0.005	0.115
	Role2Vec	0.044	0.004	0.171	0.006	
	Combined	0.083	0.004	0.207	0.005	

TABLE V: Grade Prediction Performance of Embeddings with Default Parameter Settings

Network	Course Count			Withdrawal Count			Fail Count			Degree		
	Coef.	SE	p	Coef.	SE	p	Coef.	SE	p	Coef.	SE	p
BU	0.213	0.002	0.000	-0.006	0.001	0.000	-0.013	0.000	0.000	0.002	0.001	0.001
HS	0.098	0.001	0.000	-0.003	0.000	0.000	-0.003	0.00	0.000	-0.000	0.000	0.025
IT	0.108	0.002	0.000	-0.005	0.001	0.000	-0.010	0.001	0.000	0.001	0.001	0.018

TABLE VI: Predictive Performance of Covariate & Degree Features with Standardized Coefficients

almost one in the classification models, and the extremely small coefficients in the regression models (see Tables IV and VI) imply that taking courses with large numbers of people has minimal impact on student dropout and students' grades.

1) *Parameter Sensitivity of Embeddings*: The results of our sensitivity analysis indicate that the predictive power of our embeddings is highly robust to changes in their parameterization. In the case of the number of random walks and their length, this is likely due to the inter-connected structure of co-enrolment networks. Performing truncated random walks allows us to induce a distribution over nodes in the graph, given a particular node [71]. This distribution is, of necessity, an approximation, but converges in probability to the true distribution in the random walk length limit [71]. Nevertheless, the robustness of predictive performance with regards to the parameterization of these random walks implies that, even with atypical settings, this approximation is accurate. This is likely due to the dense connectivity structure characteristic of a co-enrolment network, where communities exhibit a degree of homogeneity.

2) *Implications*: The muted importance of structural features in our findings may be a result of the co-enrolment structure itself: since all students enrolled in the same course are linked, there is little scope within a given community for distinct structural roles. Nevertheless, our results provide substantial empirical evidence in favour of the theoretical predictions made by Tinto's [12]. This is especially the case with regards to student dropout, where all models achieved impressive results, and demonstrated weakly additive performance when combined with covariate and degree features. Furthermore, given the scale at which our analysis was conducted, the impressive F_1 -scores and the ability to explain up to 24% of variance in students' final GPA using only co-enrolment ties and simple covariate features presents a convincing case for the future use of graph embeddings within learning analytics research.

Regarding the parameter sensitivity of our embeddings, our result has important implications for practice as it demonstrates the robustness of these methods. This implies that good predictive performance may be attained in computationally cheap terms. This is particularly salient for research on enrolment where nodes and edges can number in the hundreds of thousands or millions [34].

B. RQ 2: Impact of Network Corruption

The impact of random network corruption – specifically the removal of edges – had *limited discernible impact on the predictive performance of the embeddings across both dropout*

and grade prediction tasks. Considering how the random corruption functions in the context of densely connected co-enrolment networks, and the impact this has on implicit matrix factorization techniques such as DeepWalk and Role2Vec [71], such results are unsurprising. For instance, within a co-enrolment network, the majority of edges exist within dense, fully-connected communities (such as a cohort of students on a programme with minimal elective modules). Accordingly, when removing an edge there is a high probability of selecting an edge within one such connected cluster rather than an edge bridging separate clusters. Removing an edge within a densely connected community will have limited impact on both local and global graph structure, which will not significantly impact the powers of the adjacency matrix nor node positions within the derived embedding. Given that embedded positions are largely unchanged, it is unsurprising that neither dropout nor grade prediction tasks suffered as a greater proportion of nodes were pruned from the network.

This investigation into the impact of network corruption was framed as a simulated violation of our attendance assumption. Since the data used in this study conflates university attendants with distance learners, an unknown proportion of the co-enrolment edges connect students that did not study together. While such corruption raises concerns regarding the validity of any modeling, the removal of edges allows us to evaluate whether this impacts on the predictive capacity of our embeddings. Given that both the dropout and grade prediction tasks were robust to the removal of up to 50% of edges, it is clear the violation of our attendance assumption does not undermine the utility or the validity of our method. However, our method is not only robust to this, but also outperforms existing co-enrolment approaches in the literature. For instance, while Gardner et. al [34] are able to explain approximately 8% of the variance in student GPA across semester, our approach explains up to 24% of variance in student GPA.

1) *Implications*: While a number of studies have analysed co-enrolment networks [34], [40], [54], [59], they have attracted limited attention compared to alternative network structures, such as discussion forums where social ties may be derived from individual interactions. This may in part be due to the substantial noise that co-enrolment networks may induce; a student may never interact with a large portion of their cohort. In the present study, the conflation of university attendants and distance students induces considerable noise. The robust performance of our models to the simulation of this noise amply demonstrates the practical and research viability of co-enrolment networks within higher education.

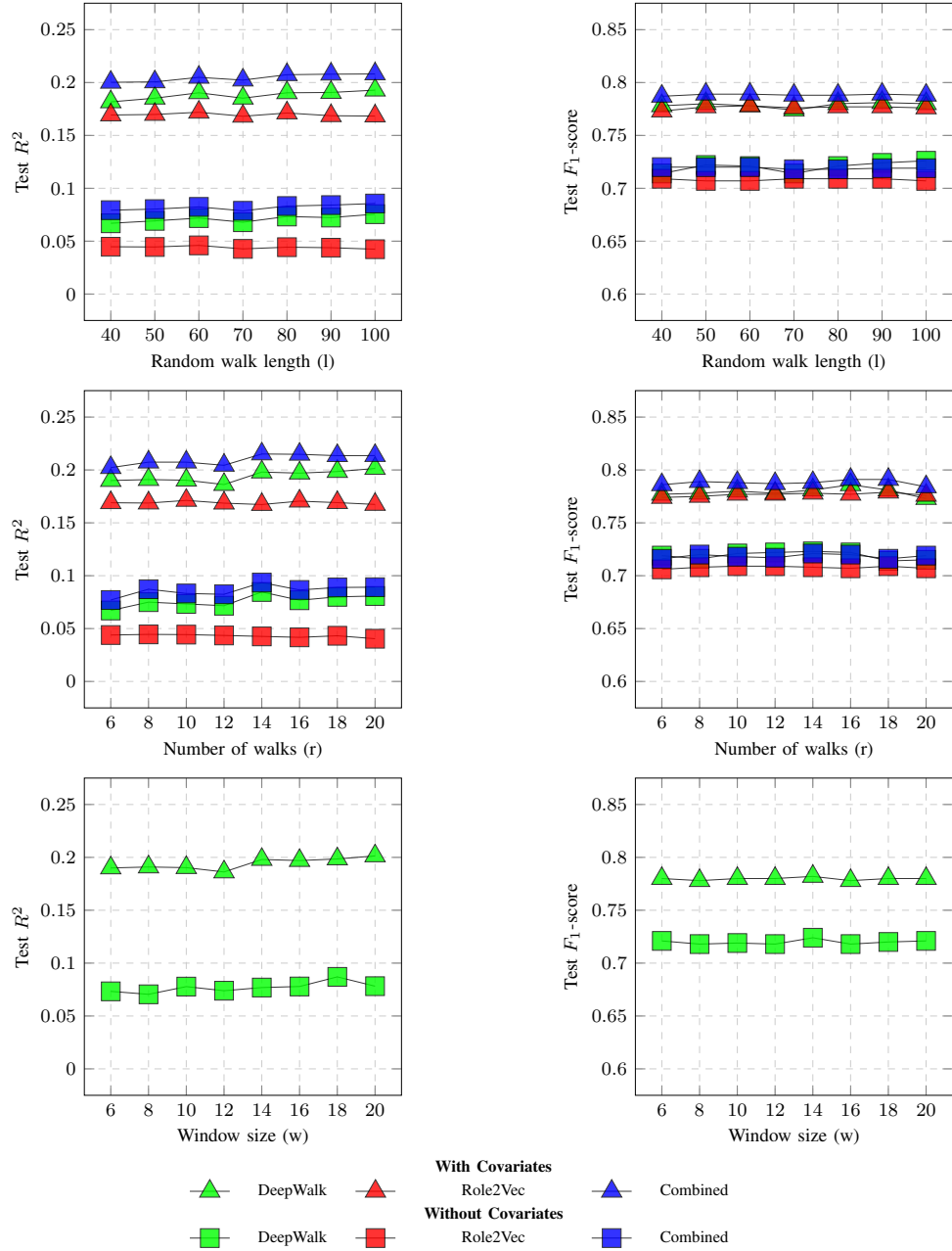


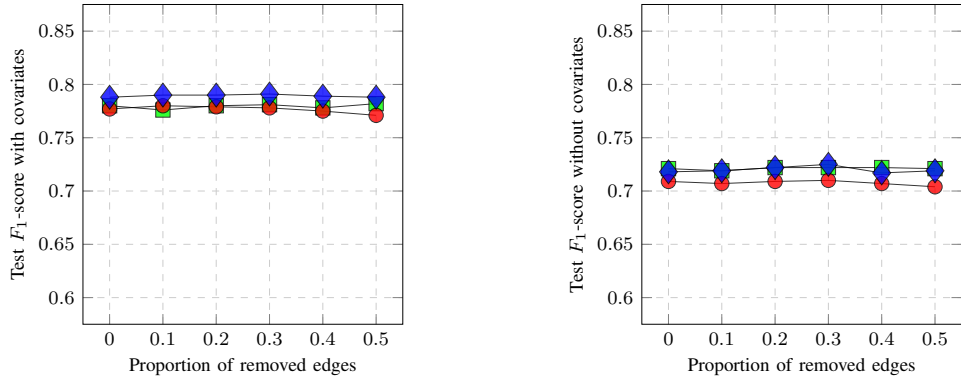
Fig. 2: Parameter sensitivity of grade and dropout prediction tasks – tested on the school of IT.

VII. LIMITATIONS

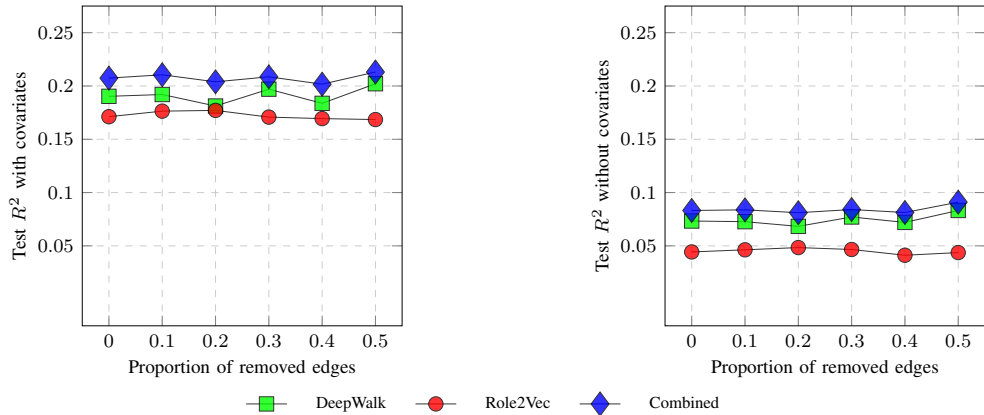
The graph embedding approach we propose permits the creation of student representations in an online manner (for instance, an entire new cohort or just a single visiting student). However, as the number of new students added to the trained representation grows, the representation quality obtained for these updates decreases. This is because the addition of new

nodes could have a considerable impact on the overall network structure. Accordingly, the optimal decision is to refit the entire embedding; a computationally demanding task. It is then necessary to refit the downstream models that depend on these embedded features. These limitations can be addressed by modern inductive graph representation learning techniques [72].

In our grade prediction, we sought to control for the



(a) Dropout classification task.



(b) Grade prediction task.

Fig. 3: Performance sensitivity using the school of IT network under perturbed data.

influence of a number of covariates, such as the number of courses that students took, or how many courses they withdrew from. However, there were a number of possible confounding variables that we either did not control for (such as cohort or programme effects), or could not control for (such as teacher, or grade curving effects). Future work should seek to capture and control for these effects, as they may provide our conclusions with substantial nuance.

Furthermore, in our grade and dropout prediction tasks, our choice of models assumed that students' embedding position contributed to students' outcomes in a linear fashion. This imposed a strong bias on our embeddings and future work should investigate the extent to which non-linear structure may exist within these embeddings, and whether more advanced, non-linear models can exploit it to improve classification or predictive performance.

In addition, our analysis did not investigate multi-resolution descriptions of the neighbourhood structure. Multi-scale node embedding procedures [73], [74] are able to characterise the location of nodes in the embedding space such that different proximities are described by certain components of the representation vector. It is likely that calculating such multi-scale representations would improve the effectiveness of any

analytic framework that builds on the features extracted from co-enrolment networks.

In evaluating Tinto's [12] model, we only investigated a specific operationalisation of social integration, namely, that existing within the classroom. Other forms of social integration, no less critical to students' dropout decisions, were omitted due to a lack of data. These include, but are not limited to bullying, relationships, and cheating.

Student co-enrolment networks can provide rich node-level attributes, the predictive utility of which has been well demonstrated in the literature [34], [40], [54]. However, we did not show how effective attributed node embedding procedures can be regarding feature extraction. These methods take into account the generic node features when a network embedding is created [75], [76], [77], [78], and is a promising future extension of our work to investigate the predictive power of such embedded features. This is particularly relevant for further investigation of Tinto's [12] model which, in addition to the predictions investigated in this study, emphasises the importance of contextual information such as academic background in forming students' initial goal commitments. Accounting for such student-level attributes might not only increase the predictive power – and thus remedial potential –

of our method, but also allow for a more nuanced investigation of Tinto's model.

VIII. CONCLUSION

This study has made several contributions towards understanding how students' social networks are associated with not only their academic performance, but also their persistence in their degree programme. In doing so, our analysis has been framed by Tinto's Student Integration Model [10], [11], [12], which presents a conceptual schema for understanding how students' commitment to their academic studies is modified by their social integration and academic performance. Our investigation not only finds substantial empirical evidence for Tinto's key claims, but provides an important contribution to the limited literature predicting student dropout using only information drawn from co-enrolment ties. Finally, we also present a novel use case for unsupervised graph embedding techniques to capture salient properties of student co-enrolment networks at an unprecedented scale. We find that not only are these techniques robust to changes in their parameterization, but also up to 10% of the variance in students' GPA can be explained simply by the other students with whom they enrol. This is a significant result and a considerable improvement over the existing literature, without requiring any node-level attributes [34], [40], [54]. The inclusion of a simple set of attributes, however, substantially improves this performance to 24%; approximately three times the variance explained by studies with datasets of a comparable scale [34].

In addition to the robustness of our methods to changes in their parameterization, which permits cheap computation at little to no cost to model performance, our perturbation study demonstrated that our method was robust to corruption within the dataset. Given this result, an interesting avenue of future research is to investigate the predictive performance of co-enrolment networks over the course of students' university studies. Of particular interest, would be assessing the scope for interventions in group design and the extent to which this is mediated by assortativity [60]. Furthermore, Tinto's model emphasises the role of contextual factors and describes a fundamentally temporal process [12]. These are key features of the model and future work should focus on incorporating them to further understand student persistence and inform remedial interventions.

REFERENCES

- [1] A. Calderon, "Massification of higher education revisited," http://cdn02.pucp.edu.pe/academico/2018/08/23165810/na_mass_revis_230818.pdf, RMIT University, Tech. Rep., 2018, [Online]; accessed 1-August-2019].
- [2] B. Carolan, *Social Network Analysis and Education: Theory, Methods and Applications*. SAGE Publications, Inc., 2014. [Online]. Available: <http://methods.sagepub.com/book/social-network-analysis-and-education>
- [3] L. C. Freeman, "Centrality in social networks conceptual clarification," *Social Networks*, vol. 1, no. 3, pp. 73–86, 1979. [Online]. Available: [https://doi.org/10.1016/0378-8733\(78\)90021-7](https://doi.org/10.1016/0378-8733(78)90021-7)
- [4] S. Wasserman and P. Pattison, "Logit models and logistic regressions for social networks: I. An introduction to Markov graphs and p," *Psychometrika*, vol. 61, no. 3, pp. 401–425, 1996.
- [5] E. Fincham, D. Gašević, J. Jovanović, and A. Pardo, "From Study Tactics to Learning Strategies: An Analytical Method for Extracting Interpretable Representations," *IEEE Transactions on Learning Technologies*, vol. 12, no. 1, pp. 59–72, 2019. [Online]. Available: <https://ieeexplore.ieee.org/document/8331946>
- [6] J. Jovanović, D. Gašević, S. Dawson, A. Pardo, and N. Mirriahi, "Learning analytics to unveil learning strategies in a flipped classroom," *The Internet and Higher Education*, vol. 33, pp. 74–85, 2017.
- [7] V. Tinto, "Research and practice of student retention: What next?" *Journal of College Student Retention: Research, Theory & Practice*, vol. 8, no. 1, pp. 1–19, 2006. [Online]. Available: <https://doi.org/10.2190/4YNU-4TMB-22DJ-AN4W>
- [8] J. Flood, "Read all about it: Online learning facing 80% attrition rates," *The Turkish Online Journal of Distance Education*, vol. 3, 2002. [Online]. Available: <http://tojde.adolu.edu.tr/yonetim/icerik/makaleler/48-published.pdf>
- [9] R. D. Johnson, S. Hornik, and E. Salas, "An empirical examination of factors contributing to the creation of successful e-learning environments," *International Journal of Human-Computer Studies*, vol. 66, no. 5, pp. 356–369, May 2008. [Online]. Available: <http://dx.doi.org/10.1016/j.ijhcs.2007.11.003>
- [10] V. Tinto, "Dropout from higher education: A theoretical synthesis of recent research," *Review of Educational Research*, vol. 45, no. 1, pp. 89–125, 1975. [Online]. Available: <http://www.jstor.org/stable/1170024>
- [11] —, *Leaving College: Rethinking the Causes and Cures of Student Attrition (1st ed.)*. University of Chicago Press, 1987.
- [12] —, *Leaving College: Rethinking the Causes and Cures of Student Attrition (2nd ed.)*. University of Chicago Press, 1993.
- [13] B. H. Munro, "Dropouts from higher education: Path analysis of a national sample," *American Educational Research Journal*, vol. 18, no. 2, pp. 133–141, 1981. [Online]. Available: <https://doi.org/10.3102/00028312018002133>
- [14] E. T. Pascarella and P. T. Terenzini, "Predicting freshman persistence and voluntary dropout decisions from a theoretical model," *The Journal of Higher Education*, vol. 51, no. 1, pp. 60–75, 1980. [Online]. Available: <http://www.jstor.org/stable/1981125>
- [15] A. M. Amaury Nora, L. C. Attinasi, "Testing qualitative indicators of precollege factors in tinto's attrition model: A community college student population," *The Review of Higher Education*, vol. 13, no. 3, pp. 337–355, 1990. [Online]. Available: <https://doi.org/10.1353/rhe.1990.0021>
- [16] A. Nora and L. I. Rendon, "Determinants of predisposition to transfer among community college students: A structural model," *Research in Higher Education*, vol. 31, no. 3, pp. 235–255, 1990. [Online]. Available: <http://www.jstor.org/stable/40195941>
- [17] P. T. Terenzini and E. T. Pascarella, "Voluntary freshman attrition and patterns of social and academic integration in a university: A test of a conceptual model," *Research in Higher Education*, vol. 6, no. 1, pp. 25–43, 1977. [Online]. Available: <https://doi.org/10.1007/BF00992014>
- [18] J. E. Eckles and E. G. Stradley, "A social network analysis of student retention using archival data," *Social Psychology of Education*, vol. 15, no. 2, pp. 165–180, Jun 2012. [Online]. Available: <https://doi.org/10.1007/s11218-011-9173-z>
- [19] K. L. Cela, M. Á. Sicilia, and S. Sánchez, "Social network analysis in e-learning environments: A preliminary systematic review," *Educational Psychology Review*, vol. 27, no. 1, pp. 219–246, 2015. [Online]. Available: <https://doi.org/10.1007/s10648-014-9276-0>
- [20] S. Dawson, D. Gašević, G. Siemens, and S. Joksimović, "Current state and future trends: A citation network analysis of the learning analytics field," in *Proceedings of the Fourth International Conference on Learning Analytics & Knowledge*, ser. LAK '14. ACM, 04 2014, pp. 231–240.
- [21] K. Stepanyan, K. Borau, and C. Ullrich, "A social network analysis perspective on student interaction within the twitter microblogging environment," *10th IEEE International Conference on Advanced Learning Technologies*, pp. 70–72, 2010.
- [22] L. M. Vaquero and M. Cebrián, "The rich club phenomenon in the classroom," in *Scientific reports*, 2013, p. 1174. [Online]. Available: <https://doi.org/10.1038/srep01174>
- [23] S. Joksimović, A. Manataki, D. Gašević, S. Dawson, V. Kovanović, and I. Friss De Kereki, "Translating network position into performance: Importance of centrality in different network configurations," in *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge (LAK 2016)*. ACM Press, 4 2016, pp. 314–323. [Online]. Available: <http://doi.acm.org/10.1145/2883851.2883928>
- [24] S. Goodreau, J. Kitts, and M. Morris, "Birds of a Feather, or Friend of a Friend? Using Exponential Random Graph Models to Investigate

- Adolescent Social Networks,” *Demography*, vol. 46, no. 1, pp. 103–125, 2009.
- [25] N. Dowell, O. Skrypnik, S. Joksimović, A. C. Graesser, S. Dawson, D. Gašević, P. de Vries, T. Hennis, and V. Kovanović, “Modeling learners’ social centrality and performance through language and discourse,” 2015, presented at the Proceedings of the 8th International Conference on Educational Data Mining. [Online]. Available: <https://eric.ed.gov/?id=ED560532>
- [26] P. Cui, X. Wang, J. Pei, and W. Zhu, “A survey on network embedding,” *IEEE Transactions on Knowledge and Data Engineering*, 11 2017. [Online]. Available: <https://dx.doi.org/10.1109/TKDE.2018.2849727>
- [27] N. K. Ahmed, R. Rossi, J. B. Lee, X. Kong, T. L. Willke, R. Zhou, and H. Eldardiry, “Learning role-based graph embeddings,” *arXiv preprint arXiv:1802.02896*, 2018. [Online]. Available: <https://arxiv.org/pdf/1802.02896.pdf>
- [28] B. Perozzi, R. Al-Rfou, and S. Skiena, “Deepwalk: Online learning of social representations,” in *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ser. KDD ’14. New York, NY, USA: ACM, 2014, pp. 701–710. [Online]. Available: <http://doi.acm.org/10.1145/2623330.2623732>
- [29] K. Henderson, B. Gallagher, L. Li, L. Akoglu, T. Eliassi-Rad, H. Tong, and C. Faloutsos, “It’s who you know: Graph mining using recursive structural features,” in *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ser. KDD ’11. ACM, 2011, pp. 663–671. [Online]. Available: <http://doi.acm.org/10.1145/2020408.2020512>
- [30] K. Henderson, B. Gallagher, T. Eliassi-Rad, H. Tong, S. Basu, L. Akoglu, D. Koutra, C. Faloutsos, and L. Li, “Rolx: structural role extraction & mining in large graphs,” in *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 2012, pp. 1231–1239. [Online]. Available: <http://doi.acm.org/10.1145/2339530.2339723>
- [31] L. F. Ribeiro, P. H. Saverese, and D. R. Figueiredo, “struc2vec: Learning node representations from structural identity,” in *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 2017, pp. 385–394. [Online]. Available: <https://dx.doi.org/10.1145/3097983.3098061>
- [32] C. Donnat, M. Zitnik, D. Hallac, and J. Leskovec, “Learning structural node embeddings via diffusion wavelets,” in *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 2018, pp. 1320–1329. [Online]. Available: <http://doi.acm.org/10.1145/3219819.3220025>
- [33] D. Yang, T. Sinha, D. Adamson, and C. P. Rose, “Turn on, tune in, drop out: Anticipating student dropouts in massive open online courses,” in *NIPS Data-Driven Education Workshop*, 2013, pp. 1–14.
- [34] J. Gardner, C. Brooks, and W. Li, “Learn from your (markov) neighbor: Coenrollment, assortativity, and grade prediction in undergraduate courses,” *The Journal of Learning Analytics*, vol. 5, no. 3, pp. 42–59, 2018. [Online]. Available: <http://dx.doi.org/10.18608/jla.2018.53.4>
- [35] M. Granovetter, “The impact of social structure on economic outcomes,” *Journal of Economic Perspectives*, vol. 19, no. 1, pp. 33–50, March 2005. [Online]. Available: <http://www.aeaweb.org/articles?id=10.1257/0895330053147958>
- [36] H. L. Horowitz, *Campus Life: Undergraduate Cultures from the End of the Eighteenth Century to the Present*. University of Chicago Press, 1987.
- [37] M. Moffatt and L. David, *Coming of Age in New Jersey: College and American Culture*. Rutgers University Press, 1989.
- [38] D. Kember, “A longitudinal-process model of drop-out from distance education,” *The Journal of Higher Education*, vol. 60, no. 3, pp. 278–301, 1989. [Online]. Available: <http://www.jstor.org/stable/1982251>
- [39] —, “An illustration, with case studies, of a linear-process model of drop-out from distance education,” *Distance Education*, vol. 10, no. 2, pp. 196–211, 1989. [Online]. Available: <https://doi.org/10.1080/0158791890100205>
- [40] D. Gašević, A. Zouaq, and R. Janzen, ““choose your classmates, your gpa is at stake!”: The association of cross-class social ties and academic performance,” *American Behavioral Scientist*, vol. 57, no. 10, pp. 1460–1479, 2013. [Online]. Available: <https://doi.org/10.1177/0002764213479362>
- [41] H. Cho, G. Gay, B. Davidson, and A. Ingraffea, “Social networks, communication styles, and learning performance in a cscl community,” *Computers and Education*, vol. 49, no. 2, pp. 309–329, Sep. 2007. [Online]. Available: <http://dx.doi.org/10.1016/j.compedu.2005.07.003>
- [42] E. Fincham, D. Gašević, and A. Pardo, “From social ties to network processes: Do tie definitions matter?” *Journal of Learning Analytics*, vol. 5, no. 2, pp. 9–28, 2018. [Online]. Available: <http://dx.doi.org/10.18608/jla.2018.52.2>
- [43] U. Brandes, “A faster algorithm for betweenness centrality,” *The Journal of Mathematical Sociology*, vol. 25, no. 2, pp. 163–177, 2001. [Online]. Available: <https://doi.org/10.1080/0022250X.2001.9990249>
- [44] A. Grover and J. Leskovec, “node2vec: Scalable feature learning for networks,” in *Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2016, pp. 855–864.
- [45] J. Tang, M. Qu, M. Wang, M. Zhang, J. Yan, and Q. Mei, “Line: Large-scale information network embedding,” in *Proceedings of the 24th International Conference on World Wide Web*. International World Wide Web Conferences Steering Committee, 2015, pp. 1067–1077. [Online]. Available: <https://dx.doi.org/10.1145/2736277.2741093>
- [46] W. L. Hamilton, R. Ying, and J. Leskovec, “Representation learning on graphs: Methods and applications,” *IEEE Data Engineering Bulletin*, vol. 40, pp. 52–74, 2017. [Online]. Available: <https://arxiv.org/pdf/1709.05584.pdf>
- [47] P. Goyal and E. Ferrara, “Graph embedding techniques, applications, and performance: A survey,” *Knowledge-Based Systems*, vol. 151, pp. 78–94, 2018. [Online]. Available: <https://doi.org/10.1016/j.knsys.2018.03.022>
- [48] M. Belkin and P. Niyogi, “Laplacian eigenmaps and spectral techniques for embedding and clustering,” in *Proceedings of the 14th International Conference on Neural Information Processing Systems: Natural and Synthetic*, ser. NIPS’01. Cambridge, MA, USA: MIT Press, 2001, pp. 585–591. [Online]. Available: <http://dl.acm.org/citation.cfm?id=2980539.2980616>
- [49] S. T. Roweis and L. K. Saul, “Nonlinear dimensionality reduction by locally linear embedding,” *Science*, vol. 290, no. 5500, pp. 2323–2326, 2000. [Online]. Available: <http://science.sciencemag.org/content/290/5500/2323>
- [50] A. Ahmed, N. Shervashidze, S. Narayanamurthy, V. Josifovski, and A. J. Smola, “Distributed large-scale natural graph factorization,” in *Proceedings of the 22nd International Conference on World Wide Web*, ser. WWW ’13. New York, NY, USA: ACM, 2013, pp. 37–48. [Online]. Available: <http://doi.acm.org/10.1145/2488388.2488393>
- [51] A. Wimmer and K. Lewis, “Beyond and below racial homophily: Erg models of a friendship network documented on facebook,” *American Journal of Sociology*, vol. 116, no. 2, pp. 583–642, 2010. [Online]. Available: <https://doi.org/10.1086/653658>
- [52] C. Lee, T. Schermerell, and M. J. Barber, “Investigating an online social network using spatial interaction models,” *Social Networks*, vol. 33, no. 2, pp. 129 – 133, 2011. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0378873310000584>
- [53] A. Traud, E. Kelsic, P. Mucha, and M. Porter, “Comparing community structure to characteristics in online collegiate social networks,” *SIAM Review*, vol. 53, no. 3, pp. 526–543, 2011. [Online]. Available: <https://doi.org/10.1137/080734315>
- [54] J. Gardner and C. Brooks, “Coenrollment networks and their relationship to grades in undergraduate education,” in *Proceedings of the 8th International Conference on Learning Analytics and Knowledge*, ser. LAK ’18. New York, NY, USA: ACM, 2018, pp. 295–304. [Online]. Available: <http://doi.acm.org/10.1145/3170358.3170373>
- [55] M. S. Granovetter, “The strength of weak ties,” *The American Journal of Sociology*, pp. 1360–1380, 1973.
- [56] J. Hommes, O. Arah, W. de Grave, S. A. Schuwirth, L., and G. Bos, “Medical students perceive better group learning processes when large classes are made to seem small,” *PLOS ONE*, vol. 9, no. 4, 4 2014. [Online]. Available: <https://doi.org/10.1371/journal.pone.0093328>
- [57] J. Hommes, B. Rienties, W. de Grave, G. Bos, L. Schuwirth, and A. Scherpbier, “Visualising the invisible: a network approach to reveal the informal social side of student learning,” *Advances in Health Sciences Education*, vol. 17, no. 5, pp. 743–757, Dec 2012. [Online]. Available: <https://doi.org/10.1007/s10459-012-9349-0>
- [58] B. Rienties, T. Lewis, R. McFarlane, Q. Nguyen, and L. Toeteneel, “Analytics in online and offline language learning environments: the role of learning design to understand student online engagement,” *Computer Assisted Language Learning*, vol. 31, no. 3, pp. 273–293, 2018. [Online]. Available: <https://doi.org/10.1080/09588221.2017.1401548>
- [59] T. T. Baldwin, M. D. Bedell, and J. L. Johnson, “The social fabric of a team-based m.b.a. program: Network effects on student satisfaction and performance,” *The Academy of Management Journal*, vol. 40, no. 6, pp. 1369–1397, 1997. [Online]. Available: <http://www.jstor.org/stable/257037>
- [60] D. S. Lyle, “Estimating and interpreting peer and role model effects from randomly assigned social groups at west point,” *The Review of Economics and Statistics*, vol. 89, no. 2, pp. 289–299, 2007.

- [Online]. Available: <https://www.mitpressjournals.org/doi/pdf/10.1162/rest.89.2.289>
- [61] S. Biancani and D. McFarland, "Social networks research in higher education," *Educational Studies*, no. 4, pp. 85–126, 2013. [Online]. Available: <https://EconPapers.repec.org/RePEc:nos:voprob:2013:i:4:p:85-126>
- [62] B. Weisfeiler and A. Lehman, "A reduction of a graph to a canonical form and an algebra arising during this reduction," *Nauchno-Tekhnicheskaya Informatsia*, vol. 2, no. 9, pp. 12–16, 1968.
- [63] G. Kelly, "Lecture attendance rates at university and related factors," *Journal of Further and Higher Education*, vol. 36, pp. 17–40, 02 2012. [Online]. Available: <https://dx.doi.org/10.1080/0309877X.2011.596196>
- [64] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient estimation of word representations in vector space," *CoRR*, vol. abs/1301.3781, 2013. [Online]. Available: <http://arxiv.org/abs/1301.3781>
- [65] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean, "Distributed representations of words and phrases and their compositionality," in *Advances in Neural Information Processing Systems 26*, C. J. C. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K. Q. Weinberger, Eds., 2013, pp. 3111–3119. [Online]. Available: <http://dl.acm.org/citation.cfm?id=2999792.2999959>
- [66] Q. Le and T. Mikolov, "Distributed representations of sentences and documents," in *Proceedings of the 31st International Conference on International Conference on Machine Learning - Volume 32*, ser. ICML'14, 2014, pp. 1188–1196. [Online]. Available: <http://dl.acm.org/citation.cfm?id=3044805.3045025>
- [67] F. Wilcoxon, "Individual comparisons by ranking methods," *Biometrics Bulletin*, vol. 1, no. 6, pp. 80–83, 1945. [Online]. Available: <http://www.jstor.org/stable/3001968>
- [68] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 58, no. 1, pp. 267–288, 1996. [Online]. Available: <http://www.jstor.org/stable/2346178>
- [69] H. W. Marsh, K.-T. Hau, and Z. Wen, "In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing hu and bentler's (1999) findings," *Structural Equation Modeling: A Multidisciplinary Journal*, vol. 11, no. 3, pp. 320–341, 2004. [Online]. Available: https://doi.org/10.1207/s15328007sem1103_2
- [70] J. Leskovec and A. Krevl, "SNAP Datasets: Stanford large network dataset collection," <http://snap.stanford.edu/data>, jun 2014.
- [71] J. Qiu, Y. Dong, H. Ma, J. Li, K. Wang, and J. Tang, "Network embedding as matrix factorization: Unifying deepwalk, line, pte, and node2vec," in *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining*, ser. WSDM '18. ACM, 2018, pp. 459–467. [Online]. Available: <http://doi.acm.org/10.1145/3159652.3159706>
- [72] W. L. Hamilton, R. Ying, and J. Leskovec, "Inductive representation learning on large graphs," in *Proceedings of the 31st International Conference on Neural Information Processing Systems*, ser. NIPS'17, 2017, pp. 1025–1035.
- [73] S. Cao, W. Lu, and Q. Xu, "Grarep: Learning graph representations with global structural information," in *Proceedings of the 24th ACM International on Conference on Information and Knowledge Management*. ACM, 2015, pp. 891–900. [Online]. Available: <http://doi.acm.org/10.1145/2806416.2806512>
- [74] B. Perozzi, V. Kulkarni, H. Chen, and S. Skiena, "Don't walk, skip!: Online learning of multi-scale network embeddings," in *Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2017*, ser. ASONAM '17. New York, NY, USA: ACM, 2017, pp. 258–265. [Online]. Available: <http://doi.acm.org/10.1145/3110025.3110086>
- [75] C. Yang, Z. Liu, D. Zhao, M. Sun, and E. Y. Chang, "Network representation learning with rich text information," in *Proceedings of the 24th International Conference on Artificial Intelligence*, ser. IJCAI'15. AAAI Press, 2015, pp. 2111–2117. [Online]. Available: <http://dl.acm.org/citation.cfm?id=2832415.2832542>
- [76] S. Pan, J. Wu, X. Zhu, C. Zhang, and Y. Wang, "Tri-party deep network representation," in *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence*, ser. IJCAI'16. AAAI Press, 2016, pp. 1895–1901. [Online]. Available: <http://dl.acm.org/citation.cfm?id=3060832.3060886>
- [77] X. Huang, J. Li, and X. Hu, "Accelerated attributed network embedding," in *Proceedings of the 2017 SIAM International Conference on Data Mining*. SIAM, 2017, pp. 633–641. [Online]. Available: <https://doi.org/10.1137/1.9781611974973.71>
- [78] L. Liao, X. He, H. Zhang, and T.-S. Chua, "Attributed social network embedding," *IEEE Transactions on Knowledge and Data Engineering*, 2018. [Online]. Available: <https://arxiv.org/pdf/1705.04969.pdf>
- [79] S. Cavallari, V. W. Zheng, H. Cai, K. C.-C. Chang, and E. Cambria, "Learning community embedding with community detection and node embedding on graphs," in *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*, ser. CIKM '17. New York, NY, USA: ACM, 2017, pp. 377–386. [Online]. Available: <http://doi.acm.org/10.1145/3132847.3132925>
- [80] M. U. Gutmann and A. Hyvärinen, "Noise-contrastive estimation of unnormalized statistical models, with applications to natural image statistics," *Journal of Machine Learning Research*, vol. 13, no. 1, pp. 307–361, 2012. [Online]. Available: <http://dl.acm.org/citation.cfm?id=2503308.2188396>



Ed Fincham is a PhD student in the School of Informatics at the University of Edinburgh. His research focuses on developing computational methods to derive interpretable, actionable insights from complex student data.



Benedek Rózemerczki is a PhD student in the School of Informatics at the University of Edinburgh. His research focuses on graph representation learning, large-scale network science, and network analytics.



Vitomir Kovanović is a Research Fellow at the School of Education, University of South Australia and a Data Scientist at the Teaching Innovation Unit, University of South Australia. His research focuses on examining students learning strategies and self-regulated learning, and assessing complex skills such as critical thinking and collaboration.



Srećko Joksimović is a Data Scientist at the Teaching Innovation Unit and a Research Fellow at the School of Education, University of South Australia. His research interests center around the analysis of teaching and learning in networked learning environments.



Jelena Jovanović is a Professor at the Department of Software Engineering, University of Belgrade, Serbia. Her research interests include semantic technologies, learning analytics, and technology enhanced learning. She is particularly interested in combining human and machine intelligence to shed insight on the learning process.



Dragan Gašević is Professor of Learning Analytics at the Faculty of Education at Monash University. His research involves developing computational methods that can shape next-generation learning technologies and advance our understanding of information seeking, sense-making, and self-regulated and collaborative learning.

APPENDIX A GRAPH EMBEDDINGS

A. Bipartite Projection

Since each student attends a course in a given semester, we treat attendance as a bipartite graph $G = (V, U, \tilde{E})$ where $V = \{v_1, \dots, v_m\}$ and $U = \{u_1, \dots, u_n\}$ are two distinct sets of nodes representing students and courses, and $\tilde{E} = \{\{v_i, u_j\} : v_i \in V, u_j \in U\}$ is the set of edges between students and the courses they take. From this bipartite graph, a weighted projection is calculated containing only V , such that all edges $E = \{\{v_i, v_j\} : v_i \in V, v_j \in V\}$ form a student-student graph where edges represent students who attended courses together, and weights represent the number of courses they shared. From hence we refer to this student-student graph as $G = (V, E)$.

Following this procedure we derive a co-enrolment graph. However, to investigate the predictive capacity of the co-enrolment networks using standard statistical techniques, it is first necessary to encode these relations within continuous vector space. This is achieved using two distinct graph embedding methods to capture neighbourhood-based and structural relations between students, respectively.

B. Neighbourhood Features

The first embedding procedure we employ, *DeepWalk* [28], is an unsupervised feature learning technique that uses information obtained from truncated random walks across a graph to learn a latent representation. Social relations are encoded within a continuous vector space such that the latent features characterize the neighbourhood similarity and community membership of the vertices. Accordingly, students from the same cohort who took a large number of courses together should be proximal in the embedding space. Beyond first-order connections, high-order proximity is also preserved and students with high neighbourhood overlaps should also be close together (see Fig. 1a).

DeepWalk arose as a generalization of language models, such as *Word2Vec* [64], [65], which serve to estimate the likelihood of a given sequence of words appearing within a corpus. In this case, a graph G is explored through r uniformly sampled random walks per vertex of length l , with the intention of estimating the likelihood of observing vertex v_i , given the previous vertices visited thus far in the walk:

$$Pr(v_i | (v_1, \dots, v_{i-1})) \quad (1)$$

However, as the length of the random walk grows, computing this likelihood becomes infeasible. This may be addressed by reformulating the problem such that the ordering of the walk is irrelevant [64], [65]; rather than using only the previous vertices visited to predict the next vertex, we may use one vertex to predict all other vertices in a given random walk, regardless of their offset from the given vertex. Moreover, the aim of *DeepWalk* is not only to estimate the probability distribution of node co-occurrences, but also to learn a latent representation [28]. Accordingly, the model includes a mapping function $\Phi : v \in V \mapsto \mathbb{R}^{|V| \times d}$ that characterizes the latent representation for each vertex v in d dimensions. This

problem – maximizing the probability of any vertex appearing in a random walk given a specific vertex within that walk – yields the following optimization problem:

$$\underset{\Phi}{\text{minimize}} \quad -\log Pr(\{v_{i-w}, \dots, v_{i+w}\} \setminus v_i | \Phi(v_i)) \quad (2)$$

In an iterative process, *DeepWalk* generates a random walk and, using the *SkipGram* algorithm [64], updates the mapping function Φ in accordance with the objective function in Equation (2) [28]. More specifically, for a given random walk, *SkipGram* maximizes the co-occurrence probability of vertices appearing within a window, w and approximates the conditional probability in Equation (2) by making the following independence assumption:

$$Pr(\{v_{i-w}, \dots, v_{i+w}\} \setminus v_i | \Phi(v_i)) = \prod_{\substack{j=i-w \\ j \neq i}}^{i+w} Pr(v_j | \Phi(v_i)) \quad (3)$$

The incorporation of Equation (3) into Equation (2) results in the optimization problem formulated by Equation (4).

$$\underset{\Phi}{\text{minimize}} \quad -\log \left(\prod_{\substack{j=i-w \\ j \neq i}}^{i+w} Pr(v_j | \Phi(v_i)) \right) \quad (4)$$

Following this procedure, the mapping function Φ is updated to maximize the probability of v_i co-occurring with its context $\{v_{i-w}, v_{i+w}\}$, where the context is the local graph as characterized by a series of random walks originating from v_i . The probabilities are parametrized by inner products of the embedding vectors and transformed with the softmax function. Accordingly, the trained model generates representations that, for each vertex v , capture similarities in local graph structure between vertices, such that vertices with similar neighbourhoods will be situated in close proximity within the embedding space. In contrast to more traditional network centrality metrics, *DeepWalk* generates such representations in linear time [26].

As an implicit network factorization technique, the *DeepWalk* procedure leverages community structure and has been used for community detection [79]. In the context of co-enrolment networks, such communities may include students who study similar majors and, accordingly, take similar courses.

While *DeepWalk* was originally designed for unweighted graphs, it may be seen as a corner case of *Node2Vec* [44], which generalizes to weighted graphs. In the analysis of co-enrolment networks, the ability to account for weighted vertices is especially relevant, as students who took similar courses will share strongly weighted connections.

C. Structural Features

In contrast to neighbourhood based node embeddings, structural role embeddings create representations in a latent space such that the structural similarity of nodes is preserved within the embedding space [29], [30], [31], [32]. That is, nodes

which have a similar degree, centrality, and clustering coefficient value and distribution in their neighbourhood are located close to each other in the embedding space. In the case of co-enrolment networks, this would imply that students who bridge disparate neighbourhoods would be close to each other within the latent space. The same would hold for students located in the core of highly interconnected communities. Importantly, this method implies that proximity within the graph does not necessarily translate to proximity within the embedding (see Fig. 1b).

The structural node embedding procedure that we adopted, *Role2Vec* [27], is a generalization of the node and document embedding models *DeepWalk* [28] and *Paragraph2Vec* [66]. We assume that we have a multi-labelled graph, $G(V, E, F)$, where F is a set of sets and $F(v) \in F$ is the set of structural role description labels specific to $v \in V$. Our goal is to learn a structural role embedding $\Phi : v \in V \mapsto \mathbb{R}^{|V| \times d}$ and a feature embedding $\Omega : f \in \mathcal{F} \mapsto \mathbb{R}^{|\mathcal{F}| \times d}$ jointly, where $\mathcal{F} = \bigcup_{v \in V} F(v)$. This learning task is formulated as an optimization problem in Equation (5) – we minimize the negative log-likelihood of observing the feature sets of vertices appearing in a random walk within a window w of node v_i , given the node representation $\Phi(v_i)$.

$$\underset{\Phi}{\text{minimize}} \quad -\log Pr(\{F(v_{i-w}), \dots, F(v_{i+w})\} \setminus F(v_i) | \Phi(v_i)) \quad (5)$$

In order to factorize the probability of observing the structural feature sets $\{F(v_{i-w}), \dots, F(v_{i+w})\}$ in the neighbourhood of v_i , conditioned on the node representations, we make the following independence assumption:

$$\underset{\Phi}{\text{minimize}} \quad -\log \left(\prod_{\substack{j=i-w \\ j \neq i}}^{i+w} Pr(F(v_j) | \Phi(v_i)) \right) \quad (6)$$

In order to further factorize Equation (6), we make an additional independence assumption regarding the features appearing within the feature sets:

$$\underset{\Phi, \Omega}{\text{minimize}} \quad -\log \left(\prod_{\substack{j=i-w \\ j \neq i}}^{i+w} \prod_{f \in F(v_j)} Pr(f | \Phi(v_i), \Omega(f)) \right) \quad (7)$$

The feature observation probabilities are parametrized by softmax transformed inner products of node and feature representations. Finally, in order to make the optimization problem tractable, Equation (7) is approximated by negative sampling [64], [80].

Since a structural role labelling mechanism is required for *Role2Vec* to generate $F(v)$, $\forall v \in V$ we opted for the *Weisfeiler-Lehman* procedure [62]. Given an initial node labelling, this procedure iteratively relabels the vertices. Concretely: for each node, labels of the node's neighbours are taken and sorted, and the sorted label sets are then used as the new labels. Nodes that have the same labelled neighbourhoods receive the same new label. Note that if the labelling is based

upon some structural property, this procedure can capture the structural similarity of neighbourhoods. In order to capture this neighbourhood structure, we log-bin the degree of each vertex and treated the bin membership as an initial label for each node within the graph.

To create this structural embedding, we first perform *Weisfeiler-Lehman* relabelling k times to generate node labels. In this way, each node has $k + 1$ labels which describe the structural role of the node itself, and the structure of vertex

neighbourhoods up to k hops distant. After this labelling procedure, we apply the *Role2Vec* embedding method, which performs r attributed truncated random walks from every source node with length l . From these attributed random walks we extract features within a window of w and learn a d dimensional structural role embedding according to the model specified in Equation (7). This embedding is generated in linear time, with respect to the number of nodes and edges.

3.3 Summary

In this chapter, we presented a partial, empirical investigation of Tinto's (1975) model. Specifically, we evaluated the theoretical claim that students' *social* interactions are associated with not only their academic performance, but also their dropout decisions.

In doing so, the present chapter makes a number of contributions to the learning analytics literature. For example, while a substantial body of research has investigated the association between various types of social ties and students' academic outcomes (Joksimović et al., 2016; Schreurs, Teplovs, Ferguson, de Laat, & Buckingham Shum, 2013; Skrypnyk, Joksimović, Kovanović, Gašević, & Dawson, 2015), the association between these ties and students' persistence in their studies has received markedly less attention (Eckles & Stradley, 2012). Furthermore, in assessing dropout, we conduct an analysis of students' social ties at a scale unprecedented in the wider learning analytics literature. In doing so, we present a novel methodological approach utilising graph-embedding techniques. We not only demonstrate that such methods can outperform more traditional SNA techniques, but also find considerable evidence for the validity of our approach. The robustness of our results to changes in the parameterisation of our models implies that good predictive performance may be attained in computationally cheap terms.

In conclusion, our investigation of Tinto's (1975) model found substantial evidence in favour of the theoretical claims made with regards to students' social domain. Namely, we find that students' social networks are not only predictive of their academic performance, but also facilitate the classification of students who have dropped out. In the context of learning analytics, the novelty of both our method and our results warrants future work not only to provide further validation, but also investigate the predictive performance of such embedded networks over the course of students' university studies.

4

Interpretable Representations of Student Behaviour

4.1 Introduction

HAVING investigated students' social interactions, we now turn our attention to students' interactions in the academic domain. One of the central challenges in understanding students' interactions in such a context is that, in pursuing their studies, students' possess diverse motivations. These motivations are in turn manifested by distinct behavioural patterns. To account for this diversity of behavioural activity, we turned to the educational research literature and identified "learning strategies" (Weinstein et al., 2012) as a viable construct for condensing this variety into interpretable types.

Following Weinstein et al. (2012), a learning strategy may be defined as any behaviours that facilitate the acquisition, understanding, or later transfer of new knowledge. In keeping with the theoretical literature, we postulated that these strategies were themselves composed of shorter-term "tactics", indicating cognitive routines for performing specified tasks (Alexander et al., 1998; Kirby, 1988). One of the practical benefits of this multi-level approach was that each level offers a transparent taxonomy of student behaviour.

Traditionally, research into learning strategies has relied upon self-reported data collected through think-aloud protocols and surveys (Bannert, Reimann, & Sonnenberg, 2014; P. Winne, 2013). Although this data provides invaluable information about students' perception of their own learning, they fail to measure how students actually employ study tactics and learning strategies (P. Winne & Jamieson-Noel, 2002). To address this, Hadwin et al. (2007) proposed that the analysis of student activity trace data (collected from the tools and services with which students interact during the learning process) is a vital resource for understanding the actual learning strategies that students adopt.

Furthermore, when comparing self-reports and trace data, the literature reveals that while these two instruments are designed to capture the same construct, their associations are not consistently observed. For instance, P. Winne and Jamieson-Noel (2002) demonstrated that learners are poor at calibrating their self-reported measures with actual measures of their behaviour, and often overestimate. This inconsistency is often attributed to poor learner self-reflection, and trace-based measures

of student achievement are often found to have much stronger associations with learning outcomes than self-reported measures (Zhou & Winne, 2012). Due to this finding, as well as the validity concerns that arise from the inconsistent associations observed between self-reports and trace data with regards to learning outcomes, our analysis focuses on the use of digital traces to measure student interactions.

4.1.1 Extracting Learning Strategies

Since Hadwin et al.'s (2007) proposal that future research should focus on developing more sophisticated statistical techniques for capturing students' learning strategies, a considerable body of literature has investigated the ability of educational data mining techniques to extract representations of learning strategies from digital trace data (Jeong, Biswas, Johnson, & Howard, 2010; Jovanović, Gašević, Dawson, Pardo, & Mirriahi, 2017; Kovanović, Gašević, Joksimović, Hatala, & Adesope, 2015; Matcha, Gašević, Uzir, Jovanović, & Pardo, 2019; Kovanović et al., 2019). Within such data, a learning strategy is a latent construct so, in order to extract interpretable representations, appropriate analytical methods are required. Our approach utilises hidden Markov models (HMMs), and extracts latent representations of students' behaviour which can be described and interpreted across two theoretically inspired levels: that of learning strategies, and the study tactics that compose them (Kirby, 1988). To achieve this, students' interactions with the academic resources were segmented into study sessions, where each session corresponds to a particular tactic. In fitting the HMM, we identify the optimal number of states, where each state is taken to represent a study tactic. Sequences of these states were then generated for each student, and were clustered to identify a set of learning strategies.

This methodology was developed with the intention of providing interpretable representations of student behaviour. Since the parameters of the HMM can be inspected, each state offers a intuitive description of the *type* of activity it captures. Furthermore, by clustering sequences of these study tactics into common sequence patterns, we can assess how students' behaviour changes over time. More concretely, in designing this approach, we intended to answer three research questions. Firstly, can a set of study tactics and, consequently, learning strategies be identified from students' study sessions? Secondly, is there an association between students' behaviour, as measured by these learning strategies, and their course performance? Finally, this analysis was conducted in the context of a feedback intervention which was incrementally implemented over the course of three years. Accordingly, our final research question was whether significant associations could be found between the type of feedback strategy and students' choice of learning strategy.

4.2 Publication: From Study Tactics to Learning Strategies: An Analytical Method for Extracting Interpretable Representations

The following section includes the verbatim copy of the following publication:

Fincham, E., Gašević, D., Jovanović, J., & Pardo, A. (2019). From study tactics to learning strategies: An analytical method for extracting interpretable representations. *IEEE Transactions on Learning Technologies*, 12(1), 59–72. doi:10.1109/TLT.2018.2823317

From Study Tactics to Learning Strategies: An Analytical Method for Extracting Interpretable Representations

Ed Fincham¹, Dragan Gašević, Jelena Jovanović², and Abelardo Pardo³

Abstract—Research into self-regulated learning has traditionally relied upon self-reported data. While there is a rich body of literature that has extracted invaluable information from such sources, it suffers from a number of shortcomings. For instance, it has been shown that surveys often provide insight into students’ perceptions about learning rather than how students actually employ study tactics and learning strategies. Accordingly, recent research has sought to assess students’ learning strategies and, by extension, their self-regulated learning via trace data collected from digital learning environments. A number of studies have amply demonstrated the ability of educational data mining and learning analytics methods to identify patterns indicative of learning strategies within trace log data. However, many of these methods are limited in their ability to describe and interpret differences between extracted latent representations at varying levels of granularity (for instance, in terms of the underlying data of student actions and behavior). To address this limitation, the present study proposes a new methodology whereby interpretable representations of student’s self-regulating behavior are derived at two theoretically inspired levels: that of learning strategies, and the study tactics that compose them.

Index Terms—Learning strategies, study tactics, self-regulated learning, learning analytics, flipped classroom.

I. INTRODUCTION

THE importance of creating active learning environments for learners in higher education has been consistently emphasised in contemporary educational research [25], [26]. The positive impact of such environments on learners’ experiences and academic outcomes has been widely documented. For instance, in a meta-analysis of active learning studies, Freeman et al. [15] found that students in active learning settings on average earned higher grades and were less likely to fail than peers in traditional lecture based models. While active learning has clear benefits, implementations are often fraught with

difficulty [27]. In many classrooms, students often assume a passive role rather than actively create knowledge [19]. Inverting the traditional model requires consideration on the part of instructors regarding activity structure, learner knowledge and motivation, and overall curriculum design.

The flipped classroom (FC), an active learning design, is a blended learning environment that retains some features of a traditional classroom model but augments it with activities that require active student interaction both before and during face-to-face classes [38]. While numerous studies have reported greater student satisfaction associated with FC implementation, there remains a paucity of research comparing academic outcomes in a traditional classroom model with a FC model. FC encourages autonomy on the part of students and requires them to organise and regulate their own learning [32]. However, students often lack the requisite skills to modify their learning strategies to suit novel learning situations [39].

A. Learning Strategies and Self-Regulated Learning

Prior studies in educational research provide an abundance of definitions of learning strategies. For the purposes of this study, we rely on the broad account provided by Weinstein et al. [12, p. 227] whereby a learning strategy represents any thoughts, behaviors, beliefs or emotions that facilitate the acquisition, understanding, or later transfer of new knowledge and skills. In the literature, the terms learning strategy and study tactic are often used interchangeably but they are in fact different [33]. While study tactics are existing cognitive routines for performing specified tasks, learning strategies are the means of selecting, combining, or redesigning these cognitive routines, directed by a learning goal [1], [33], [60], [63], [70].

There has been considerable research into learning strategies. For instance, Pintrich and de Groot [51] investigated the relationship between self-regulation, motivation, students’ choice of cognitive strategy, and their performance on academic tasks in the classroom. Their findings suggest that the relationship between a student’s cognitive strategy and academic outcomes is mediated by self-regulation. This is to say, not only must students be aware of possible learning strategies, they must also know how and when to apply different strategies [64]. Given these findings, we understand students’ choice of learning strategies as manifestations of self-regulated learning processes. Following Winne [65], we view self-regulated learning as a set of intentional actions and processes

Manuscript received August 29, 2017; revised March 15, 2018; accepted April 1, 2018. Date of publication April 5, 2018; date of current version April 4, 2019.

E. Fincham is with the School of Informatics, University of Edinburgh, Edinburgh EH8 9YL, U.K (e-mail: ed.fincham@ed.ac.uk).

D. Gašević is with the Faculty of Education, Monash University, Clayton, VIC 3800, Australia, and also with the School of Informatics, University of Edinburgh, Edinburgh EH8 9YL, U.K (e-mail: dgasevic@acm.org).

J. Jovanović is with the Faculty of Organizational Sciences, University of Belgrade, Belgrade 11000, Serbia (e-mail: jeljov@fon.bg.ac.rs).

A. Pardo is with the School of Electrical and Information Engineering, University of Sydney, Camperdown, NSW 2006, Australia (e-mail: abelardo.pardo@sydney.edu.au).

Digital Object Identifier 10.1109/TLT.2018.2823317

1939-1382 © 2018 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

that are planned and deployed for the purposes of learning new skills and acquiring new knowledge. This implies a degree of learner agency regarding the monitoring, evaluation, and subsequent adjustment of learning activity.

Research into self-regulated learning and learning strategies has traditionally relied upon self-reported data collected through think aloud protocols and surveys [3], [65]. While such data have provided invaluable information about students' perceptions of learning, they crucially fail to measure how students actually employ study tactics and learning strategies [67]. Accordingly, Hadwin et al. [20] have proposed that the analysis of student activity trace data (collected from the tools and services the students interact with during the learning process) is vital for furthering our understanding of self-regulated learning. Moreover, such data streams offer a number of benefits. For instance, traces track learning events without interfering with a student's thinking or navigation through content and so provide a direct and reliable indicator of cognition [61].

While the literature has shown that self-reported instruments capture information about relatively stable propensities to engage in self-regulated learning [20], there are several shortcomings. First, self-reports can be plagued by inaccuracy due to poor recall of students' prior behavior [67]. Second, think aloud protocols place additional cognitive load on students and provide no guarantee that pertinent, unbiased information regarding students' applied study tactics and learning strategies is expressed [65]. Third, even though an analysis of self-reported data can identify overarching trends in how students' learning strategies develop [65], they fail to capture the strategic adaptations and developments students make within and across study sessions [62].

Comparing self-reports and trace data, however, reveals that associations between the two instruments on the same construct are not consistently observed. For example, Winne and Jamieson-Noel [67] showed that learners are inaccurate in calibrating their self-reported measures with actual measures of their use of specific study tactics. The authors found that learners have a tendency to overestimate their use of specific study tactics. Zhou and Winne [68] attribute this inaccuracy to poor learner self-reflection and found that trace based measures of student achievement goal orientation had much stronger associations with learning outcomes than self-reported measures. The authors interpreted the self-reports as measuring student intentions while trace data measured their realised intentions [68].

To fully mine the potential trace data logs, Hadwin and Winne [20] argued that future research should focus on developing more sophisticated statistical techniques and methods for examining patterns across groups of students. Since then, numerous studies have amply demonstrated the ability of educational data mining and learning analytics methods to extract representations of learning strategies from trace log data over varying time frames [4], [6], [28], [30], [36], [39], [48]. The present study contributes to this body of literature and proposes a new methodology for extracting and interpreting patterns of student activity that characterise different self-regulating behaviors. Compared to existing methods, however,

our approach extracts latent representations of student's self-regulating behavior which can be described and interpreted across two theoretically inspired levels: that of learning strategies, and the study tactics that compose them [33].

The present study applied this novel methodology to examine students' learning strategies using three years worth of trace data from the preparatory activities students undertook prior to scheduled lectures in a first-year undergraduate FC course in computer engineering. The data was collected from the University's Learning Management System (LMS) and was comprised of 13 weeks worth of trace logs, results on a mid-term test, and results on a final exam. The three years comprise a baseline and two iterations of a feedback intervention where students were sent weekly feedback, customised on the type and extent of their engagement, and their performance.

B. Analytical Methods for Pattern Detection in Student Behavior

Existing research into the identification and analysis of learning strategies in online and blended learning environments makes frequent use of trace data [20], [65]. Within these data, a student's learning strategy is a latent construct so, in order to extract interpretable and meaningful representations, appropriate analytical methods are required. For instance, Jeong et al. [29] used Hidden Markov Models (HMMs) to identify patterns in the learning actions of middle school science students. The students were engaged in learning-by-teaching activities, where they taught a computer agent graphical representations called concept maps. The analysis of these trace data found clear associations between students' learning performance and the identified patterns of their interaction with the system. In a later study, Jeong et al. [28] demonstrated the generality of the HMM approach by applying it to examine the learning behavior of professionals in an asynchronous learning environment that, compared to the learning-by-teaching environment, provided a greater deal of learner control. Specifically, Jeong et al. [28] examined the differences in behaviors, represented by the HMM states, between high and low performing students. The authors found that higher performance was associated with linear, consistent patterns of behavior.

Clustering methods have also been used in conjunction with online and blended environment trace data to identify different patterns of learner behavior, representing different underlying learning strategies. For instance, Valle and Duffy [11] used clustering to identify different approaches to learning within a distance education environment and found that, in spite of widespread concern about student time management under such conditions, the majority of learners were very effective in their choice of learning strategies. Kovanović et al. [36] used a similar clustering approach to identifying student learning strategies in an online graduate software engineering course. Their study used the Communities of Inquiry model as a theoretical framework and sought to answer how learning strategies could be explained in terms of self-regulation, goal-orientation, and cognitive presence. The authors found a significant association between effective learning strategies and high levels of cognitive presence in social knowledge construction.

The two aforementioned analytical approaches may also be used in conjunction with other methods. Perera et al. [48] combined sequence mining with clustering to investigate how students work in small groups, and how group performance relates to the use of online collaborative tools. Their study found clear, early patterns indicating effective and poor practices that could be used to inform timely remediation. Berland et al. [4] also utilised sequential analysis with clustering to investigate mechanisms of how students begin to learn to program in a relaxed, collaborative learning setting. Specifically, when the students changed their code a snapshot was saved, from which the likelihood of each code state transitioning to every other code state was calculated. The study found that students typically progressed from exploration, through tinkering, to refinement.

Research into learning strategies has amply demonstrated the ability of sequential analytical methods to identify latent constructs within trace data. However, within this field there is considerable methodological breadth and the underlying unit of analysis can range from sequences of student actions within a study session [30] to week or multi-week “phases” [39]. However, there has been little exploration as to whether, and to what extent, learning strategies are manifested in sequences of data over narrower time frames. If learning strategies could be identified robustly from sequences of study conducted over a narrow time frame, this would facilitate research into how students adjust their learning strategies and would allow for more accurate assessment of how interventions impact on student behavior. To explore this, we examine whether patterns in student behavior can be detected at the level of study sessions, and whether sequences of such sessions can robustly identify learning strategies. This approach has the benefit of offering a close approximation of how students engage with a learning environment. Accordingly, we defined our first research question as follows:

Research Question 1: *Can we detect robust study tactics on the basis of study sessions? Given a sequence of detected study tactics over a specified time-period, can patterns indicative of learning strategies, be identified?*

C. Learning Strategies in a Flipped Classroom

It is widely acknowledged that effective regulation of learning strategies can lead to higher academic achievement [2], [34], [51], [69]. In online or blended learning environments, an important decision students face is how best to utilise the available learning resources. Prior work has found that many learners struggle to self-regulate effectively in online learning environments [58]. In a blended learning environment, Lust et al. [40] identified three student learning strategy profiles: 1) no-users who made no use of the available face-to-face tools and made very limited use of the LMS, 2) intensive users that frequently used the majority of the tools and resources available in the LMS, and 3) incoherent users who used only online tools and made little to no use of the available face-to-face tools. Regarding academic performance, both intensive and incoherent strategies had significantly higher academic performance than no-users. One of the reasons for this observed

difference between learning strategy profiles was later found to be students’ self-regulation of tool use [39]. The study showed that while the majority of students actively regulated their learning, only 3 percent of them regulated their study in accordance with the course objectives. Instead, a majority of the students (59 percent) used a very limited set of the available tools, suggesting an inability to regulate their learning activity effectively [49], [64].

In the case of FC, there are reasons to suspect that students would be similarly incapable of effectively regulating their learning activities. This may in part be due to FC not being a dominant pedagogic method, such that many learners find themselves unfamiliar with its features and requirements. Hattie et al. [22] and Hattie and Donoghue [24] noted that while study skills could be readily transferred between similar learning contexts, this did not hold for disparate learning environments. Given the substantial difference between a FC model and a traditional classroom, it is reasonable to expect students unfamiliar with FC to experience difficulties with strategy regulation.

Should students successfully adapt to a FC model, research suggests that academic achievement is comparable or better than under a traditional model [42]. Such claims, however, are far from conclusive: in a scoping review of the literature, O’Flaherty et al. [44] found conflicting evidence for the benefit of FC over a traditional classroom model. Given our proposed methodology, validating whether and how the regulation of pre-lecture activities relates to course performance leads us to our second research question:

Research Question 2: *Is there an association between the identified patterns in student behavior when preparing for face-to-face sessions and course performance? In other words, do differences in identified learning strategies between students relate to differences in academic outcomes?*

D. Learning Strategy and Feedback

Providing opportunities for the improvement of students’ regulation of learning strategies is critical in addressing the limitations of students’ study skills. Many studies have found that feedback can lead students to engage in more self-regulated learning activities, which is associated with significant differences in learning outcomes [23], [34], [43]. In terms of Winne and Hadwins model of self-regulated learning [66], students set goals for their learning, and monitor how their learning strategies are aiding or hindering their progress towards these goals. This monitoring provides an internal feedback loop that relies on both internal and external feedback to help students regulate their learning. Feedback affects learners evaluation of the products of their learning and the effectiveness of study tactics and strategies used. However, there is limited understanding of the association between feedback interventions, learning strategies, and academic outcomes in FC settings. Indeed, there is a dearth of research considering whether learning strategies change over several offerings of the same course following the implementation of an intervention. This is especially the case in the context of practical implementations of interventions within learning environments based on FC principles.

The contemporary literature posits that, within an FC setting, feedback can be personalised at scale by making use of data-rich learning environments [45], [46]. Given how recent these developments are, the existing literature has only studied the associations of such feedback with student satisfaction and academic performance. However, there has been limited research considering the associations of such feedback with the choices of study tactics and learning strategies. Therefore, understanding the impact of feedback in FC settings prompts our third research question:

Research Question 3: *What is the impact of the personalised feedback interventions on the choices of learning strategies? What, if any, associations are there between the personalized feedback interventions, students' choice of learning strategies, and academic outcomes?*

This last research question primarily aims to demonstrate a potential utility of the proposed method for detection of learning tactics and strategies from trace data recorded by digital learning environments.

II. METHODS

A. Study Context

In this study, a FC model was applied to a first-year engineering course in Computer systems at an Australian higher education institution. The course lasted 13 weeks and, over the three years for which data were collected, had an enrollment of approximately 1,300 students. Trace data were available for 1,138 of these students (290, 371, and 477 students in each year, respectively). The course also included a mid-term test and a final exam.

The FC design was composed of two elements: a set of online resources intended to be completed in preparation for the plenary session (the lecture), and the re-framing of the plenary session to embrace an active learning design requiring students' preparation and participation in collaborative problem solving tasks [47]. Trace data were collected from the LMS and form the basis of this study. The resources for class preparation retained the same structure throughout the course, and included:

- Videos with multi-choice questions (MCQs): short videos introduced and explained relevant concepts. These were immediately followed by a set of MCQs promoting simple factual recall. The MCQs were framed as formative assessments. Students could answer a question, have their answer evaluated, and were then given the choice of either seeing the solution or trying again.
- Documents with embedded MCQs: students were required to study the provided reading materials and, similar to the video resources, were presented with MCQs embedded within the text. These MCQs were also framed as formative assessments.
- Problem sequences: the activity consists of answering a sequence of more complex problems about the topic covered in previous MCQs. Students sequentially attempt a set of exercises framed as summative MCQs. If an exercise is solved correctly, the student's score is

increased and the exercise is removed from the set. Alternatively, a new exercise is randomly selected and the current problem remains in the sequence. Students received exercises randomly until they solved all of them correctly. To be counted towards their final course grade, the exercises had to be solved before the plenary session. This requirement was introduced as an incentive for students to timely engage with the preparatory learning materials.

A more comprehensive description of the learning design, including examples of the tasks students faced is provided by Pardo and Mirriahi [47].

B. Feedback Design and Iterations

Across all three years of the course, students were provided with real-time feedback on their level of engagement with the preparation activities and their performance via an analytics dashboard [31]. This dashboard enabled students to monitor their engagement with the video resources, their performance on the video-related MCQs, their performance on the MCQs associated with the reading materials, as well as the percentage of correctly solved problem sequences. In addition to the student's personal scores, the dashboard displayed the overall class scores, thus facilitating social comparison. This information was updated every 15 minutes, and was reset each week in line with the changing course content.

The three years also provide a baseline and two iterations of an additional, personalised feedback intervention. The personalised feedback was generated as follows: for each activity in the course design, instructors prepared in advance a set of feedback messages for a variety of levels of interaction with the learning resources. For each week with n activities and m students, instructors would write $k \times n$ comments, where k is the number of engagement categories instructors are interested in. In this dataset, four categories were used (i.e., $k = 4$). At the end of each week, for each student and for each activity, an algorithm would select the appropriate feedback message based on the level of the student's participation in that activity, to generate a personalised email. By focusing on k categories rather than m students, the approach is scalable and does not depend on the number of students in the cohort.

In the first year (2014), no additional, personalised feedback was given beyond what the students could access via the analytics dashboard. In the second year (2015), personalised feedback was provided weekly to students for the first half of the course and, in the third year (2016), similar personalised feedback was weekly provided albeit for both halves of the course. Further details about the generation of personalised feedback is provided by Pardo and his colleagues [45], [46].

C. Learning Traces

The study is based on student interaction data in the form of LMS trace logs obtained from student engagement with the preparatory learning resources during weeks 2-13 of the three editions of the course. Each event is represented by a quadruple of event id, anonymised student id, type of learning action,

TABLE I
STUDY SESSION FEATURES

Action Code	Description
<u>FA_PERC</u>	Percentage of formative assessment actions within the study session
<u>FA_CO_PERC</u>	Percentage of correct formative assessment actions
<u>SA_PERC</u>	Percentage of summative assessment actions within the study session
<u>SA_CO_PERC</u>	Percentage of correct summative assessment actions
<u>VID_PERC</u>	Percentage of video play/pause actions within the study session
<u>READ_PERC</u>	Percentage of reading (content access) actions within the study session
<u>METACOG_PERC</u>	Percentage of meta-cognitive actions within the study session

and timestamp. These events were classified into formative assessment (FA) actions, summative assessment (SA) actions, video actions, content access (reading) actions, and meta-cognitive actions. The latter was defined as engagement with the dashboard and syllabus materials, where students could monitor their engagement with video resources, their success in answering video MCQs, reading MCQs, as well as the percentage of correctly solved problem sequences.

Study sessions were extracted from the events data and are here defined as continuous sequences of events where any two events are within 21 minutes of one another. As there is no standard for estimation of time on task based on trace data [35], this delimiter was chosen because, given the length of videos, the 96th percentile of 17.5 minutes seemed insufficiently short, whereas the 98th percentile of 41.2 minutes seemed overly long. The delimiter accounts for the length of videos while excluding excessive periods of inactivity. The segmentation resulted in 55,710 study sessions across 1,138 unique students for the 12 active weeks of the three course editions. To gain an insight into the general pattern of study sessions we removed outliers. Specifically, study sessions comprised of a single event were removed, along with students with excessive study session counts (some students had in excess of 100 sessions, compared to a median of 22). Additionally, study sessions were only retained if a student had at least one study session in both halves of the course. Removing these outliers resulted in 54,400 study sessions across 1,114 students (28,630 sessions in the first half of the course, 25,770 in the latter half). The data were aggregated across years to facilitate inter-year comparisons on the basis of the same units of analysis: learning strategies. These were extracted following the method described in the next subsection.

D. Feature Computation and Hidden Markov Models

Under this analytic framework, an HMM state corresponds to a study session and can be considered representative of the study tactic that a student applied during that session. To this end, the proportions of each learning action for each study session are used as features. These are first ordered by student and then chronologically ordered. Table I documents the features computed for each study session.

An HMM was fitted to the features of student study sessions using the `depmixS4` R package [59]. This method was selected as it enables us to identify some of the students' general behavioral patterns from sequences of their interactions, segmented into study sessions. After the parameters that define

the states of the HMM and the possible state transitions are initialised, the Expectation Maximization (EM) algorithm was applied iteratively until the model parameters converge [52]. The metrics we used to assess model fit were the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC). These take into account both the complexity of the derived model (that is, how many states it is composed of) and how likely it is, given the data (the log likelihood), to find a model that best strikes a balance between high likelihood and parsimony.

Previous research has found that student learning behaviour alters during examination weeks [50]. Furthermore, the literature has found that learning design influences students' choice of learning strategy [17]. In our dataset, the learning design changed in weeks 6 and 13 (mid-term and final exams, respectively); the FC design was suspended so as to allow students to focus on revision and offered no preparation activities. Accordingly, we excluded weeks 6 and 13 from our analysis. An HMM was fit to the remaining data with 8 states. As a result, each session had a list of probabilities representing assignment to each of the HMM states and were labeled with the most probable state assignment.

E. Clustering

Unlike the HMM analysis, which identified a number of study tactics common to both halves of the course, clustering was done on each half of the course separately. This decision was made because the feedback intervention was phased in incrementally over the two halves of the course and clustering them separately enabled us to assess the impact of the feedback intervention and answer our third research question.

After identifying study tactics with the states of an HMM, a sequence of such states was created for each student according to the chronological order of that student's sessions (note that each study session has a corresponding study tactic). These sequences were then clustered. Following the proposal made by Kovanović [36], we used agglomerative hierarchical clustering, based on Ward's method [21]. Computation of the distance between sequences of HMM states, required for the clustering algorithm, is based on the optimal matching distance metric [16]. According to this metric, the distance between two sequences of states is the minimal cost, in terms of insertions, deletions, and/or state substitutions required to transform one sequence into another. These computed distances were then normalised to account for differences in sequence lengths. To select the optimal number of clusters, dendrograms were used to identify the most plausible segmentations of the tree structure. This method grouped similar sequences in each half of the course to detect students' adopted learning strategies and so answer our first research question. Student cluster assignments enabled us to group students and so identify whether different learning strategies relate to differences in academic performance, and thus address our second research question.

In this analysis, sequences of HMM states correspond to sequences of study tactics while the clusters of these sequences are representative of the learning strategies students

TABLE II
TRANSITION MATRIX FOR EIGHT STATE HMM

	Tactic 1	Tactic 2	Tactic 3	Tactic 4	Tactic 5	Tactic 6	Tactic 7	Tactic 8
Tactic 1	0.183	0.101	0.105	0.192	0.048	0.186	0.039	0.146
Tactic 2	0.038	0.228	0.136	0.299	0.095	0.091	0.053	0.058
Tactic 3	0.045	0.150	0.135	0.315	0.060	0.154	0.047	0.093
Tactic 4	0.044	0.108	0.088	0.421	0.059	0.139	0.050	0.090
Tactic 5	0.030	0.127	0.101	0.229	0.214	0.135	0.075	0.090
Tactic 6	0.046	0.090	0.126	0.275	0.061	0.218	0.045	0.138
Tactic 7	0.059	0.108	0.090	0.306	0.096	0.154	0.075	0.112
Tactic 8	0.036	0.128	0.108	0.309	0.137	0.136	0.059	0.087

adopted. In both the first and latter half of the course, the sequence clustering algorithm identified four clusters.

F. MANOVA Analysis

To confirm that the learning strategies identified as part of our first research question were indeed distinct in terms of their tactic composition, a multivariate analysis of variance (MANOVA) was used. Cluster assignment was treated as the single, independent variable and the 8 HMM states were treated as the dependent variables. Before running a MANOVA, in keeping with the work of Lust et al. [39], [40], the homogeneity of variances assumption was tested using Levene’s test and homogeneity of covariances was checked using Box’s M test [14]. Following Bray and Maxwell [8], we opted for the Pillai-Bartlett test statistic as this is the most robust to violations of the test assumptions. In the case of a significant MANOVA result, a follow up univariate ANOVA is conducted on each of the dependent variables. This use of a univariate test following a multivariate test is often considered a “protection” from the Type I errors arising from the direct use of multiple ANOVAs [7]. To further control for Type I error rate inflation due to the multiple comparisons, the conservative Bonferroni correction was adopted. Before running follow-up ANOVAs, the homogeneity of variance was checked using Levene’s test and, where violated, a non-parametric Kruskal-Wallis test was used. This provided robustness to any violations of normality. Significant Kruskal-Wallis tests were followed up by pairwise comparisons of clusters (also using Bonferroni corrections). Finally, in the case of significant ANOVAs, Tukey’s honest significant difference (HSD) test was used to check for differences among cluster pairs.

G. Learning Strategies and Academic Outcomes

To examine whether there was a significant difference between the student groups regarding academic performance, and so answer our second research question, Kruskal-Wallis tests were conducted, followed by pairwise Mann Whitney U tests. These tests were separately applied on the basis of both the final exam and the mid-term exam and provide robustness to any violations of normality. The False Discovery Rate (FDR) correction was used for preventing alpha inflation when doing multiple tests [10].

H. Assessing the Impact of Interventions

To answer our third research question—assessing the impact of the interventions over the three years of the course—we introduced the year as a variable of interest. As the feedback was incrementally implemented, if it had an effect on students’ choice of learning strategy, we would expect the distribution of cluster assignments (that is, learning strategies) to be dependent on the year variable. To answer this research question, we conducted chi-squared tests on the cluster assignments for both halves of the course. To assess the impact of interventions on learning strategies on an inter-year basis, for both halves of the course we tabulated the distribution of students across learning strategies, stratified by year. As the interventions were phased in differently over the two halves of the course, we also considered intra-year learning strategy transitions. To assess these transitions, we tabulated learning strategy transition matrices for each year of the study.

III. RESULTS

A. Research Question 1: Identifying Learning Strategies

1) *Identifying Study Tactics:* The transition matrix for the HMM is shown in Table II. The eight identified study tactics are:

- Tactic 1: Video actions take up 60-100 percent of a session. Content access actions also feature, and account for up to 40 percent of a session.
- Tactic 2: SA actions account for 90-100 percent of a session. Of this, correct answers range from 30-60 percent.
- Tactic 3: FA actions account for 90-100 percent of a session. Of this, correct answers range from 20-60 or 90-100 percent. Content access actions comprise up to 10 percent of a session.
- Tactic 4: Content access actions account for 90-100 percent of a session.
- Tactic 5: Meta-cognitive actions account for 90-100 percent of a session.
- Tactic 6: FA actions account for 20-90 percent of a session. Of that, correct answers range from 30-70 or 90-100 percent. Content access actions account for up to 30 percent of a session and there is a small chance that meta-cognitive actions comprise up to 10 percent of a session.
- Tactic 7: Content access actions are most likely, accounting for 40-90 percent of a session. Meta-cognitive actions are also likely and comprise 10-40 percent of a session. There is a chance that video actions account for 50-60 percent of sessions.
- Tactic 8: SA actions are most likely, accounting for 20-70 percent of a session. Of this, correct answers range from 30-60 percent. FA actions are also likely, comprising up to 50 percent of a session and with correct answers ranging from 30-60 percent. Content access actions and video actions both account for up to 20 percent of a session, and there is a chance of that up to 10

TABLE III
SUMMARY STATISTICS FOR THE PROPORTION OF THE EIGHT TACTIC TYPES WITHIN CLUSTERS FOR THE FIRST HALF OF THE COURSE

	Median (Q1, Q3)	Median (Q1, Q3)	Median (Q1, Q3)	Median (Q1, Q3)
Cluster	1 - Diverse	2 - Highly Active	3 - FA-Content-oriented	4 - Disengaged
Frequency	293	233	336	233
Tactic 1	0.04 (0.00, 0.08)	0.04 (0.02, 0.08)	0.00 (0.00, 0.08)	0.00 (0.00, 0.00)
Tactic 2	0.06 (0.03, 0.10)	0.03 (0.00, 0.06)	0.07 (0.00, 0.17)	0.12 (0.00, 0.23)
Tactic 3	0.08 (0.04, 0.12)	0.10 (0.06, 0.14)	0.11 (0.06, 0.17)	0.12 (0.00, 0.20)
Tactic 4	0.34 (0.28, 0.42)	0.36 (0.26, 0.44)	0.29 (0.20, 0.38)	0.29 (0.14, 0.40)
Tactic 5	0.07 (0.03, 0.14)	0.09 (0.03, 0.15)	0.06 (0.00, 0.13)	0.08 (0.00, 0.15)
Tactic 6	0.18 (0.12, 0.24)	0.19 (0.14, 0.26)	0.17 (0.10, 0.25)	0.10 (0.00, 0.20)
Tactic 7	0.06 (0.03, 0.09)	0.06 (0.03, 0.10)	0.05 (0.00, 0.10)	0.00 (0.00, 0.11)
Tactic 8	0.10 (0.06, 0.15)	0.07 (0.05, 0.10)	0.12 (0.06, 0.17)	0.11 (0.00, 0.20)
Final Exam	22.00 (16.00, 30.00)	21.67 (15.00, 31.00)	18.00 (12.25, 26.08)	14.67 (10.00, 19.00)
Mid-term Exam	15.00 (12.00, 17.00)	16.00 (13.00, 18.00)	13.00 (10.75, 16.00)	12.00 (8.00, 15.00)

percent of sessions are composed of meta-cognitive actions.

Accounting for a third of all sessions, Tactic 4 was the most likely destination for all states, which indicates that after almost any study tactic, students were most likely to return to the tactic of complete focus on content access (reading materials). The most probable transition was from Tactic 4 to itself, implying that after a content access study session, students were most likely to continue accessing content (i.e., reading materials) in the subsequent session.

2) *Clustering Tactics to Identify Learning Strategies*: The second step towards addressing our first and second research questions is the clustering of students on the basis of tactic sequences. By grouping students on their tactic sequences, we could identify learning strategies and assess whether such patterns are associated with differences in academic outcomes. The clustering analysis identified 4 clusters in the first half of the course and 4 clusters in the second. On the basis of the 8 HMM states (i.e., study tactics), the 4 clusters for the first half of the course may be described as follows:

- Cluster 1 – Diverse: In the median, 34 percent of study sessions consist of content access (Tactic 4), 18 percent consist of FA/content access/meta-cognitive (Tactic 6), and 10 percent of sessions are a mixture of actions (Tactic 8). In terms of study session sequence length, students in this cluster have the second longest sequences behind Intensive students.
- Cluster 2 – Highly Active: the most active group. The distribution of tactics was very similar to that of Diverse. In the median, 36 percent of study sessions consist of content access (Tactic 4), 19 percent consist of FA/content access/meta-cognitive (Tactic 6), and 7 percent of sessions are a mixture of actions (Tactic 8). The most apparent feature of this group was the length of study sequences: almost twice that of the second longest group (Diverse).

- Cluster 3 – FA-Content-oriented: content access actions accounted for only 29 percent of sessions, and students focused comparatively more on the other learning resources, particularly FA and SA materials.
- Cluster 4 – Disengaged: study sequences of students in this group were typically half as long as FA-Content-oriented students. In the median, 29 percent of study sessions are content access (Tactic 4), only 10 percent consist of FA/content access/meta-cognitive (Tactic 6), 12 percent consist of SA activity (Tactic 2), and 11 percent are accounted for by a mixture of actions (Tactic 8).

Table III describes the resulting clusters in terms of the proportion of the eight HMM states used for clustering, the mid-term exam scores across clusters, and the final exam scores. For all variables, the table shows the median, the 25th and 75th percentiles.

On the basis of the eight state HMM, the four clusters for the second half of the course may be described as follows:

- Cluster 1 – Highly Content-oriented: on average, over 40 percent of study sessions were composed of Tactic 4, representing an exclusive focus on content access actions. There was also a moderate focus on Tactic 6 (mixed actions), and Tactic 2 (SA actions). Students in this cluster were associated with the longest, most diverse study session sequences.
- Cluster 2 – Intensive: this group engaged similarly to Highly Content-oriented students but focused less on content access (Tactic 4) and typically spent 17 percent of their time in Tactic 2 (SA actions) and 11 percent of their time in Tactic 3 (FA/content access). Study sessions sequences were approximately as long as that of the Highly Content-oriented, but the distribution over states was far less varied in the longest sequences.
- Cluster 3 – Assessment-oriented: in the median, students spent only 25 percent in Tactic 4 (content access), instead focusing on Tactic 2 (SA actions) 33 percent of the time. Additionally, 15 percent of sessions were

TABLE IV
SUMMARY STATISTICS FOR THE PROPORTION OF THE EIGHT TACTIC TYPES WITHIN CLUSTERS FOR THE SECOND HALF OF THE COURSE

	Median (Q1, Q3)	Median (Q1, Q3)	Median (Q1, Q3)	Median (Q1, Q3)
Cluster	1 - Highly Content-oriented	2 - Intensive	3 - Assessment-oriented	4 - Highly Assessment-oriented
Frequency	176	383	224	223
Tactic 1	0.05 (0.02, 0.11)	0.03 (0.00, 0.08)	0.00 (0.00, 0.06)	0.00 (0.00, 0.00)
Tactic 2	0.11 (0.07, 0.14)	0.17 (0.09, 0.23)	0.33 (0.27, 0.39)	0.36 (0.22, 0.50)
Tactic 3	0.09 (0.05, 0.13)	0.11 (0.04, 0.15)	0.15 (0.07, 0.21)	0.16 (0.00, 0.27)
Tactic 4	0.41 (0.33, 0.47)	0.33 (0.23, 0.42)	0.25 (0.17, 0.36)	0.17 (0.07, 0.27)
Tactic 5	0.03 (0.02, 0.08)	0.05 (0.00, 0.12)	0.05 (0.00, 0.11)	0.00 (0.00, 0.11)
Tactic 6	0.14 (0.09, 0.19)	0.10 (0.04, 0.17)	0.05 (0.00, 0.08)	0.00 (0.00, 0.09)
Tactic 7	0.03 (0.02, 0.06)	0.04 (0.00, 0.07)	0.00 (0.00, 0.06)	0.00 (0.00, 0.00)
Tactic 8	0.08 (0.04, 0.12)	0.10 (0.04, 0.17)	0.05 (0.00, 0.09)	0.09 (0.00, 0.23)
Final Exam	26.67 (18.67, 34.00)	21.17 (15.00, 30.00)	17.50; (13.00, 26.00)	15.00 (10.00, 20.50)
Mid-term Exam	16.00 (14.00, 18.00)	15.00 (12.00, 17.00)	13.00 (11.00, 16.00)	13.00 (10.00, 15.00)

spent on Tactic 3 (FA/reading). Study session sequences, however, were the second shortest.

- Cluster 4 – Highly Assessment-oriented: compared to their contemporaries, this group spent around 36 percent of their time engaging with Tactic 2 (SA actions) and 16 percent of their time on Tactic 3 (FA/content access); they tended to focus more on assessment than on content access and students were associated with the shortest study sequences.

Table IV describes the resulting clusters in terms of the proportion of the eight HMM states used for clustering, the mid-term exam scores and the final exam scores across clusters. For all variables, the table shows the median, the 25th and 75th percentiles.

3) *MANOVA Analysis*: For both halves of the course, Box’s M test was not accepted so, following Bray and Maxwell [8], we opted for the Pillai-Bartlett test statistic as it is the most robust to violations of the test assumptions.

For the first half of the course a statistically significant MANOVA effect was obtained, Pillai’s Trace = 0.46 $F(3, 1091) = 24.41$, $p = 5.90e-99$. The multivariate effect was estimated at $\eta^2 = 0.15$, implying that 15 percent of the variance in the dependent variables was accounted for by differences in the student cluster assignments which, according to Cohen [9], is a medium effect size. To follow up, a series of one-way ANOVA tests were conducted, with Bonferroni corrections. First, the assumption of homogeneity of variance was tested using Levene’s F test and was found to be violated for all variables but Tactic 5. The ANOVA model for Tactic 5 was found to be insignificant, but Kruskal-Wallis tests for all other dependent variables were significant.

For the second half of the course, a statistically significant MANOVA effect was also obtained, Pillai’s Trace = 0.55, $F(3, 1002) = 27.7$, $p = 0$. The multivariate effect was estimated at $\eta^2 = 0.18$, implying that 18 percent of the variance in the dependent variables was accounted for by the differences in the student cluster assignments. Following Cohen [9], this is

considered a medium effect size. To follow up, a series of one-way ANOVA tests were conducted, with Bonferroni corrections. First, the assumption of homogeneity of variance was tested using Levene’s F test and was found to be violated for all variables but Tactics 1 and 6. The ANOVA models for Tactics 1 and 6 were found to be significant, and Kruskal-Wallis tests for all other dependent variables except Tactic 5 were significant.

Tables documenting these results, identifying which states (tactics) vary significantly across clusters (learning strategies) may be found in the supplemental materials, which can be found on the IEEE Xplore Digital Library at <https://ieeexplore.ieee.org/document/8331946/>, along with further discussion.

B. Research Question 2: Learning Strategies and Academic Outcomes

1) *Pairwise Comparisons of Learning Strategies and Course Performance*: To answer our second research question, we conducted pairwise comparisons of the identified clusters with respect to scores in the mid-term and final exams. As shown in Tables V and VI, for the first half of the course, all cluster pairs, except for the pair Diverse-Highly Active in the final exam, were significantly different. Effect sizes (r) ranged from small to medium.

For the second half of the course, we conducted pairwise comparisons of clusters with respect to scores in the final exam, as displayed in Table VII. All cluster pairs were significantly different. Effect sizes (r) ranged from small to medium.

C. Research Question 3: Impact of Interventions

1) *Intervention Analysis and Changes in Learning Strategies*: To address our third research question and ascertain whether the additional feedback intervention, absent in the first year but phased in over the second and third, had an impact on students’ choice of learning strategies, for both

4. INTERPRETABLE REPRESENTATIONS OF STUDENT BEHAVIOUR

TABLE V
PAIRWISE CLUSTER COMPARISON FOR THE FIRST HALF OF THE COURSE
BASED ON MID-TERM EXAM SCORES

Cluster 1	Cluster 2	Z	p	r	alpha
Diverse	Disengaged	8.321	0.000*	0.363	0.008
Highly Active	FA-Content-oriented	6.168	0.000*	0.259	0.017
Highly Active	Disengaged	9.537	0.000*	0.442	0.025
FA-Content-oriented	Disengaged	4.974	0.000*	0.209	0.033
Diverse	FA-Content-oriented	4.373	0.000*	0.174	0.042
Diverse	Highly Active	-2.068	0.039	0.090	0.050

TABLE VI
PAIRWISE CLUSTER COMPARISON FOR THE FIRST HALF OF THE COURSE
BASED ON FINAL EXAM SCORES

Cluster 1	Cluster 2	Z	p	r	alpha
Diverse	Disengaged	9.109	0.000*	0.397	0.008
Highly Active	Disengaged	8.481	0.000*	0.393	0.017
FA-Content-oriented	Disengaged	5.310	0.000*	0.223	0.025
Diverse	FA-Content-oriented	4.660	0.000*	0.186	0.033
Highly Active	FA-Content-oriented	4.301	0.000*	0.180	0.042
Diverse	Highly Active	-0.135	0.893	0.006	0.050

TABLE VII
PAIRWISE CLUSTER COMPARISON FOR THE SECOND HALF OF THE COURSE
BASED ON FINAL EXAM SCORES

Cluster 1	Cluster 2	Z	p	r	alpha
Highly Content-oriented	Assessment-oriented	7.584	0.000*	0.379	0.008
Highly Content-oriented	Highly Assessment-oriented	9.861	0.000*	0.494	0.017
Intensive	Highly Assessment-oriented	7.834	0.000*	0.318	0.025
Highly Content-oriented	Intensive	4.509	0.000*	0.191	0.033
Intensive	Assessment-oriented	4.379	0.000	0.178	0.042
Assessment-oriented	Highly Assessment-oriented	3.789	0.000*	0.179	0.050

halves of the course, chi-squared tests of independence were performed to examine the relation between year and cluster assignment. For the first half of the course the association between these variables was significant, $\chi^2(6, N = 1095) = 44.70, p < .01$. However, in the latter half the association was not significant, $\chi^2(6, N = 1095) = 6.91, p = 0.33$. Thus, in the first half of the course, we reject the null of independence and find that cluster assignments were dependent upon the year variable but, in the second half of the course, we do not reject the null of independence. To further ascertain the association of the interventions with years of implementation, Tables VIII and IX provide distribution of cluster assignments stratified by year.

Following Lust et al. [39], we place the cluster labels for both halves of the course into a series of transitional matrices. These matrices, presented in Tables X through XII, depict how students' learning strategies, as characterised by their cluster assignments, changed between the two halves of the course. Specifically, each row represents a cluster in the first half of the course, and column values represent the proportion of that cluster that transitioned to the cluster (in the 2nd half of the course) represented by that column.

TABLE VIII
DISTRIBUTION OF STUDENTS ACROSS CLUSTERS, STRATIFIED BY YEAR FOR
THE FIRST HALF OF THE COURSE

	Diverse	Highly Active	FA-Content-oriented	Disengaged
2014	0.194	0.168	0.369	0.269
2015	0.287	0.295	0.216	0.202
2016	0.297	0.177	0.338	0.188

TABLE IX
DISTRIBUTION OF STUDENTS ACROSS CLUSTERS, STRATIFIED BY YEAR FOR
THE SECOND HALF OF THE COURSE

	Highly Content-oriented	Intensive	Assessment-oriented	Highly Assessment-oriented
2014	0.167	0.361	0.206	0.266
2015	0.195	0.393	0.201	0.210
2016	0.164	0.382	0.249	0.204

TABLE X
CLUSTER TRANSITION MATRIX IN 2014 WITH NO INTERVENTION

	Highly Content-oriented	Intensive	Assessment-oriented	Highly Assessment-oriented
Diverse	0.185	0.481	0.241	0.093
Highly Active	0.468	0.404	0.043	0.085
FA-Content-oriented	0.092	0.449	0.214	0.245
Disengaged	0.000	0.040	0.320	0.640

In 2014, Highly Active and Diverse—the top two performing groups of students—were most likely to transition to Highly Content-oriented and Intensive. For the lower performing FA-Content-oriented students, the most likely transition was to the Intensive strategy, a strategy associated with improved academic performance. However, the transition to a Highly Assessment-oriented strategy was also likely, which is associated with worse academic performance. For Disengaged students, the most likely transition was to the Highly Assessment-oriented strategy, although the transition to Assessment-oriented (associated with improved academic performance) was also likely. In 2015, while Diverse students remained most likely to transition to Intensive, for Highly Active the most likely transition was also to the Intensive strategy, associated with a decline in academic performance (though Highly Content-oriented remained a highly likely transition). For FA-Content-oriented students, the most likely transition was to the Highly Assessment-oriented strategy, associated with a decline in performance, though there was also a high chance of transitioning to the Intensive strategy which was associated with the opposite. Similar to 2014, for Disengaged students, the most likely transition was to the Highly Assessment-oriented strategy, although the transition to Assessment-oriented (associated with improved academic performance) was also likely. In 2016, Content-Oriented and Diverse—the top two performing groups of students—were most likely to transition to Highly Content-Oriented and Intensive. For the lower performing FA-Content-oriented students, the most likely transitions were to the Intensive strategy and the Assessment-oriented strategy. These transitions were associated with an improvement and little change in academic performance, respectively. Similar to previous years, for Disengaged students, the most likely transition was to the Highly

TABLE XI
CLUSTER TRANSITION MATRIX IN 2015 WITH PARTIAL INTERVENTION

	Highly Content-oriented	Intensive	Assessment-oriented	Highly Assessment-oriented
Diverse	0.103	0.546	0.227	0.124
Highly Active	0.427	0.466	0.087	0.019
FA-Content-oriented	0.027	0.288	0.315	0.370
Disengaged	0.000	0.176	0.255	0.569

TABLE XII
CLUSTER TRANSITION MATRIX IN 2016 WITH FULL INTERVENTION

	Highly Content-oriented	Intensive	Assessment-oriented	Highly Assessment-oriented
Diverse	0.175	0.518	0.204	0.102
Highly Active	0.463	0.427	0.098	0.012
FA-Content-oriented	0.041	0.320	0.347	0.293
Disengaged	0.000	0.148	0.333	0.519

Assessment-oriented strategy, although the transition to Assessment-oriented (associated with improved academic performance) was increasingly likely.

IV. DISCUSSION

A. Research Question 1: From Study Sessions to Learning Strategies

The results of clustering confirm the existence of well differentiated patterns in student learning behavior examined at the level of study sessions. As manifestations of students' learning strategies [65], these clusters provide insight as to how students opted to interact with the learning resources. The identified groups reflect those reported in previous research [11], [18], [30], [39] and are well described by the different approaches to learning (deep versus surface) as outlined by Biggs [5]:

- Group characterised by a high activity level and a thorough engagement with a variety of learning materials. In the first half of the course, Diverse and Highly Active students matched this group as they were highly active and engaged with a variety of resources. The same could be said for Highly Content-oriented and Intensive students in the latter half of the course. Compared to their less successful peers, the extent of these students' engagement with the reading resources was one of the most prominent characteristics of their learning strategies. The fact that these students performed highly in both the mid-term and final exams suggests that they were successful in regulating their learning.
- Group with a low activity level, focused on a narrow range of the available resources. Students in this group adopt a selective approach to the learning materials and are performance oriented. Disengaged and FA-Content-oriented students in the first half of the course may be considered shallow learners: they engaged the least with the reading materials and were instead preoccupied with summative and, to a lesser extent, formative assessment resources. The same may be said of Assessment-oriented and Highly Assessment-oriented students in the second half of the course. Though these students appear to have focused more on assessment than on reading materials—indicating performance orientation—their low exam performance evidences that their regulation of learning was far from optimal.

Given that our study was performed on the same course as Jovanović et al. [30] (albeit three years of data rather than 2014 alone), a closer comparison of our respective research questions, methodologies, and results is warranted. While our first and second research questions were very similar—

investigating whether patterns in student learning behaviour that are indicative of learning strategies can be identified and whether such patterns are associated with course performance—our research embarked on a third, more qualitative research question to try and ascertain the association, if any, between learning strategies and feedback interventions. More pertinent, however, are the methodological differences between the two studies. While Jovanović et al. [30] clustered processed sequences of learning actions to identify distinct learning strategies, in our study, we first segmented these sequences of learning actions into study sessions before fitting an HMM to these segmentations. In this study, chronologically ordered sequences of HMM states (or study tactics) were clustered for each of the two halves of the course to identify learning strategies followed by individual learners.

Jovanović et al. [30], however, first clustered individual study sessions each composed of action sequences. This was followed by clustering of students based on the counts of the occurrences of each session cluster. While differing numbers of learning strategies were identified in the two studies, upon analysis, these differences are well described by the methodological differences. Yet, conceptually, the results of these two studies are related and can be interpreted by Biggs' [5] two approaches to learning (deep versus surface). The rationale for our multi-step analysis is that by distinguishing between study tactics and learning strategies, it provides researchers with more granular, low-level interpretations of students' behavioral patterns (through analysing the parameters of HMM states). Comparing our results with those of Jovanović et al. [30], the fact that these two distinct methodologies provide similar high-level interpretations may be seen as partial validation of our method.

Though our interpretation of the identified clusters is very similar to Jovanović et al. [30], the number of actual clusters (or learning strategies) identified differs. Regarding our methodology, this may raise a concern about the extent to which our results are dependent upon the learning context in which they arise. However, this should come as no surprise: research has found that learning strategies may be shaped by course design, and that instructional conditions must be taken into account when estimating the effects of specific LMS features on academic outcomes [17]. The methodology we propose does not aim to identify learning strategies that are independent of the context within which they arise, but rather assesses the ability of students to take advantage of the educational environment within which they find themselves. That is, a proxy for their ability to effectively self-regulate.

The results of the MANOVA analyses indicate that the learning strategies identified are significantly different in terms of study tactic compositions. For the first half of the course we found a medium effect size, as cluster assignments

accounted for 15 percent of the variance in the dependent variables, indicating an important relationship between learning strategies and study sessions. For Tactics 4 and 7, the Kruskal-Wallis and Post-Hoc results indicate that while there are no significant differences between the top performing clusters (Highly Active and Diverse), these are significantly different from the remaining two (FA-Content-oriented and Disengaged), which are themselves significantly different. As Tactic 4 is representative of content access actions and Tactic 7 is representative of content access and meta-cognitive actions, this indicates that learning strategies associated with higher academic performance tend to engage more with course content and that the use of meta-cognitive resources is associated with more effective self-regulated learning.

For the second half of the course we found a medium effect size, as the cluster assignment accounted for 18 percent of the variance in the dependent variables. For Tactics 2 and 3, the Kruskal-Wallis and Post-Hoc results indicated that the top performing clusters (Highly Content-oriented and Intensive) were significantly different from the remaining two (Assessment-oriented and Highly Assessment-oriented) and, in the case of Tactic 2, were significantly different from each other. As Tactics 2 and 3 are representative of SA and FA actions, respectively, this indicates that learning strategies associated with higher academic performance tend to spend less time engaging with assessment materials. Instead, there were significant differences between all clusters regarding Tactic 4, which indicates that students whose learning strategies were associated with higher academic success tend to engage more with the recommended readings. Tables documenting these results may be found in the supplemental materials, available online, along with further discussion.

B. Research Question 2: Learning Strategies and Course Performance

Comparison of the identified learning strategies with respect to students' mid-term and final exam scores (Tables V, VI, and VII) demonstrate that for the first half of the course, academic performance was significantly different for all strategy pairs except for the pair Diverse-Highly Active in the final exam. This indicates that even in the first half of the course, students' choice of learning strategy was significantly associated with their final exam performance. In the second half of the course, strategies were only compared on the basis of the final exam, where all learning strategy pairs were significantly different. This was indicative of an association between the learning strategies adopted by students and their academic performance. Specifically, study sessions which were characterised by more diverse engagements patterns (Highly Active and Diverse in the first half, and Highly Content-oriented and Intensive in the second half) were associated with higher academic performance. By contrast, students who predominantly engaged least with content, preferring assessment materials (FA-Content-oriented and Disengaged in the first half and Assessment-oriented and Highly Assessment-oriented in the latter half) were associated with poor performance. This finding is consistent with previous research: students who

experiment with a variety of learning strategies are engaged in meta-cognitive monitoring and, thereby, more active self-regulated learning [13], [18], [20].

Our findings are also consistent with empirical studies that have examined learning strategies and how these relate to performance outcomes [30]. In keeping with Biggs [5], in our clusters, we recognise two broad approaches: a deep approach whereby students engage critically and thoroughly with the resources, and a shallow approach whereby students "skated along the surface" [41] of the materials and lacked critical engagement. Learning strategies adopted by the top performing groups in both halves of the course may be considered deep approaches, both in terms of the number of study sessions and the variety of actions of which their sessions were composed. By contrast, the lowest performing groups in both halves seemed to be practicing surface approaches. Across both halves of the course, the performance of identified strategies is consistent with the extent or absence of students' deep approach to the learning resources. This finding is consistent with meta-analysis by [55] that found positive, albeit small, correlations between students' performance and deep learning approaches, while surface approaches were found to be negatively correlated with academic outcomes.

C. Research Question 3: Intervention Analysis, Learning Strategies, and Course Performance

The feedback intervention for the first half of the course was phased in over the three years, being absent, present, present with respect to the years under investigation. Therefore, we may expect a change from the first to the second year, but less change from the second to the third. Tables VIII and IX provide partial support of this hypothesis: comparing the first two years, there is a large increase in the proportion of students in the highest performing group (Highly Active), a drop in mid-performing students (FA-Content-oriented), and a drop in the proportion of poorly performing students (Disengaged). Comparing the last two years, however, only part of this improvement remains (see Table VIII). The proportion of top performing students' transitions had been attenuated, and a larger proportion of students were found in the second highest performing cluster (Diverse). While a smaller proportion of students were found in the worst performing cluster (Disengaged), a larger proportion are found in the mid-range cluster (FA-Content-oriented).

In the second half of the course, the feedback intervention was phased in differently, being absent, present, with respect to the years under study. When comparing the first two years, similar to the first half of the course, there was an increase in the proportion of students in the highest performing group (Highly Content-oriented). In both low performing groups (Assessment-oriented and Highly Assessment-oriented), the proportion of students dropped between the first and second year (see Table IX). This may be in part due to lingering effects of the intervention in the first half of the course, but may also reflect cohort differences, a covariate for which we are unable to control due to the lack of permission to access additional variables. Comparing the second and third

year, however, while the proportion of students in the top two performing groups dropped, the number of students in the third highest group (Assessment-oriented) increased. This is a very similar pattern to that of the first half of the course and suggests that cohort effects may be hard to extract from the changes fostered by the interventions.

The three transition matrices (Tables X to XII), indicate that students can be broadly categorised into two groups: high and low performers. In the first year (2014), the two groups were comparatively rigid: both high and low performers tended to transition to their respective groups, though over a third of FA-Content-oriented students transitioned to an Intensive strategy, associated with improved performance. In the second year when the intervention was implemented, however, this pattern began to break down. In particular, while a larger proportion of FA-Content-oriented students transitioned to the Intensive strategy, a moderate proportion of Diverse students transitioned to Assessment-oriented, associated with lower academic performance, a trend which continues in the third year.

Prior studies have found that immediate feedback is often more effective at the level of study tactics [23], [37]. As the feedback was distributed on a weekly basis, and was based on the students' engagement and performance in the preceding week, it could meaningfully inform tactic changes, and gradually lead to strategy changes. Therefore, it is perhaps unsurprising that we observed such meaningful transitions regarding students' choice of learning strategies. Moreover, previous research has found that students who received more frequent feedback tended to be more successful than those who received feedback less frequently [56], [57].

V. LIMITATIONS

To identify learning strategies within trace data we adopted HMMs and clustering techniques. However, both of these are unsupervised machine learning methods and introduce an element of subjectivity. In particular, while the dendrograms indicated plausible clustering solutions, these were neither objectively right nor wrong and the choice was both unavoidably subjective and could have potentially impacted the study findings. Furthermore, while our choice of features (the proportion of different learning actions within study sessions) led to distinct study tactics, this removed all information about the chronological ordering of actions within study tactics. However, our methodology could be expanded to include intra-tactic transition graphs to provide greater insight into the patterns of student activity and self-regulating activity and further inform our choice of features.

Though we found significant differences in the tactic composition of the identified learning strategies, when comparing the distribution across these strategies on a year by year basis, our analysis was limited by our data. In particular, we did not have permission to access prior educational factors nor data on student demographics which could enable us to better estimate the actual impact of the interventions. Had such data been available, such variables would be useful additions to our analysis as confounding variables.

Furthermore, the techniques used in this analysis can provide descriptive accounts of regularities and have, in this case, identified clusters that were predominantly different in both composition and academic outcomes. However, these methods offer limited theoretical explanation of the identified patterns [54]. Questions such as the motivation or objectives of individual students, which could explain some of the variance in students' choice of learning strategies, remain unanswered. As Reimann et al. [54] suggest, this limitation could be addressed in the future with a multi-modal study, combining the techniques used here with data from other sources, such as think aloud protocols, student written reports, or self-reports. However, a challenge with the use of these data collection methods is that they can activate meta-cognitive processes that would not otherwise be triggered, can be susceptible to self-selection biases (self-reports), and present distorted memories about actual experiences (written reports) [18], [65].

Finally, as the study was based on the design based research method, where interventions are framed as within-subjects, repeated measurement designs rather than randomised controlled trials [53], the results must be treated with caution as claims to causality are not truly possible. The observed inter-year differences in the adopted learning strategies and intra-year strategy transitions may in part be due to cohort differences such as academic ability, for which we were unable to control. Accordingly, the analysis of strategy transitions should be met with the proviso that interpretations are qualitative. In future studies, this could be addressed by identifying confounding variables during intervention design and controlling for them.

VI. CONCLUSIONS

Within the context of an undergraduate FC course in computer engineering with a feedback intervention phased in over three years, the methodology that we propose in this paper enabled us to:

- Identify patterns in student learning behavior on the basis of study sessions. The benefit of such an approach is that by grounding the identification of learning strategies in narrower time frames, it facilitates research into how students adjust their learning strategies and would allow for more accurate assessment of how interventions impact on student behaviour. The analytical method used was able to discern a variety of study tactics and strategies that students adopted in preparation for face-to-face sessions. The composition of these learning strategies closely corresponds to those reported in previous research, particularly the deep versus surface approaches summarised in [5]. Significant differences in study tactic composition were found between the identified learning strategies. Compared to previous research, however, our method provided interpretable representations of student's self-regulating behavior at two theoretically inspired levels: that of learning strategies, and the study tactics that compose them [33].
- Identify a significant association between student learning strategies and academic outcomes. Consistent with

previous research, we found that more active, self-regulated learning was positively associated with academic performance [20], [30].

- Analysed the impact of the feedback intervention in terms of yearly distributions over learning strategies, and the transitions between the two halves of the course. Rather than being static, learning strategy transitions changed as the phased intervention was implemented. However, to estimate the impact of the intervention, information on students' prior educational ability, not present in our data, is required.

The consistency of our results with the existing literature demonstrates that the the identification of study tactics using study sessions is viable for future research. It also merits convenience, offering a close approximation of how students engage with a learning environment. Overall, our methodology provides interpretable explanations of study tactics and how learning strategies vary in their composition. Future research may expand upon this basis by further analysing intra-tactic transitions and transition graphs for different learning strategies.

The results of this study could be used to aid instructors in the monitoring of student engagement, to identify students at risk, and to guide students towards more effective study tactics and strategies. Since student learning strategies prior to the mid-term can well differentiate students' final exam scores, methods such as those used in this study could inform timely interventions to guide students to more successful engagement strategies.

REFERENCES

- [1] P. A. Alexander, S. Graham, and K. R. Harris, "A perspective on strategy research: Progress and prospects," *Educational Psychology Rev.*, vol. 10, no. 2, pp. 129–154, 1998.
- [2] R. Azevedo and J. G. Cromley, "Does training on self-regulated learning facilitate students' learning with hypermedia?" *J. Educational Psychology*, vol. 96, no. 3, pp. 523–535, 2004.
- [3] M. Bannert, P. Reimann, and C. Sonnenberg, "Process mining techniques for analysing patterns and strategies in students' self-regulated learning," *Metacognition Learn.*, vol. 9, no. 2, pp. 161–185, Aug. 2014.
- [4] M. Berland, T. Martin, T. Benton, C. P. Smith, and D. Davis, "Using learning analytics to understand the learning pathways of novice programmers," *J. Learn. Sci.*, vol. 22, no. 4, pp. 564–599, 2013.
- [5] J. Biggs, "What the student does: Teaching for enhanced learning," *Higher Education Res. Develop.*, vol. 18, no. 1, pp. 57–75, 1999.
- [6] P. Blikstein, M. Worsley, C. Piech, M. Sahami, S. Cooper, and D. Koller, "Programming pluralism: Using learning analytics to detect patterns in the learning of computer programming," *J. Learn. Sci.*, vol. 23, no. 4, pp. 561–599, 2014.
- [7] R. D. Bock, *Multivariate Statistical Methods in Behavioral Research*. Scientific Software, New York, NY, USA: McGraw-Hill, 1985.
- [8] J. H. Bray and S. E. Maxwell, *Multivariate Analysis of Variance*. Thousand Oaks, CA, USA: SAGE Publications, 1985 no. 54.
- [9] J. Cohen, *Statistical Power Analysis for the Behavioral Sciences*. Abingdon, U.K.: Taylor & Francis, 2013.
- [10] A. O. J. Cramer *et al.*, "Hidden multiplicity in exploratory multiway ANOVA: Prevalence and remedies," *Psychonomic Bulletin Rev.*, vol. 23, no. 2, pp. 640–647, Apr. 2016.
- [11] R. del Valle and T. M. Duffy, "Online learning: Learner characteristics and their approaches to managing learning," *Instructional Sci.*, vol. 37, no. 2, pp. 129–149, Mar. 2009.
- [12] C. EllenWeinstein, J. Husman, and D. Dierking, "Chapter 22. Self-Regulation Interventions with a Focus on Learning Strategies," in *Handbook of Self-Regulation*. Cambridge, MA, USA: Academic Press, 2012.
- [13] N. Entwistle, *Teaching for Understanding at University: Deep Approaches and Distinctive Ways of Thinking*. London, U.K.: Macmillan Education, 2009.
- [14] A. Field, J. Miles, and Z. Field, *Discovering Statistics Using R*. Thousand Oaks, CA, USA: SAGE Publications, 2012.
- [15] S. Freeman *et al.*, "Active learning increases student performance in science, engineering, and mathematics," *Proc. Nat. Acad. Sci.*, vol. 111, no. 23, pp. 8410–8415, Jun. 2014.
- [16] A. Gabadinho, G. Ritschard, N. Mller, and M. Studer, "Analyzing and visualizing state sequences in r with traminer," *J. Statist. Softw. Articles*, vol. 40, no. 4, pp. 1–37, 2011.
- [17] D. Gašević, S. Dawson, T. Rogers, and D. Gašević, "Learning analytics should not promote one size fits all: The effects of instructional conditions in predicting academic success," *Internet Higher Edu.*, vol. 28, pp. 68–84, 2016.
- [18] D. Gašević, J. Jovanovic, A. Pardo, and S. Dawson, "Detecting learning strategies with analytics: Links with self-reported measures and academic performance," *J. Learn. Analytics*, vol. 4, no. 2, pp. 113–128, 2017.
- [19] R. M. Gillies and M. Boyle, "Teachers' reflections on cooperative learning: Issues of implementation," *Teaching Teacher Edu.*, vol. 26, no. 4, pp. 933–940, 2010.
- [20] A. F. Hadwin, J. C. Nesbit, D. Jamieson-Noel, J. Code, and P. H. Winne, "Examining trace data to explore self-regulated learning," *Metacognition Learn.*, vol. 2, no. 2, pp. 107–124, Dec. 2007.
- [21] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning. Data Mining, Inference, and Prediction*. New York, NY, USA: Springer-Verlag, 2009.
- [22] J. Hattie, J. Biggs, and N. Purdie, "Effects of learning skills interventions on student learning: A meta-analysis," *Rev. Educational Res.*, vol. 66, no. 2, pp. 99–136, 1996.
- [23] J. Hattie and H. Timperley, "The power of feedback," *Rev. Educational Res.*, vol. 77, no. 1, pp. 81–112, 2007.
- [24] J. A. C. Hattie and G. M. Donoghue, "Learning strategies: A synthesis and conceptual model," *npj Sci. Learn.*, vol. 1, Aug. 2016, Art. no. 16013.
- [25] C. Hockings, S. Cooke, H. Yamashita, S. McGinty, and M. Bowl, "Switched off? A study of disengagement among computing students at two universities," *Res. Papers Edu.*, vol. 23, no. 2, pp. 191–201, 2008.
- [26] H.-T. Hung, "Flipping the classroom for English language learners to foster active learning," *Comput. Assisted Lang. Learn.*, vol. 28, no. 1, pp. 81–96, 2015.
- [27] W. Hung, "Theory to reality: A few issues in implementing problem-based learning," *Educational Technol. Res. Develop.*, vol. 59, no. 4, pp. 529–552, Aug. 2011.
- [28] H. Jeong, G. Biswas, J. Johnson, and L. Howard, "Analysis of productive learning behaviors in a structured inquiry cycle using hidden Markov models," in *Proc. 3rd Int. Conf. Educational Data Mining*, Jun. 2010, pp. 81–90.
- [29] H. Jeong, A. Gupta, R. Roscoe, J. Wagster, G. Biswas, and D. Schwartz, *Using Hidden Markov Models to Characterize Student Behaviors in Learning-by-Teaching Environments*. Berlin, Germany: Springer, 2008, pp. 614–625.
- [30] J. Jovanović, D. Gašević, S. Dawson, A. Pardo, and N. Mirriahi, "Learning analytics to unveil learning strategies in a flipped classroom," *Internet Higher Edu.*, vol. 33, pp. 74–85, 2017.
- [31] I. Khan and A. Pardo, "Data2U: Scalable real time student feedback in active learning environments," in *Proc. 6th Int. Conf. Learn. Analytics Knowl.*, 2016, pp. 249–253.
- [32] M. K. Kim, S. M. Kim, O. Khera, and J. Getman, "The experience of three flipped classrooms in an urban university: An exploration of design principles," *Internet Higher Edu.*, vol. 22, pp. 37–50, 2014.
- [33] J. R. Kirby, *Style, Strategy, and Skill in Reading*. Boston, MA, USA: Springer, 1988, pp. 229–274.
- [34] R. F. Kizilcec, M. Pérez-Sanagustín, and J. J. Maldonado, "Self-regulated learning strategies predict learner behavior and goal attainment in Massive Open Online Courses," *Comput. Edu.*, vol. 104, pp. 18–33, 2017.
- [35] V. Kovanovic, D. Gašević, S. Dawson, S. Joksimovic, and R. Baker, "Does time-on-task estimation matter? implications on validity of learning analytics findings," *J. Learn. Analytics*, vol. 2, no. 3, pp. 81–110, 2016.
- [36] V. Kovanović, D. Gašević, S. Joksimović, M. Hatala, and O. Adesope, "Analytics of communities of inquiry: Effects of learning technology use on cognitive presence in asynchronous online discussions," *Internet Higher Edu.*, vol. 27, pp. 74–89, 2015.
- [37] J. A. Kulik and C.-L. C. Kulik, "Timing of feedback and verbal learning," *Rev. Edu. Res.*, vol. 58, no. 1, pp. 79–97, 1988.

- [38] M. J. Lage, G. J. Platt, and M. Treglia, "Inverting the classroom: A gateway to creating an inclusive learning environment," *J. Econ. Edu.*, vol. 31, no. 1, pp. 30–43, 2000.
- [39] G. Lust, J. Elen, and G. Clarebout, "Regulation of tool-use within a blended course: Student differences and performance effects," *Comput. Edu.*, vol. 60, no. 1, pp. 385–395, 2013.
- [40] G. Lust, M. Vandewaetere, E. Ceulemans, J. Elen, and G. Clarebout, "Tool-use in a blended undergraduate course: In Search of user profiles," *Comput. Edu.*, vol. 57, no. 3, pp. 2135–2144, 2011.
- [41] F. Marton and R. Säljö, "On qualitative differences in learning: Ioutcome and process," *Brit. J. Edu. Psychol.*, vol. 46, no. 1, pp. 4–11, 1976.
- [42] G. S. Mason, T. R. Shuman, and K. E. Cook, "Comparing the effectiveness of an inverted classroom to a traditional classroom in an upper-division engineering course," *IEEE Trans. Educ.*, vol. 56, no. 4, pp. 430–435, Nov. 2013.
- [43] D. C. Moos and C. Bonde, "Flipping the classroom: Embedding self-regulated learning prompts in videos," *Technol. Knowl. Learn.*, vol. 21, no. 2, pp. 225–242, Jul. 2016.
- [44] J. O'Flaherty and C. Phillips, "The use of flipped classrooms in higher education: A scoping review," *Internet Higher Educ.*, vol. 25, pp. 85–95, 2015.
- [45] A. Pardo, "A feedback model for data-rich learning experiences," *Assessment Evaluation Higher Edu.*, in press. [Online]. Available: <https://doi.org/10.1080/02602938.2017.1356905>
- [46] A. Pardo, J. Jovanović, S. Dawson, D. Gašević, and N. Mirriahi, "Using learning analytics to scale the provision of personalised feedback," *British J. Educational Technol.*, 2017, in press. [Online]. Available: <http://dx.doi.org/10.1111/bjet.12592>
- [47] A. Pardo and N. Mirriahi, *Design, Deployment and Evaluation of a Flipped Learning First-Year Engineering Course*. Singapore: Springer, 2017, ch. 11, pp. 177–191.
- [48] D. Perera, J. Kay, I. Koprinska, K. Yacef, and O. R. Zaiane, "Clustering and sequential pattern mining of online collaborative learning data," *IEEE Trans. Knowl. Data Eng.*, vol. 21, no. 6, pp. 759–772, Jun. 2009.
- [49] D. N. Perkins, "The fingertip effect: How information-processing technology shapes thinking," *Edu. Res.*, vol. 14, no. 7, pp. 11–17, 1985.
- [50] R. Phillips, D. Maor, G. Preston, and W. Cumming-Potvin, "Exploring learning analytics as indicators of study behaviour," in *Proc. EdMedia: World Conf. Edu. Media Technol.*, Jun. 2012, pp. 2861–2867.
- [51] P. R. Pintrich and E. V. de Groot, "Motivational and self-regulated learning components of classroom academic performance," *J. Edu. Psychology*, vol. 82, pp. 33–40, 1990.
- [52] L. R. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proc. IEEE*, vol. 77, no. 2, pp. 257–286, Feb. 1989.
- [53] P. Reimann, "Connecting learning analytics with learning research: The role of design-based research," *Learn.: Res. Practice*, vol. 2, no. 2, pp. 130–142, 2016.
- [54] P. Reimann, L. Markauskaite, and M. Bannert, "e-Research and learning theory: What do sequence and process mining methods contribute?" *British J. Edu. Technol.*, vol. 45, no. 3, pp. 528–540, 2014.
- [55] M. Richardson, C. Abraham, and R. Bond, "Psychological Correlates of University Students' Academic performance: A systematic review and meta-analysis," *Psychological Bull.*, vol. 138, pp. 353–387, 2012.
- [56] A. W. Salmoni, R. A. Schmidt, and C. B. Walter, "Knowledge of results and motor learning: A review and critical reappraisal," *Psychological Bulletin*, vol. 95, no. 3, pp. 355–386, 1984.
- [57] R. A. Schmidt and R. A. Bjork, "New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training," *Psychological Sci.*, vol. 3, no. 4, pp. 207–218, 1992.
- [58] R. A. Susanne and P. Lajoie, *Teaching and Learning in Technology-Rich Environments*. Abingdon, U.K.: Routledge, 2006, ch. Chapter 35.
- [59] I. Visser and M. Speekenbrink, "depmixS4: An R package for hidden Markov models," *J. Statistical Softw. Articles*, vol. 36, no. 7, pp. 1–21, 2010.
- [60] C. E. Weinstein, *Assessment and Training of Student Learning Strategies*. Boston, MA, USA: Springer, 1988, pp. 291–316.
- [61] P. H. Winne, A. F. Hadwin, and L. P. Beaudoin, "nstudy: A web application for researching and promoting self-regulated learning (version 2.0) [computer program]. Burnaby, BC, Canada: Simon Fraser University," 2010.
- [62] P. H. Winne and N. E. Perry, *Measuring Self-Regulated Learning*. Orlando, FL, USA: Academic Press, 2000, pp. 531–566.
- [63] P. H. Winne, "Self-regulated learning viewed from models of information processing," in *Self-Regulated Learning and Academic Achievement: Theoretical Perspectives*, B. Zimmerman and D. Schunk, Eds. Abingdon, U.K.: Routledge, 2001, ch. 5, pp. 145–178.
- [64] P. H. Winne, "How software technologies can improve research on learning and bolster school reform," *Educ. Psychologist*, vol. 41, no. 1, pp. 5–17, 2006.
- [65] P. H. Winne, *Learning Strategies, Study Skills, and Self-Regulated Learning in Postsecondary Education*. Dordrecht, The Netherlands: Springer, 2013, pp. 377–403.
- [66] P. H. Winne and A. F. Hadwin, "Studying as self-regulated learning," in *Metacognition in Educational Theory and Practice*, D. J. Hacker, J. Dunlosky, and A. C. Graesser, Eds. Mahwah, NJ, USA: Lawrence Erlbaum Associates Publishers, 1998, pp. 277–304.
- [67] P. H. Winne and D. Jamieson-Noel, "Exploring students' calibration of self reports about study tactics and achievement," *Contemporary Educ. Psychol.*, vol. 27, no. 4, pp. 551–572, 2002.
- [68] M. Zhou and P. H. Winne, "Modeling academic achievement by self-reported versus traced goal learning and orientation," *Learn. Instruction*, vol. 22, no. 6, pp. 413–419, 2012.
- [69] B. Zimmerman, "Self-regulated learning and academic achievement: An overview," *Educational Psychologist*, vol. 25, no. 1, pp. 3–17, 1990.
- [70] B. Zimmerman, "Academic studing and the development of personal skill: A self-regulatory perspective," *Educational Psychologist*, vol. 33, pp. 73–86, 1998.



Ed Fincham is working towards the Ph.D. degree at the School of Informatics, University of Edinburgh. His research interests include developing computational methods to identify learning processes and provide insight into the social context within which such processes take place.



Dragan Gašević is currently a Professor of learning analytics with the Faculty of Education, Monash University and a Professor of learning analytics (part-time) with the School of Informatics, University of Edinburgh. His research involves developing computational methods that can shape next-generation learning technologies and advance our understanding of information seeking, sense-making, and self-regulated and collaborative learning.



Jelena Jovanović is currently an Associate Professor with the Department of Software Engineering, University of Belgrade, Belgrade, Serbia. Her research interests include combining machine learning methods with large-scale knowledge databases to better understand learning processes. She is also interested in how the insights obtained from such data can be effectively communicated to teachers and learners, as meaningful feedback.



Abelardo Pardo is currently an Associate Professor and Associate Head of teaching and learning with the School of Electrical and Information Engineering, The University of Sydney. He specializes in technology enhanced learning with an emphasis on learning and behavioral analytics, computer supported collaborative learning, and personalization of learning experiences.

4.3 Summary

The work presented in this chapter provides an introduction to our analysis of students' integration with their educational institution's academic system. In doing so, we argued for a reliance on digital trace data, as the literature suggests this provides a more valid representation of students' *actual* behaviour, as opposed to their *perceived* behaviour (P. Winne & Jamieson-Noel, 2002). However, it should be noted that trace data alone are no panacea, but should be supplemented with information about learning design and student dispositions to understand why students display certain behavioural patterns. (Rienties, Lewis, McFarlane, Nguyen, & Toetenel, 2018; Tempelaar et al., 2017).

Furthermore, while we conducted a computational analysis of students' behaviour, this was grounded within educational theory. Given our concerns about ensuring the validity of our methods, the consistency of our findings with the existing literature – in particular, that more active learning is positively associated with academic performance (Jovanović et al., 2017; Hadwin et al., 2007) – provides our student representation with a degree of external validity.

The method presented here also represents an initial attempt at the development of models which are readily interpretable. In this context, the HMM is an attractive model, offering not only transparent state parameters, but also explicit transition probabilities between these states. However, the interpretability of such a model depends in part on the features, and in part on the ease with which they may be understood. In the present study, the states may corresponded to study tactics but the features describing these states introduce a number of problems.

For instance, the features are all proportions of different types of learning actions taken on the platform, but ignore time-on-task. Furthermore, such a representation may be substantially influenced by course design, which may bias state parameters in favour of particular types of actions (such as video or reading actions), and undermine the generality of the approach.

In analysing student interactions, the diversity of students' motivations may be characterised by considerable heterogeneity in students' behaviour. Therefore, in a bid to circumvent the challenges we identified in the present chapter's methodology, in Chapter 5 we returned to the educational research literature to examine whether or not theories surrounding student "engagement" could provide a more general model for understanding students' interactions with course materials.

5

Validating a Theoretical Model of Student Engagement

5.1 Introduction

IN our initial foray into the analysis of students' academic interactions in Chapter 4, we extracted representations of students' interactions with three instances of a blended learning course. However, in doing so, we identified a number of challenges. In particular, despite couching our investigation in the theoretical literature surrounding "learning strategies", the validity and generality of our method were difficult to evaluate. Given the sensitivity of our features to the influence of course design, it is plausible that very different learning strategy constructs would emerge across different courses.

To address this limitation, we turned to the educational research literature and sought to find firmer theoretic grounding, which could account for the diversity of students' possible interactions. Such a decision is not uncommon in the learning analytics literature, as the role of theory is increasingly recognised as essential for informing the choice of questions asked and the hypotheses tested (Rogers et al., 2016; Wise & Shaffer, 2015). In searching for a comprehensive, theoretical model of student interactions, we focused on the literature surrounding engagement, particularly within the context of online learning environments such as MOOCs.

5.1.1 Engagement

Understanding student engagement is often considered essential for modeling learning and predicting learning outcomes. In non-formal, online educational settings, however, research into student engagement has been hindered by a lack of common understanding regarding how engagement should be defined and measured. This shortcoming is not limited to the study of engagement in non-formal settings: in spite of extensive research, the literature on engagement in traditional learning environments has been hindered by a lack of consensus regarding both the definition and the number of sub-types Reschly and Christenson (2012). For instance, Reschly and Christenson (2012) propose a quadripartite model, comprised of affective and cognitive engagement, representing student perceptions, which are then manifested by academic and behavioural engagement. Taken together, the authors argue that these four facets mediate the relationship between contextual factors and

learning outcomes. In a systematic review of the literature, Joksimović et al. (2018) highlight the differences that emerge between Reschly and Christenson's (2012) model and non-formal, online settings and propose a modified operationalisation of how engagement and learning in MOOCs can be assessed. Like Reschly and Christenson (2012), the model proposed by Joksimović et al. (2018) posits that student engagement may be subdivided into four facets, but adds that each of these may be captured by a series of trace-based metrics common to the learning analytics literature. Evaluating this model forms the basis of the current chapter.

5.1.2 Chapter Overview

The work presented in this chapter joins a growing body of literature within learning analytics that emphasises the importance of grounding computational analyses in theory. In doing so, we not only explore the ability of trace data to measure theoretical constructs, but also ensure that our measurements of those constructs are valid. Accordingly, our research questions relate to exploring the extent to which our construct of interest – student engagement – can be characterised by commonly used metrics in learning analytics research. Consequently, we employ exploratory factor analysis and confirmatory factor analysis to identify a latent structure within these trace data, before qualitatively comparing it to our theorised model. We then use this model to evaluate the extent to which these latent variable are associated with academic performance, before assessing the generality of this model across multiple course domains.

5.2 Publication: Counting Clicks is Not Enough: Validating a Theorized Model of Engagement in Learning Analytics

The following section includes the verbatim copy of the following publication:

Fincham, E., Whitelock-Wainwright, A., Kovanović, V., Joksimović, S., van Staalduinen, J.-P., & Gašević, D. (2019). Counting clicks is not enough: Validating a theorized model of engagement in learning analytics. In *Proceedings of the 9th international conference on learning analytics & knowledge* (pp. 501–510). LAK19. Tempe, AZ, USA: ACM. doi:10.1145/3303772.3303775

Counting Clicks is Not Enough: Validating a Theorized Model of Engagement in Learning Analytics

Ed Fincham
School of Informatics
University of Edinburgh
Edinburgh, United Kingdom
ed.fincham@ed.ac.uk

Alexander
Whitelock-Wainwright
School of Psychology
University of Liverpool
Liverpool, United Kingdom
a.wainwright@liverpool.ac.uk

Vitomir Kovanović
Teaching Innovation Unit and School
of Education
University of South Australia
Adelaide, Australia
vitomir.kovanovic@unisa.edu.au

Srećko Joksimović
Teaching Innovation Unit and School
of Education
University of South Australia
Adelaide, Australia
srecko.joksimovic@unisa.edu.au

Jan-Paul van Staaldouin
Education and Student Affairs
Delft University of Technology
Delft, Netherlands
j.p.vanstaaldouin@tudelft.nl

Dragan Gašević
Faculty of Education
Monash University
Clayton, Australia
dragan.gasevic@monash.edu

ABSTRACT

Student engagement is often considered an overarching construct in educational research and practice. Though frequently employed in the learning analytics literature, engagement has been subjected to a variety of interpretations and there is little consensus regarding the very definition of the construct. This raises grave concerns with regards to construct validity: namely, do these varied metrics measure the same thing? To address such concerns, this paper proposes, quantifies, and validates a model of engagement which is both grounded in the theoretical literature and described by common metrics drawn from the field of learning analytics. To identify a latent variable structure in our data we used exploratory factor analysis and validated the derived model on a separate subsample of our data using confirmatory factor analysis. To analyze the associations between our latent variables and student outcomes, a structural equation model was fitted, and the validity of this model across different course settings was assessed using MIMIC modeling. Across different domains, the broad consistency of our model with the theoretical literature suggest a mechanism that may be used to inform both interventions and course design.

KEYWORDS

Engagement, MOOCs, Factor Analysis, Structural Equation Modeling, Measurement Invariance

ACM Reference Format:

Ed Fincham, Alexander Whitelock-Wainwright, Vitomir Kovanović, Srećko Joksimović, Jan-Paul van Staaldouin, and Dragan Gašević. 2019. Counting Clicks is Not Enough: Validating a Theorized Model of Engagement in

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

LAK19, March 4–8, 2019, Tempe, AZ, USA

© 2019 Association for Computing Machinery.

ACM ISBN 978-1-4503-6256-6/19/03...\$15.00

<https://doi.org/10.1145/3303772.3303775>

Learning Analytics. In *The 9th International Learning Analytics & Knowledge Conference (LAK19)*, March 4–8, 2019, Tempe, AZ, USA. ACM, New York, NY, USA, 10 pages. <https://doi.org/10.1145/3303772.3303775>

1 INTRODUCTION

Within educational research and, particularly, learning analytics the term “engagement” has been frequently employed yet rarely defined. In such cases that permit a definition, it is often used to refer to arbitrary metrics hypothesized to capture some aspect of interaction [63]. Such practices amount to conflating heterogeneous measurements under the unified banner of “engagement”, and call into question the validity of these results. Recently, however, researchers have emphasized the importance of grounding computational analyses within existing educational research and theory [22, 42, 66]. Nevertheless, in the study of Massive Open Online Courses (MOOCs), little attention has been paid to the theoretical engagement literature.

Across diverse educational settings, student engagement has long been viewed as a factor that both drives learning and predicts academic, social, and emotional learning outcomes [20, 21, 52]. While much research has focused on the development of theoretical models of student engagement [4, 21, 52], these have typically been constrained to formal educational settings and less attention has been paid to online learning environments. The rapid proliferation of online learning, particularly MOOCs, has only recently been met with the corresponding development of theoretical models of student engagement within these environments [32]. In part, such developments come on the heels of wide-ranging criticisms of MOOCs, which are often characterized by low student motivation and engagement, resulting in limited social interactions and low completion rates [26, 39].

To date, much of the existing literature has relied on self-reported instruments to assess engagement [6]. There are, however, a number of limitations to the use of such data. In particular, associations between trace and self-reported data on the same construct are not consistently observed [23]. As Winne and Jamieson-Noel [65] have argued, this may be partially attributed to learner inaccuracy in

calibrating their self-reported learning behavior with their actual learning behavior. According to Zhou and Winne [68], this discrepancy is likely to be due to poor learner reflection. Moreover, Zhou and Winne's [68] study demonstrated that trace-based measures of student achievement goal-orientation had considerably stronger associations with learning outcomes than self-reported measures; a disparity the authors interpret as the difference between intention and actual behavior. Given the shortcomings of relying on self-reported data, we focus on investigating how engagement may be measured using trace data generated by student interactions with the course and the discussion forum.

Building on the conceptualization of engagement in MOOCs as provided by Joksimović et al. [32], this study examines how this complex, often misused and overgeneralized construct [6], can be operationalized in online learning environments. Specifically, using methods drawn from learning analytics and following Joksimović et al. [32], we model academic, behavioral, cognitive, and affective engagement and investigate the association between context, engagement, and learning outcomes. In so doing, we subject our theoretical model to robust empirical validation, investigating whether or not the identified model structure is congruent with the theory; whether or not the identified structure of our model is consistent across different courses, contexts, and pedagogic approaches; and the extent to which our model is predictive of learning outcomes.

2 THEORETICAL BACKGROUND

In reviewing the learning analytics literature, many existing approaches do not fully meet the criteria commonly used for asserting construct validity [36, 46, 47] (that is, whether or not a metric measures what it purports to measure). Following Messick [46], validity may be deconstructed into three core types. These include *structural validity*, or the fidelity of the scoring structure to the structure of the construct itself [43]; *external validity*, which includes supportive or dissuasive evidence arising from related constructs [46]; and *generalizability*, or the extent to which metrics generalize across populations and contexts [13, 58].

In the case of learning analytics, and particularly the study of engagement in online settings, structural validity (that is, theoretical foundations) have received limited attention [22, 54, 66]; simple count metrics count for little if we know not what they measure. This paucity of structural grounding also impacts upon external validity since, without theoretically formulated relations, correlations between constructs cannot be tested. Finally, it is also unclear the extent to which existing results are generalizable and consequential (that is, actionable) [22]. Accordingly, in defining, identifying, and validating a model of engagement, ensuring construct validity holds is central to our approach.

2.1 Reconstituting Learning in MOOCs

In the study of MOOCs, research has predominately focused on student persistence and the development of predictive models of dropout or academic performance [32]. While the proliferation of MOOCs has provided fertile grounds for research, the field has often been criticized for being primarily observational and lacking in appropriate rigor [32]. For instance, Reich [51] argued that, in spite of the vast quantity of data collected from students' activity, MOOC

research has failed to provide causal links between the observed metrics and student learning. This shortcoming is in part attributable to the lack of theoretically informed approaches employed in the analysis of data generated within online learning environments [22, 32, 66]. Though special issues and institutional reports have offered some insight into student engagement within MOOCs, little evidence has been provided regarding factors that contribute to learning [15, 51]. As argued by Joksimović et al. [32], the paucity of insight into student engagement offered by the existing literature can be attributed to a lack of understanding that non-formal education settings profoundly differ from more traditional pedagogic structures [62]. For instance, technology has facilitated the design of courses which cater to vast quantities of students in ways that are not possible in more traditional learning environments [51].

In investigating student engagement within MOOCs, the literature has emphasized the importance of forum participation, interaction with the course materials, and participation with assessment activities [41, 56, 59]. These activities are variously referred to as "discussion behaviors" [63], simply "behavior" [48, 49] or even "engagement" itself [56, 59, 60]. In spite of the fact that this emphasis on different activities within MOOCs suggests a multidimensional construct, a number of researchers have sought to observe engagement through unidimensional metrics. For instance, engagement has been characterized as discussion forum participation [61, 63], watching video lectures [41], or completing course assessments [64, 67]. The overarching understanding is that more active engagement with the course content, and more intensive interaction with peers, leads to higher course grades, greater learning, and greater course persistence.

Several researchers, however, have moved beyond observing single metrics to operationalizing engagement in MOOCs as a complex, multidimensional construct. For instance, Ramesh et al. [48, 49] defined engagement within online learning environments as a complex interaction between behavioral, linguistic, and social cues. The authors hypothesized a tripartite latent variable structure consisting of active engagement, passive engagement, and disengagement, and showed that their model provided better predictive accuracy of course success than individual measures, such as the number of video lectures watched, or the number of forum messages posted or viewed. While comprehensive, the extent to which their modeling approach connects with existing research on student engagement across different educational settings and MOOC domains is open to question. Accordingly, it is essential to ground any analysis of student engagement in the existing educational literature.

2.2 Theorizing About Engagement

Understanding student engagement, its metrics and mediating factors, is essential for modeling learning and predicting learning outcomes in non-formal, online educational settings. To date, such research has been hindered by a lack of common understanding regarding how engagement should be defined and measured in the context of online learning environments [32]. Having a generally accepted conceptualization of engagement would enable more comprehensive insight into the factors that influence learning within

MOOCs, as well as how these factors could be generalized across different platforms or compared within diverse contexts [15, 32]. Moreover, it would allow research to move beyond observing student “click data” and explore how the quantity and quality of interactions with course content and peers could predict course outcomes and persistence.

This shortcoming, however, is not limited to the study of engagement in non-formal settings: in spite of extensive research, the study of engagement in formal learning settings has also been hindered by a lack of consensus regarding both the definition and the number of sub-types [52]. At minimum, there is broad agreement that engagement is comprised of behavior indicative of participation and some affective element [19]. While some researchers have added a cognitive component [4, 12, 21], Reschly and Christenson [52] propose a quadripartite model, comprised of affective and cognitive engagement, representing student perceptions, which are then manifested by academic and behavioral engagement. Taken together, these four facets mediate the relationship between contextual factors and learning outcomes.

This taxonomy, however, does not readily transfer to online educational settings. Accordingly, in a systematic review of the literature, Joksimović et al. [32] highlight the differences that emerge between Reschly and Christenson’s [52] model and non-formal, online settings and propose a modified operationalization of how engagement and learning in MOOCs can be studied. In particular, Joksimović et al. [32] redefine contextual factors as demographic (age, prior education, and prior experience), classroom (peers, course design, and course platform), and individual needs (prior intentions, interest in topic, interest in MOOC learning). Furthermore, Joksimović et al. [32] redefine learning outcomes as immediate and course level. Finally, the results of the review are used to categorize common methods from MOOC research into the four facets, as described below.

The study of affective and cognitive engagement draws on the analysis of student-generated artifacts. This focus on artifacts is based on the premise that in computer-mediated environments, learning is primarily expressed through the artifacts that students create [24, 33], namely, posts in a discussion forum. Accordingly, in the context of affective engagement, this body of research has relied on linguistic indices to assess positive and negative emotions extracted from forum posts [1, 60]. By contrast, cognitive engagement refers to students’ motivational goals and self-regulated learning skills [21, 52]. In lieu of measuring these constructs, the quality of discourse is often treated as a proxy of cognitive engagement [32], as manifested by linguistic indicators, such as text cohesion and narrativity [16, 31, 63].

The metrics used to measure academic engagement in online learning environments broadly align with those used in more traditional classroom settings [4, 52]. Namely, the time spent on course activities (e.g. viewing pages, completing quizzes and assignments) also known as time on task [40], course attendance (or number of logins), the accuracy and completion rate on quizzes and assignments, and the credit earned towards course completion. Behavioral engagement “draws on the idea of participation; it includes involvement in academic and social or extracurricular activities and is considered crucial for achieving positive academic outcomes” [21, p. 60]. In online learning environments, this form of engagement

may be operationalised through participation in discussion forums, viewing lectures and course videos [41], and the persistence of their participation in the course.

3 STUDY FRAMING

The present study joins an increasing body of literature that emphasizes the importance of grounding computational analyses within existing educational theory [22, 42, 66]. In doing so, we additionally explore the ability of trace data to measure theoretical constructs, whilst ensuring our measurements of these constructs are valid. The study also contributes to the “next generation of MOOC research” [51, p. 34] that seeks to explain learning processes and the factors that influence learning outcomes. Understanding the association between outcomes and learning-related constructs necessitates a more holistic approach [15]. Accordingly, framed by Reschly and Christenson’s [52] model of the association between learning context, learning processes, and learning outcomes, the present study proposes and empirically validates a modeling approach that captures the interactions of these constructs.

Like Reschly and Christenson [52], the model posited by Joksimović et al. [32] conceptualizes student engagement as mediating the association between contextual factors and learning outcomes. Accordingly, our research questions relate to exploring whether this hypothesized construct can be characterized by commonly used metrics in learning analytics research.

Though Joksimović et al. [32] provide a multi-faceted, theoretically grounded framework for engagement which encompasses the vast majority of such common metrics, the model has not been subjected to robust empirical investigation or validation. In our first research question, our study seeks to address this lacuna with the use of exploratory modeling techniques.

Research Question 1: Can we identify a latent variable model structure consistent with our theorized engagement framework for studying learning in non-formal educational settings?

Our research interests, however, are not constrained to just identifying a plausible metric-based model of engagement. Rather, we aim to situate the proposed model as a mediator between contextual variables and learning outcomes. That said, some researchers have speculated that cognitive and affective engagement may be antecedent to, and mediate the manifestation of, academic and behavioral engagement [32, 52]. To assess this, we must augment our validated model of engagement with a path analysis to investigate both how our model’s structure is predictive of course outcomes, and the interdependencies of the latent variables within the model itself.

Research Question 2: To what extent, and in what ways, are our latent variables associated with learning outcomes as measured by course grade? Furthermore, what associations exist between the latent variables themselves?

Our final research question considers the extent to which our structural model of engagement holds across contexts, in this case, different online courses. In latent variable modeling, such concerns

Table 1: Descriptive Statistics for Courses in Dataset

Course	Number of weeks	Certificate threshold	Median age	Top countries	Gender		Highest degree		
					Male	Female	High school or less	College (Bachelors or Associate)	Advanced (Masters or Doctorate)
Solving Complex Problems	5	60%	33	USA (21.5%) India (8.2%) UK (4.6%)	75.2%	24.4%	18.7%	38.9%	40.3%
Delft Design Approach	10	60%	31	USA (19.8%) India (10.9%) Netherlands (5.7%)	68.6%	31.0%	19.0%	42.0%	36.8%
Technology for Biobased Products	7	55%	30	USA (16.1%) India (11.2%) Netherlands (6.0%)	36.0%	63.6%	21.2%	36.8%	39.5%

relate to measurement invariance (explained in Section 4.2), which exists with a number of gradations. In this study, we were only interested in a less stringent form of invariance, assessing whether latent variable means varied between different subpopulations. This informs our third and final research question:

Research Question 3: *To what extent is the structure of our obtained latent variable model generalizable across different subject domains and course pedagogies?*

4 METHODS

4.1 Data Sources

This study used data from three MOOCs offered by TU Delft in 2014 and delivered on the EdX platform. These courses included Design Approach (DDA), where students learned to formulate strategic

vision, map user behaviors, and develop meaningful proposals for products and services; Technology for Biobased Products (TBP), where students gained the insights and tools needed for sustainable design of biotechnological processes; and Solving Complex Problems (SCP) where students developed analytical frameworks for deconstructing problems in complex multi-actor systems. These courses were drawn from disparate disciplines and represent different pedagogical structures. For instance, DDA involved multi-media assignments which were then subjected to peer review. Descriptive statistics of the three courses are presented in Table 1. The data included final course grade, all interaction logs, and discussion forum posts. Complete data for our selected metrics was available for 52 students in DDA, 32 students in TBP, and 146 students in SCP. Because of these limited sample sizes, student interactions were analyzed at the course level rather than, say, on a weekly basis.

4.2 Factor Analysis and Measurement Invariance

To measure student levels of engagement, 18 candidate metrics were identified and calculated (see Table 2). These were drawn from Joksimović et al. [32], who categorized commonly used metrics within learning analytics into the four facets of their theoretical model of engagement. While a number of these metrics were extracted directly from the course trace logs, some were extracted using IBM Watson’s Tone Analyzer tool [30] applied to each post in the discussion forum, while others were derived from discourse analyses using the Coh-Metrix computational linguistic facility [25, 45]. To standardize between courses and address normality concerns, variables were scaled on a course by course basis, and heavily skewed variables (namely, Videos Watched and Problem Submissions) were log transformed. The data from all three courses was aggregated and randomly split into two equal sub-samples.

To answer our first research question and identify a latent variable model structure, exploratory factor analysis (EFA) [18] was used on the first half of the dataset. While the factor structure was identified via an iterative process whereby variables that did not load or exhibited factor loadings greater than 1 were excluded [14, 35], the number of factors was selected via a parallel analysis scree plot. In brief, a parallel analysis involves the generation of a random data set of the same dimensions as the data being analyzed. Factor analysis is then performed on the random data to extract

Table 2: Candidate Metrics for Engagement Modeling

#	Metric	Description
Trace Logs		
1	Days Active	Number of days student logs in
2	Weeks Active	Number of weeks student logs in
3	Forum Posts	Number of forum posts by student
4	Videos Watched	Number of unique videos watched
5	Problem Submissions	Number of unique problem submissions
6	Time on Task	Median time on problem submission task
7	Average Problem Score	Average problem submission score (%)
IBM Tone Analyzer		
8	Sentiment	Student’s median post sentiment
9	Sadness	Student’s median post sadness
10	Joy	Student’s median post joy
11	Anger	Student’s median post anger
12	Disgust	Student’s median post disgust
13	Fear	Student’s median post fear
Coh-Metrix		
14	Narrativity	Student’s median post narrativity
15	Syntactic Simplicity	Student’s median post syntactic simplicity
16	Word Concreteness	Student’s median post word concreteness
17	Referential Cohesion	Student’s forum post referential cohesion
18	Deep Cohesion	Student’s median post deep cohesion

eigenvalues. To avoid bias, this process is repeated 20 times and, for each eigenvalue, an average is taken. These random eigenvalues are then compared with the eigenvalues of the real data, and factors in the real data are only retained if their eigenvalues are greater than the eigenvalues from the random data [27]. This analysis was conducted using the Psych package in R [53]. To allow for correlations between factors, oblimin rotation was used and, given the relative normality of our data, standardized coefficients were estimated using maximum likelihood [14]. This permitted the computation of a wide range of goodness of fit indices, and allowed testing for the significance of factor loadings and correlations as well as the computation of confidence intervals [17]. Using this identified structure, confirmatory factor analysis (CFA) was then conducted on the second half of the data, using the Lavaan package in R [55]. Since the data was treated as continuous, the MLR estimator (with robust standard errors) was used. This identified model structure was then compared and contrasted to our theoretical model.

To answer our second research question, a structural equation model (SEM) [37] was fitted to the entire dataset using the previously identified model structure. In addition to the factor structure, the SEM specifies a path analysis that enables us to evaluate associations between variables [37]. Of particular interest were the associations between our latent variables and course grades, as Reschly and Christenson [52] hypothesize that engagement may play a mediating role with regards to student outcomes. Model reliability, or the extent to which a measure produces similar results under consistent conditions, was assessed using composite reliability [50]. This method was selected because factor loadings cannot be assumed to be equal. While there is no exact criterion for reliability, a value of 0.70 is often cited as an acceptable cut off [57].

Across all models in the foregoing analysis, goodness of fit was assessed using the comparative fit index (CFI), the Tucker-Lewis index (TLI) and the root mean square error of approximation (RMSEA), with 90% confidence intervals. In addition, χ^2 statistics are reported and, though non-significant χ^2 values are desirable, in the event of a significant χ^2 test, no modifications (such as permitting factor indicator errors to correlate) were undertaken. For both the CFI and TLI, Hu and Bentler [29] argue that values close to or greater than 0.95 are indicative of good fit. Regarding the RMSEA, values in the range of 0.08 to 0.10 are indicative of a mediocre fit [10], while values below 0.06 would suggest a good fit [29]. However, it should be stated that the foregoing cut-off values are heuristics and, though widely used, they should be treated with caution [44].

The overarching aim of our analysis is to model engagement under different conditions, namely, different course contexts. This raises concerns regarding construct validity, particularly structural validity: does our model measure the same engagement construct across different courses? Since our approach is grounded in latent variable modeling, construct validity may be assessed using the latent variable modeling concept of measurement invariance [7]. In general, the question of invariance is whether or not, under different conditions, measurements yield measures of the same attributes [28]. Measurement invariance, however, is not a binary construct, but exists with a number of gradations. In addressing our third research question, we seek to validate whether or not our latent variable model has the same structure across all courses. In particular, whether the pattern of constrained and estimated

Table 3: EFA Standardized Loadings

Metric	Factor 1	Factor 2	Factor 3
Sentiment	0.93		
Joy	0.66		
Sadness	-0.64		
Problem Submissions		0.93	
Videos Watched		0.78	
Weeks Active		0.72	
Syntactic Simplicity			-0.86
Referential Cohesion			0.46
Word Concreteness			0.42

parameters holds across all courses. Due to the small sample sizes in our dataset, measurement invariance is tested using multiple indicator multiple cause (MIMIC) modeling [9, 34]. This method assesses latent mean differences across groups (courses) by incorporating grouping variables as covariates [9, 38] and regressing latent variables onto these covariates. A significant effect of a grouping covariate on a latent factor indicates population heterogeneity between groups. The MIMIC analysis was conducted using the Lavaan package in R [55]. The effects of covariates (dummy coded course assignment) were tested using the non-parametric bootstrap method with 1000 bootstrapped samples.

5 RESULTS

To answer our first research question and investigate whether a latent variable model, consistent with our quadripartite model of engagement (see Section 2.2), could be identified, we conducted a parallel analysis. The results of this analysis suggested a three-factor structure. EFA was then conducted using an iterative process whereby variables that did not load or exhibited factor loadings greater than 1 were removed [14, 35]. This resulted in a 9 variable, 3 factor model with a good fit to the data ($\chi^2(36, N = 115) = 14.66$, $p = 0.26$, CFI = 0.99, TLI = 0.97, RMSEA = 0.05 (90% CI: 0.00 - 0.11)). Standardized loadings for this model are reported in Table 3. The fit of the identified factor structure was then assessed on the second half of the data using CFA. This also resulted in a good fit to the data ($\chi^2(24, N = 115) = 26.26$, $p = 0.34$, CFI = 0.99, TLI = 0.99, RMSEA = 0.03 (90% CI: 0.00 - 0.08)). Standardized and unstandardized loadings, with standard errors and p-values, may be found in Table 4. All unstandardized loadings were statistically significant ($p < 0.05$) except for Factor 3. A number of standardized factor loadings were strongly related to the proposed factors (R^2 between 0.09 and 0.98).

To address our second research question, we augmented the identified factor structure with a path analysis (thus forming a SEM) to investigate the relationship between our latent variables with learning outcomes, as represented by course grade. The model, assessed on the entirety of the data, demonstrated mediocre fit ($\chi^2(30, N = 230) = 63.60$, $p < 0.001$, CFI = 0.96, TLI = 0.94, RMSEA = 0.07 (90% CI: 0.05 - 0.09)). Standardized and unstandardized loadings, with standard errors and p-values, may be found in Table 5. All unstandardized loadings were statistically significant ($p < 0.05$) except for Factor 3. Standardized factor loadings were strongly related to the proposed factors (R^2 between 0.33 and 0.95). The results of the path analysis, including standardized and unstandardized coefficients, standard errors, and p-values, may be found in Table 6. The path

LAK19, March 4–8, 2019, Tempe, AZ, USA

Fincham et al.

analysis indicates that, of the three factors, only factor 2 is significant and positively associated with students’ final grade. In terms of inter-factor associations, only factors 2 and 3 are significantly associated with each other. Regarding the composite reliability of the factor structure, reliability coefficients are far above the acceptable cut off of 0.7 for the overall model (CR = 0.89), Factor 1 (CR = 0.77) and Factor 2 (CR = 0.84). Factor 3, however, falls short (CR = 0.55). The structure and standardized path coefficients of our SEM are shown in Figure 1.

To validate the consistency of our latent variable model structure across the three courses in our dataset, a MIMIC analysis was conducted. The model, with course assignment as dummy coded covariates, demonstrated a mediocre fit to the data ($\chi^2(36, N =$

Table 4: CFA Factor Loadings

Factor	Metric	Loading		SE	R ²	p
		Unstd.	Std.			
1	Sentiment	1.00	0.99		.98	
1	Sadness	-0.35	-0.52	0.11	.27	.00*
1	Joy	0.36	0.57	0.10	.32	.00*
2	Videos Watched	1.00	0.60		.37	
2	Problem Submissions	2.11	0.92	0.42	.84	.00*
2	Weeks Active	0.86	0.79	0.15	.63	.00*
3	Syntactic Simplicity	1.00	0.29		.09	
3	Word Concreteness	-1.37	-0.51	0.82	.26	.09
3	Referential Cohesion	-2.44	-0.93	2.83	.87	.39

Table 5: SEM Factor Loadings

Factor	Metric	Loading		SE	R ²	p
		Unstd.	Std.			
1	Sentiment	1.00	0.96		.93	
1	Sadness	-0.39	-0.59	0.07	.35	.00*
1	Joy	0.37	0.61	0.05	.37	.00*
2	Videos Watched	1.00	0.66		.43	
2	Problem Submissions	1.95	0.95	0.18	.90	.00*
2	Weeks Active	0.75	0.78	0.08	.59	.00*
3	Syntactic Simplicity	1.00	0.65		.42	
3	Word Concreteness	-0.56	-0.45	0.30	.21	.06
3	Referential Cohesion	-0.71	-0.53	0.36	.29	.05*

Table 6: SEM Path Coefficients

Path	Loading		SE	p
	Unstd.	Std.		
Factor 1 → Grade	0.01	0.02	0.01	.60
Factor 2 → Grade	0.66	0.89	0.08	.00*
Factor 3 → Grade	-0.01	-0.02	0.04	.84
Factor 2 → Factor 1	0.02	0.01	0.19	.90
Factor 3 → Factor 1	0.13	0.09	0.16	.41
Factor 1 → Factor 2	0.01	0.01	0.04	.90
Factor 3 → Factor 2	0.20	0.29	0.09	.03*
Factor 1 → Factor 3	0.06	0.08	0.06	.38
Factor 2 → Factor 3	0.43	0.29	0.21	.05*

Table 7: MIMIC Modeling Bootstrapped Results

Factor	Covariate	Partial std. estimate	SE	p	95% CI
1	DDA	0.06	0.10	.43	[-0.14 – 0.24]
1	TBP	-0.01	0.10	.48	[-0.22 – 0.17]
2	DDA	-0.14	0.11	.31	[-0.35 – 0.09]
2	TBP	-0.00	0.10	.47	[-0.18 – 0.20]
3	DDA	0.06	0.14	.47	[-0.21 – 0.31]
3	TBP	-0.09	0.14	.44	[-0.34 – 0.17]

230) = 60.74, $p = 0.006$). Table 7 presents the mean partially standardized estimates, mean standard errors, mean p-values and 95% bootstrapped confidence intervals. By treating course assignment as a covariate, MIMIC modeling allows us to regress our latent variables onto these covariates and, in doing so, assess differences in latent variable means between courses. Table 7 shows the results of regressing DDA and TBP on SCP. The results show that there are no significant differences in factor means which implies factor means are homogeneous for all factors across the three courses under study.

6 DISCUSSION

6.1 Research Question 1: Identifying a Latent Variable Structure

The purpose of this study was to ground the analysis of student engagement in theory. In doing so, we aimed to address concerns regarding construct validity in the existing literature and propose a common framework for assessing engagement in online learning environments. This framework is grounded in the theoretical work of Reschly and Christenson [52] and adapted for learning at scale by Joksimović et al. [32]. Initial indicators, drawn from commonly used metrics in learning analytics, were selected because, conceptually, they cohered with one or more of the four facets of Joksimović et al.’s [32] model. Of these 18 metrics, the EFA utilized 9 to derive a model which, with CFI and TLI values in excess of the recommended cut-off of 0.95, fitted the data very well. RMSEA 90% confidence intervals also indicated a good fit. Likewise, the CFA, conducted on a separate data sample, found that this factor structure resulted in an excellent fit. CFI and TLI values were again in excess of the recommended cut-off, and RMSEA 90% confidence intervals indicated a very good fit. In both the EFA and the CFA, χ^2 values were insignificant, indicating that the null hypothesis of no difference between the observed data and our model is not rejected.

The factor structure that the EFA and CFA identified and validated is consistent with the theorized model proposed by Joksimović et al. [32] with one major caveat: our factor model is tripartite, whereas Joksimović and colleagues posited a quadripartite model. Inspecting our three latent variables, Factor 1, composed of the Sentiment, Sadness, and Joy exhibited in students’ forum posts, conceptually aligns with affective engagement. That is, the positive and negative emotions that students manifest during their interactions with the learning environment. Likewise, Factor 3, composed of the Syntactic Simplicity, Referential Cohesion, and Word Concreteness of students’ forum posts, aligns conceptually

with cognitive engagement. The rationale behind this facet of engagement is grounded in the premise that students' understanding of learning content in online learning environments is expressed through the artifacts that they create [24, 33]. Accordingly, the quality of their discourse may be treated as a proxy for their cognitive engagement [32].

Factor 2, composed of Problem Submissions, Videos Watched, and Weeks Active, is more problematic. While Factors 1 and 3 are consistent with affective and cognitive engagement, it is less clear whether Factor 2 describes academic engagement, behavioral engagement, or some combination of the two. This "conceptual haziness" [3, p. 382] is mirrored in the literature, which generally reflects substantial variations in terms, definitions, and coverage [2, 3, 21]. For instance, while some researchers have posited behavioral engagement as a unified construct, others have bifurcated it into academic and behavioral subtypes. To add to the confusion, these subtypes themselves are neither clearly nor consistently defined. Following Appleton et al. [4] and Christenson et al. [11], academic engagement captures a student's time on task, credit accrual, and homework completion. Behavioral engagement, by contrast draws on the ideas of attendance and participation; preparation for and involvement with academic and extra-curricular activities. Such definitions indicate that Factor 2 may partially represent both academic engagement (via Problem Submissions) and behavioral engagement (via Videos Watched and Weeks Active).

Returning to our first research question, our latent variable model, identified and validated on two distinct sub-samples of our data, is broadly consistent with our theorized framework. Even where our factor structure diverges from the model posited by Joksimović and colleagues, the discrepancy reflects the debate in the theoretical literature. Accordingly, our model provides some evidence in support of the latent structure of engagement posited by the educational research literature.

6.2 Research Question 2: Latent Variables and Learning Outcomes

Augmenting our factor structure with a path analysis allowed us to investigate the extent to which our latent variables (representing engagement) were associated with student outcomes (represented by final grade). The SEM, conducted on the entire dataset, demonstrated a moderate fit with CFI and TLI values close to the recommended cut-off while RMSEA 90% confidence intervals indicated a mediocre fit. Problematically, the χ^2 test was found to be significant, implying that further work with additional samples is required to determine whether or not the model is acceptable.

The results of the path analysis demonstrate that while Factors 1 and 3 (affective and cognitive engagement) are not significantly associated with students' final grade, Factor 2 (academic & behavioral engagement) is significantly and positively associated with grade. Given the standardized loadings on Factor 2, this implies that students' performance is positively associated with submitting more problem assignments, watching more videos and being active in more weeks. The insignificance of Factors 1 and 3, with regards to grade, is not inconsistent with the theoretical literature: rather, Reschly and Christenson [52] speculate that cognitive and affective

engagement may well mediate academic and behavioral engagement, which in turn mediate learning outcomes. In other words, when students engage, cognitive and affective change precedes academic and behavioral change. The results of the inter-factor regression provide partial support for this hypothesis: Factors 2 and 3 are significantly and positively associated with one another. However, this support should be tempered with the caveat that while SEM results may provide support for a theory being tested, they can neither prove nor disprove theory or causality [8, p. 8].

Composite reliability indicates how consistently the items of a measurement model reflect the same underlying variable [50]. A highly reliable metric is one that yields a similar result under similar conditions. Our overall model exhibits high reliability (CR = 0.89). This is notable because, while reliability and validity are distinct concepts, reliability is a necessary but not sufficient condition for validity [5], and ensuring the validity of our engagement model is central to this study. The grounding of our modeling in the theoretical literature, for instance, is designed to address concerns relating to structural validity in empirical engagement research. The broad consistency between our latent variable model and our theoretical framework lend support to the argument that our engagement model is structurally valid. However, there is more to validity than just structure; there is also generality, which we address in our third research question by investigating measurement invariance.

6.3 Research Question 3: Measurement Invariance

Our final research question relates to the construct validity of our analysis, in particular the generality of the results. Typically, when generalizing results beyond a given study, one should consider the size and representativeness of the sample used. While the dataset in this study was not large, relying on the data of 230 students across three courses, the courses themselves contained considerable differences in both content (with subjects ranging from biotechnology to design) and pedagogy (with methods ranging from standard assessment to multi-media assignments and peer-review). To test the generality of our latent variable model across these disparate courses, we assessed the extent of measurement invariance. The general question of invariance is "whether or not, under different conditions of observing and studying phenomena, measurements yield measures of the same attributes" [28, p. 117]. While measurement invariance exists with a number of increasingly strict gradations, we only analyzed the homogeneity of factor means across courses. This method assumes a similar factor structure but allows us to test for population heterogeneity [9, 34]. This was achieved using MIMIC modeling with bootstrapping. The results indicate that there are no differences in factor means across any of the courses. This provides some support for the argument that our empirical model generalizes across courses. Any conclusions, however, should be tempered with the caveat that there are far more stringent tests of measurement invariance but, given the nature of this study and our inability to control for a number of contextual factors that our theoretical model specifies, this remains a significant result.

LAK19, March 4–8, 2019, Tempe, AZ, USA

Fincham et al.

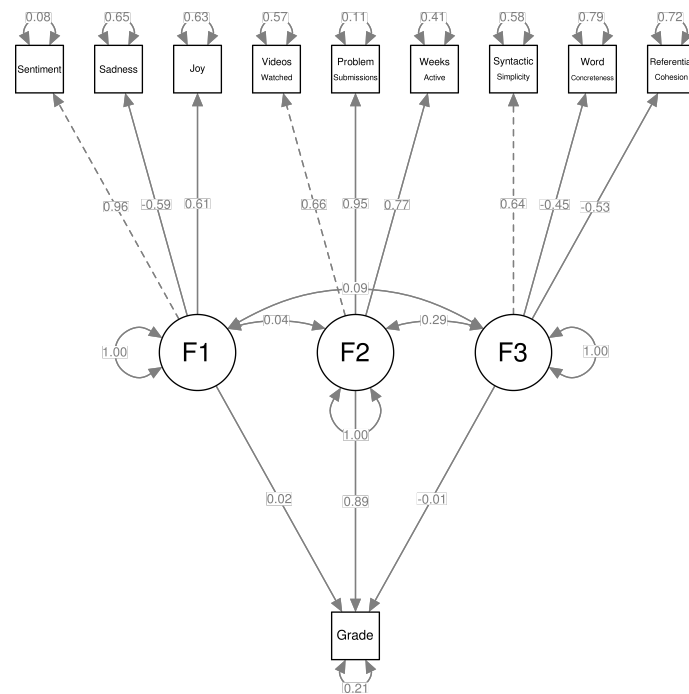


Figure 1: SEM of student engagement. Single-headed arrows represent direct influences while double-headed arrows represent correlations. All values are standardized coefficients.

7 IMPLICATIONS FOR PRACTICE

Understanding and, crucially, measuring student engagement is essential for the development of tools and practices to foster constructive actions and behavior. However, engagement is a complex, multi-dimensional construct that defies easy measurement. In spite of this complexity, the existing literature has often considered engagement through the lens of single metrics [41, 61, 63, 64, 67]. By contrast, the theorized framework adopted in this study permits such complexity by viewing engagement as a multi-faceted construct which mediates the relationship between contextual factors and learning outcomes [32]. The validation of this model marks an important step towards understanding the causal links between contextual factors, observed metrics (representing student behaviors), and student learning [51]. The influence of contextual factors may be measured at various levels, such as at the level of individual students or the overall course. Accordingly, our framework allows researchers and practitioners to analyze student engagement at both the level of the individual (how each student’s background and motivation influences their engagement) and the overall course (how course-level attributes such as course design influence students’ engagement).

While the precise relationship between these facets of engagements requires further study, our results suggest a mechanism which can inform both interventions and course design. In keeping with the predictions of the theoretical literature, we find that

student outcomes are strongly associated with academic or behavioral engagement, which in turn is associated with cognitive engagement. Though further research is required to verify that when students engage, cognitive and affective change precedes academic and behavioral change, such a mechanism would allow instructors or software systems to evaluate students in real time and intervene precisely where such interventions might have most impact. Though much research remains to be done, the measurement of student engagement provides considerable opportunities for actionable insight into students’ learning processes.

An important contribution of our engagement framework is that it permits practitioners to assess different facets of engagement, defined as theoretically grounded constructs. These facets – the latent constructs within our model – were identified in a manner consistent with the relevant practice in educational measurement, namely, EFA and CFA. The results of our analysis provide preliminary validation of these constructs, which stands in stark contrast to somewhat arbitrary metrics such as the count of page clicks or forum post views. Furthermore, the multi-dimensional structure of our framework enables practitioners to create interventions that target specific facets of engagement, which may also be personalized. For instance, a student who exhibits high levels of academic engagement and low levels of cognitive engagement may require a different intervention than a student who shows high cognitive engagement but only moderate academic engagement.

In addition, our results provide tentative evidence of model generality across courses. Although this requires further research and validation – in particular across different learning environments – such generality would allow for a context-independent terminology of engagement that would enable institutions to readily identify examples of best practice and effective learning design. This terminology would apply across courses, disciplines, and programs, and would facilitate constructive dialogue regarding the results of research informing quality enhancement across different learning environments.

8 LIMITATIONS AND FUTURE WORK

In deriving a model of engagement, student interactions were analyzed at the level of the entire course. Before investigating how information about student engagement can inform instructor interventions, however, subsequent research must first demonstrate that the results of this analysis are both valid and reliable across varying time scales.

Regarding the SEM analysis, it should be emphasized that while our results provide support for our theorized model of engagement, they cannot be taken as proof of theory or causality [8]. Furthermore, in our SEM the χ^2 test was found to be significant, indicating the need for a validation study on a larger dataset.

In assessing generality, our model may be subjected to more stringent measurement invariance tests to assess not only factor means but also indicator loadings, indicator intercepts, and indicator error variances. Our decision not to do so stemmed from our inability to control for contextual variables such as variations in course design that could impact on student engagement. Future work should investigate how such contextual variables may be controlled for and subject a more complete model of engagement to more stringent tests of measurement invariance.

9 CONCLUSION

Understanding and promoting student engagement is a central concern of educational research. While this research is not limited to formal educational settings, it is only recently that the appropriate theoretical developments have been made for studying engagement in non-formal educational settings. Accordingly, much research has been conducted without a well-defined conception of engagement itself, calling into question the validity of these results. In this study we have taken steps to address this oversight and have sought to demonstrate that theories drawn from the learning analytics literature may be subjected to empirical validation. For instance, the broad consistency of our latent variable model structure with the predictions of the theoretical literature lends support to the construct of engagement posited by Joksimović et al. [32]. The positive association between participatory, academically-oriented behavior (Factor 2) and student grades also coheres with the predictions found in the literature. Finally, the homogeneity of factor means across diverse courses provides tentative evidence of model generality. Such results have important implications for the practice of learning analytics: following further validation on larger, more diverse datasets, instructors will be able to evaluate and influence student behavior on the basis of theoretically and empirically grounded constructs.

REFERENCES

- [1] Panagiotis Adamopoulos. 2013. What makes a great MOOC? An interdisciplinary analysis of student retention in online courses. *International Conference on Information Systems (ICIS 2013): Reshaping Society Through Information Systems Design 5* (01 2013), 4720–4740.
- [2] Oqab Alrashidi, Huy Phan, and Bing Ngu. 2016. Academic Engagement: An Overview of Its Definitions, Dimensions, and Major Conceptualisations. 9 (11 2016), 41–52.
- [3] James J. Appleton, Sandra L. Christenson, and Michael J. Furlong. 2008. Student engagement with school: Critical conceptual and methodological issues of the construct. *Psychology in the Schools* 45, 5 (1 5 2008), 369–386. <https://doi.org/10.1002/pits.20303>
- [4] James J. Appleton, Sandra L. Christenson, Dongjin Kim, and Amy L. Reschly. 2006. Measuring cognitive and psychological engagement: Validation of the Student Engagement Instrument. *Journal of School Psychology* 44, 5 (1 10 2006), 427–445. <https://doi.org/10.1016/j.jsp.2006.04.002>
- [5] American Educational Research Association, American Psychological Association, National Council on Measurement in Education, Joint Committee on Standards for Educational, and Psychological Testing. 2014. *Standards for Educational and Psychological Testing*.
- [6] Roger Azevedo. 2015. Defining and Measuring Engagement and Learning in Science: Conceptual, Theoretical, Methodological, and Analytical Issues. *Educational Psychologist* 50, 1 (2015), 84–94. <https://doi.org/10.1080/00461520.2015.1004069>
- [7] A. Alexander Beaujean. 2014. *Latent variable modeling using R: A step-by-step guide*. Routledge.
- [8] Natasha K. Bowen and Shenyang Guo. 2011. *Structural Equation Modeling*. Oxford University Press.
- [9] Timothy A. Brown. 2015. *Methodology in the social sciences. Confirmatory factor analysis for applied research (2nd ed.)*. New York, NY, US: Guilford Press.
- [10] Robert C. MacCallum, Michael W. Browne, and Hazuki M. Sugawara. 1996. Power Analysis and Determination of Sample Size for Covariance Structure Modeling. 1 (06 1996), 130–149.
- [11] Sandra L. Christenson, Amy L. Reschly, James J. Appleton, Sarah Berman, Deanna M. Spanjers, and P. Varro. 2008. *Best practices in fostering student engagement*. National Association of School Psychologists, 1099–1120.
- [12] Sandra L. Christenson and Amy R. Anderson. 2002. Commentary: The centrality of the learning context for students' academic enabler skills. *School Psychology Review* 31 (2002), 378–393.
- [13] Thomas D Cook and Donald T Campbell. 1979. *Quasi-Experimentation: Design and Analysis Issues for Field Settings*. Houghton Mifflin.
- [14] Anna B. Costello and Jason W. Osborne. 2005. Best Practices in Exploratory Factor Analysis: Four Recommendations for Getting the Most From Your Analysis. 10 (01 2005), 1–9.
- [15] Jennifer DeBoer, Andrew D. Ho, Glenda S. Stump, and Lori Breslow. 2014. Changing 'course': Reconceptualizing educational variables for massive open online courses. *Educational Researcher* 43, 2 (2014), 74–94.
- [16] Nia Dowell, Oleksandra Skrypnyk, Srećko Joksimović, Arthur C. Graesser, Shane Dawson, Dragan Gašević, Pieter de Vries, Thieme Hennis, and Vitomir Kovanović. 2015. Modeling Learners' Social Centrality and Performance through Language and Discourse. (2015). <https://eric.ed.gov/?id=ED560532> presented at the Proceedings of the 8th International Conference on Educational Data Mining.
- [17] Leandre Fabrigar, Duane Wegener, Robert C. MacCallum, and Erin J. Strahan. 1999. Evaluating the use of Exploratory Factor Analysis in psychological research. 4 (01 1999), 272–299.
- [18] Andy Field, Jeremy Miles, and Zoe Field. 2012. *Discovering Statistics Using R*. SAGE Publications.
- [19] Jeremy D. Finn. 1989. Withdrawing From School. *Review of Educational Research* 59, 2 (1989), 117–142. <https://doi.org/10.3102/00346543059002117>
- [20] Charles W. Fisher and David C. Berliner. 1985. *Perspectives on instructional time*. New York: Longman.
- [21] Jennifer A Fredricks, Phyllis C Blumenfeld, and Alison H Paris. 2004. School Engagement: Potential of the Concept, State of the Evidence. *Review of Educational Research* 74, 1 (2004), 59–109. <https://doi.org/10.3102/00346543074001059>
- [22] Dragan Gašević, Shane Dawson, and George Siemens. 2015. Let's not forget: Learning analytics are about learning. *TechTrends* 59, 1 (Jan 2015), 64–71.
- [23] Dragan Gašević, Jelena Jovanović, Abelardo Pardo, and Shane Dawson. 2017. Detecting Learning Strategies with Analytics: Links with Self-reported Measures and Academic Performance. *Journal of Learning Analytics* 4, 2 (2017).
- [24] Peter Goodyear. 2002. *Psychological Foundations for Networked Learning*. Springer London, London, 49–75. https://doi.org/10.1007/978-1-4471-0181-9_4
- [25] Arthur C. Graesser, Danielle S. McNamara, and Jonna M. Kulikowich. 2011. Coh-Matrix: Providing Multilevel Analyses of Text Characteristics. *Educational Researcher* 40, 5 (2011), 223–234. <http://www.jstor.org/stable/41238943>
- [26] Christian Gütl, Vanessa Rizzardini, Rocael Hernándezang Chang, and Miguel Morales. 2014. Attrition in MOOC: Lessons Learned from Drop-Out Students. In *Learning Technology for Education in Cloud, MOOC and Big Data*, Lorna Uden,

LAK19, March 4–8, 2019, Tempe, AZ, USA

Fincham et al.

- Jane Sinclair, Yu-Hui Tao, and Dario Liberona (Eds.). Springer International Publishing, Cham, 37–48.
- [27] James C. Hayton, David G. Allen, and Vida Scarpello. 2004. Factor Retention Decisions in Exploratory Factor Analysis: a Tutorial on Parallel Analysis. *Organizational Research Methods* 7, 2 (2004), 191–205. <https://doi.org/10.1177/1094428104263675>
- [28] John L. Horn and John J. Mcardle. 1992. A practical and theoretical guide to measurement invariance in aging research. *Experimental Aging Research* 18, 3 (1992), 117–144. <https://doi.org/10.1080/03610739208253916>
- [29] Li-tze Hu and Peter M. Bentler. 1999. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal* 6, 1 (1999), 1–55. <https://doi.org/10.1080/10705519909540118>
- [30] IBM Tone Analyzer. n.d. <https://www.ibm.com/watson/developercloud/tone-analyzer.html>.
- [31] Srećko Joksimović, Nia Dowell, Oleksandra Skrypnik, Vitomir Kovanović, Dragan Gašević, Shane Dawson, and Arthur C. Graesser. 2015. How Do You Connect?: Analysis of Social Capital Accumulation in Connectivist MOOCs. In *Proceedings of the Fifth International Conference on Learning Analytics And Knowledge (LAK '15)*. ACM, New York, NY, USA, 64–68. <https://doi.org/10.1145/2723576.2723604>
- [32] Srećko Joksimović, Oleksandra Poquet, Vitomir Kovanović, Nia Dowell, Caitlin Mills, Dragan Gašević, Shane Dawson, Arthur C. Graesser, and Christopher Brooks. 2018. How Do We Model Learning at Scale? A Systematic Review of Research on MOOCs. *Review of Educational Research* 88, 1 (2018), 43–86. <https://doi.org/10.3102/0034654317740335>
- [33] Chris Jones. 2008. Networked Learning - A social practice perspective. In *Proceedings of the 6th International Conference on Networked Learning*. 616–623. <http://oro.open.ac.uk/20659/>
- [34] Karl Jöreskog and Arthur S. Goldberger. 1975. Estimation of a Model with Multiple Indicators and Multiple Causes of a Single Latent Variable. 70 (1975), 631–639.
- [35] Karl G Jöreskog. 1999, June 22. How Large Can a Standardized Coefficient be? <http://www.ssicentral.com/lisrel/techdocs/HowLargeCanaStandardizedCoefficientbe.pdf>
- [36] Michael T. Kane. 2013. Validating the interpretations and uses of test scores. *Journal of Educational Measurement* 50, 1 (2013), 1–73.
- [37] David Kaplan. 2009. *Structural Equation Modeling (2nd ed.): Foundations and Extensions*. SAGE Publications. <https://doi.org/10.4135/9781452226576>
- [38] Eun Sook Kim, Myeongsun Yoon, and Taehun Lee. 2012. Testing Measurement Invariance Using MIMIC: Likelihood Ratio Test With a Critical Value Adjustment. *Educational and Psychological Measurement* 72, 3 (2012), 469–492. <https://doi.org/10.1177/0013164411427395>
- [39] René Kizilec, Chris Piech, and Emily Schneider. 2013. Deconstructing disengagement: Analyzing learner subpopulations in massive open online courses. In *Proceedings of the Third International Conference on Learning Analytics & Knowledge (LAK '13)*. ACM, 170–179.
- [40] Vitomir Kovanović, Dragan Gašević, Shane Dawson, Srećko Joksimović, and Ryan Baker. 2016. Does Time-on-task Estimation Matter? Implications on Validity of Learning Analytics Findings. *Journal of Learning Analytics* 2, 3 (2016), 81–110. <https://doi.org/10.18608/jla.2015.23.6>
- [41] Nan Li, Lukasz Kidziński, Patrick Jermann, and Pierre Dillenbourg. 2015. *MOOC video interaction patterns: What do they tell us?* Lecture Notes in Computer Science, Vol. 9307. Springer, Berlin, Germany, 197–210. https://doi.org/10.1007/978-3-319-24258-3_15
- [42] Jason Lodge and Melinda Lewis. 2012. Pigeon pecks and mouse clicks: Putting the learning back into learning analytics. In *ASCLITE - Australian Society for Computers in Learning in Tertiary Education Annual Conference 2012*, M. Brown, M. Hartnett, and T. Stewart (Eds.). Australasian Society for Computers in Learning in Tertiary Education. <https://www.learnlib.org/p/42645>
- [43] Jane Loevinger. 1957. Objective Tests as Instruments of Psychological Theory. *Psychological Reports* 3, 3 (1957), 635–694. <https://doi.org/10.2466/pr0.1957.3.3.635>
- [44] Herbert W. Marsh, Kit-Tai Hau, and Zhonglin Wen. 2004. In Search of Golden Rules: Comment on Hypothesis-Testing Approaches to Setting Cutoff Values for Fit Indexes and Dangers in Overgeneralizing Hu and Bentler's (1999) Findings. *Structural Equation Modeling: A Multidisciplinary Journal* 11, 3 (2004), 320–341. https://doi.org/10.1207/s15328007sem1103_2
- [45] Danielle S. McNamara, Arthur C. Graesser, Philip M. McCarthy, and Zhiqiang Cai. 2014. *Automated Evaluation of Text and Discourse with Coh-Metrix*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511894664>
- [46] Samuel Messick. 1995. Standards of validity and the validity of standards in performance assessment. *Educational Measurement: Issues and Practice* (1995), 5–8.
- [47] Samuel Messick. 1995. Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American psychologist* 50, 9 (1995), 741.
- [48] Arti Ramesh, Dan Goldwasser, Bert Huang, Hal Daume, III, and Lise Getoor. 2014. Uncovering Hidden Engagement Patterns for Predicting Learner Performance in MOOCs. In *Proceedings of the First ACM Conference on Learning @ Scale Conference (L@S '14)*. ACM, New York, NY, USA, 157–158. <https://doi.org/10.1145/2556325.2567857>
- [49] Arti Ramesh, Dan Goldwasser, Bert Huang, Hal Daume III, and Lise Getoor. 2014. Learning Latent Engagement Patterns of Students in Online Courses. (2014). <https://www.aaai.org/ocs/index.php/AAAI/AAAI14/paper/view/8571>
- [50] Tenko Raykov. 1997. Estimation of Composite Reliability for Congeneric Measures. *Applied Psychological Measurement* 21, 2 (1997), 173–184. <https://doi.org/10.1177/01466216970212006>
- [51] Justin Reich. 2015. Rebooting MOOC Research. *Science* 347, 6217 (2015), 34–35. <https://doi.org/10.1126/science.1261627>
- [52] Amy L. Reschly and Sandra L. Christenson. 2012. *Jingle, Jangle, and Conceptual Haziness: Evolution and Future Directions of the Engagement Construct*. Springer US, Boston, MA, 3–19. https://doi.org/10.1007/978-1-4614-2018-7_1
- [53] William Revelle. 2018. *psych: Procedures for Psychological, Psychometric, and Personality Research*. Northwestern University, Evanston, Illinois. <https://CRAN.R-project.org/package=psych> R package version 1.8.4.
- [54] Tim Rogers, Dragan Gašević, and Shane Dawson. 2016. Learning analytics and the imperative for theory driven research. In *The SAGE Handbook of E-Learning Research, 2nd edition*, Caroline Haythornthwaite, Richard Andrews, Jude Fransma, and Eric Meyers (Eds.). SAGE Publications Ltd, London, UK, 232–250.
- [55] Yves Rosseel. 2012. lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software* 48, 2 (2012), 1–36. <http://www.jstatsoft.org/v48/i02/>
- [56] Jose Luis Santos, Joris Klerkx, Erik Duval, David Gago, and Luis Rodriguez. 2014. Success, Activity and Drop-outs in MOOCs: An Exploratory Study on the UNED COMA Courses. In *Proceedings of the Fourth International Conference on Learning Analytics And Knowledge (LAK '14)*. ACM, New York, NY, USA, 98–102. <https://doi.org/10.1145/2567574.2567627>
- [57] Albert H. Segars. 1997. Assessing the unidimensionality of measurement: a paradigm and illustration within the context of information systems research. *Omega* 25, 1 (1997), 107 – 121. [https://doi.org/10.1016/S0305-0483\(96\)00051-5](https://doi.org/10.1016/S0305-0483(96)00051-5)
- [58] Lee S. Shulman. 1970. Reconstruction of Educational Research. *Review of Educational Research* 40, 3 (1970), 371–396. <https://doi.org/10.3102/00346543040003371>
- [59] Tanmay Sinha and Justine Cassell. 2015. Connecting the Dots: Predicting Student Grade Sequences from Bursty MOOC Interactions over Time. In *Proceedings of the Second (2015) ACM Conference on Learning @ Scale (L@S '15)*. ACM, New York, NY, USA, 249–252. <https://doi.org/10.1145/2724660.2728669>
- [60] Conrad Tucker, Barton K. Pursell, and Aanna Divinsky. 2014. Mining student-generated textual data in MOOCs and quantifying their effects on student performance and learning outcomes. (2014). <https://peer.asee.org/22840> Paper presented at ASEE Annual Conference, Indianapolis, Indiana.
- [61] Duy Vu, Philippa Pattison, and Garry Robins. 2015. Relational event models for social learning in MOOCs. *Social Networks* 43 (2015), 121–135. <https://doi.org/10.1016/j.socnet.2015.05.001>
- [62] Sukaina Walji, Andrew Deacon, Janet Small, and Laura Czerniewicz. 2016. Learning through engagement: MOOCs as an emergent form of provision. *Distance Education* 37, 2 (2016), 208–223. <https://doi.org/10.1080/01587919.2016.1184400> arXiv:<https://doi.org/10.1080/01587919.2016.1184400>
- [63] Xu Wang, Diyi Yang, Miaomiao Wen, Kenneth R. Koedinger, and Carolyn Penstein Rosé. 2015. Investigating How Student's Cognitive Behavior in MOOC Discussion Forum Affect Learning Gains. In *Proceedings of the 8th International Conference on Educational Data Mining*. 226–233.
- [64] Jacob Whitehill, Joseph Williams, Glenn Lopez, Cody Austun Coleman, and Justin Reich. 2015. Beyond Prediction: First Steps Toward Automatic Intervention in MOOC Student Stopout. In *Proceedings of the 8th International Conference on Educational Data Mining (EDM'15)*. 171–178.
- [65] Philip H Winne and Dianne Jamieson-Noel. 2002. Exploring students' calibration of self reports about study tactics and achievement. *Contemporary Educational Psychology* 27, 4 (2002), 551–572.
- [66] Alyssa Friend Wise and David Shaffer. 2015. Why theory matters more than ever in the age of big data. 2 (01 2015), 5–13. <https://doi.org/10.18608/jla.2015.22.2>
- [67] Cheng Ye, John S. Kinnebrew, Gautam Biswas, Brent J. Evans, Douglas H. Fisher, Gayathri Narasimham, and Katherine A. Brady. 2015. Behavior Prediction in MOOCs Using Higher Granularity Temporal Information. In *Proceedings of the Second (2015) ACM Conference on Learning @ Scale (L@S '15)*. ACM, New York, NY, USA, 335–338. <https://doi.org/10.1145/2724660.2728687>
- [68] Mingming Zhou and Philip H. Winne. 2012. Modeling academic achievement by self-reported versus traced goal orientation. *Learning and Instruction* 22, 6 (2012), 413–419.

5.3 Summary

In this chapter we have emphasised the importance of grounding computational analyses in theory. In addition, we have demonstrated that theories drawn from the learning analytics literature may be subjected to empirical validation. For instance, in the context of student engagement, the broad consistency of our latent variable model structure with Joksimović et al.'s (2018) theorised model provides some validation of this model. Furthermore, the positive association between participatory, academically-oriented behaviour (captured by our second factor in the latent variable model) and students' academic performance also coheres with the predictions of the model.

While the present chapter does provide a methodology that may be used for preliminary investigations of theoretical constructs, it is not without its limitations. For instance, the analysis was conducted on only three courses, and while these spanned a broad array of disciplines, the cross-course differences in student numbers likely influenced the results of the analysis. Furthermore, while Tinto's (1975) model does make claims regarding academic performance, its central focus is on student persistence. This is an outcome which the current study has overlooked. Finally, while both this chapter and Chapter 4 argue that the student representations generated may be used to provide insight about students' learning, this claim has not been thoroughly investigated. Accordingly, developing a model that not only predicts student dropout, but also provides an interpretable explanation for those predictions forms the central basis of Chapter 6.

6

Towards Interpretable Insight

6.1 Introduction

THE present chapter concludes our investigation of the academic domain, as specified by Tinto’s (1975) model. While Chapter 5 focused on empirically validating a theoretical model of student interactions, it did not address dropout, which is Tinto’s (1975) primary concern. However, our research is not limited to evaluating this model, but is also concerned with providing interpretable insight. For the relevant stakeholders, unschooled in statistical analysis, interpretability is the *sine qua non* of a useful model. To address this, the current chapter proposes a model of students’ interactions with academic resources that represents an important step towards this goal. In doing so, we argue that a student’s interactions with course resources are an inherently temporal process, and the explicit modeling of these dynamics not only enables us to provide richly descriptive models that are interpretable, but also facilitates the prediction of student dropout. To validate this claim, we use our model’s representations to predict student outcomes, and further investigate the generality of our model by evaluating its cross-course predictive performance.

6.1.1 Model Interpretability

In recent years, model interpretability has become an increasingly prominent discussion within the machine learning community (Rosenfeld & Richardson, 2019; Gilpin et al., 2018). In this context, interpretability is loosely defined as the “ability to explain or to present in understandable terms to a human” (Doshi-Velez & Kim, 2017, p. 2). One of the challenges in this field is that while a number of researchers have proposed stringent codes of practice, such that a machine learning algorithm should not only provide a prediction, but also an explanation for this prediction (Doshi-Velez & Kim, 2017), there are currently no agreed standards as to what constitutes interpretability. In lieu of consensus, the literature argues that certain proxies must be relied upon (Doshi-Velez & Kim, 2017).

In the present chapter we argue that the dimensionality of a model’s representation may be used as a proxy: while there is no definite threshold beyond which a model becomes uninterpretable, lower dimensional representations provide more readily digestible information. Furthermore, since

students' interactions with a course are an inherently temporal process, we argue that the manner in which a model represents temporality will influence its interpretability. To provide a baseline for assessing model interpretability, we compare our approach to the work of Coleman, Seaton, and Chuang (2015), who adopted the topic modeling approach of LDA to identify behavioural patterns within students of a MOOC.

6.1.2 Chapter Overview

In the present chapter we propose a novel methodology which provides interpretable representations of student behaviour. These representations, however, are derived from a set of model parameters and thus require expert knowledge to interpret. Nevertheless, this work is an important first step towards providing users, such as course instructors, with representations of students' behaviour that may help them to identify students at-risk of dropping out. To ensure the validity of our representation, we perform a dropout prediction task, conducted on a weekly basis. To evaluate the generality of our method, we conduct the same prediction task, but with models trained in one course and evaluated in another.

6.2 Publication: The Road Not Taken: Preempting Dropout in MOOCs

The following section includes the verbatim copy of the following publication:

Fincham, E., Hoernle, N., Gal, K., & Gašević, D. (2019). The road not taken: Preempting dropout in moocs. *Manuscript submitted for publication*

The Road Not Taken: Preempting Dropout in MOOCs

Ed Fincham · Nick Hoernle · Kobi Gal · Dragan Gašević

Abstract The students of Massive Open Online Courses are characterised by considerable variation in their motivations and intentions. To date, the literature has addressed this diversity by either (1) identifying descriptive patterns of student activity or (2) using interaction data to build predictive models of students' learning outcomes. Limited work, however, has sought to bridge these two trends. We propose a novel methodology that provides interpretable representations of students' interactions which are also predictive of learning outcomes. Using a hidden Markov model (HMM), our approach learns a set of common interaction patterns which cohere with those found within the descriptive literature. The HMM also accounts for time, enabling our approach to provide temporally grounded insight into students' learning processes. To demonstrate the superior interpretability of our methodology, we compare it to a similar approach that uses Latent Dirichlet Allocation (LDA). We compare these two methods on the task of predicting dropout on a weekly basis and find that, in addition to outperforming the LDA method, and generalising across courses to a remarkable degree, our approach can distinguish between behavioural patterns that would be ambiguous under LDA. In summary, our

Ed Fincham
School of Informatics, University of Edinburgh, United Kingdom. E-mail: ed.fincham@ed.ac.uk

Nick Hoernle
School of Informatics, University of Edinburgh, United Kingdom. E-mail: n.s.hoernle@sms.ed.ac.uk

Kobi Gal
School of Informatics, University of Edinburgh, United Kingdom. E-mail: kgal@inf.ed.ac.uk

Dragan Gašević
Faculty of Information Technology, Monash University, Monash University, Australia.
E-mail: dragan.gasevic@monash.edu

method offers interpretable insight into learners' behaviour, their learning processes, and has the potential to identify students at-risk, even in lieu of course specific training data.

Keywords Interpretability · Dropout Prediction · Massive Open Online Courses

1 Introduction

Over the last decade, the proliferation of Massive Open Online Courses (MOOCs) has provided millions of students with unprecedented access to open educational resources. But, from their conception, MOOCs have attracted widespread criticism due to their low completion rates and students' limited interactions [22, 33]. Research has shown, however, that the heavy attrition of participants does not so much reflect the limitations of MOOCs, but rather the diverse motivations of their student-base [33]. For instance, while many students may intend to complete a course, others may enrol simply to review certain content before intentionally dropping out. Accordingly, the diversity of students' motivations should call into question any assumptions as to which academic outcomes are most valuable for learning within MOOCs [29, 33]. Furthermore, the diversity of students' motivation and, consequently, their behaviour on MOOC platforms, presents researchers with the challenge of developing predictive models of student behaviour that can not only account for such variety, but also provide interpretable insights to course instructors.

In providing course instructors with relevant insights, the MOOC research literature can be broadly categorised into two types. On the one hand, researchers have sought to understand the diversity of students' motivations by reducing students' interactions into a set of "prototypical", richly descriptive patterns of student activity [33]. On the other, researchers have foregone these descriptive representations, and have instead developed increasingly accurate predictive models of learning outcomes (such as dropout, course completion, etc.) [13, 27, 32, 49]. However, there has been limited research into the bridging of these disparate fields, leaving a substantial lacuna for the development of models that offer predictive accuracy *as well as* explanatory power [43].

In the present study, we address this gap in the literature by introducing an interpretable representation of students' interactions. This novel methodology (1) captures common interaction types which, similar to Kizilcec et al. [33], provide a high level description of students' interactions with the platform and (2) accommodates for the diversity of students' motivations via unique trajectories through this shared space of common interaction types. While students' motivations may be opaque, they play a mediating role in students' subsequent behaviour. This behaviour is observable, and may be directly modeled. Furthermore, students' motivations and behaviour also dictate their learning outcomes; as such, any representation which models students' behaviour ought also to predict their learning outcomes.

In employing representations that purport to both describe students' interactions with MOOCs and capture salient information about students' learning outcomes, it is essential to confirm their *construct validity* [40]. That is, we must ensure that our measurements are faithful to the construct they claim to measure [37]. Accordingly, these representations of behaviour should also be predictive of the consequences of this behaviour, namely, learning outcomes. For instance, Coleman et al. [9] presented a representation of student interactions, which was validated on the task of predicting course certification. In this study, while we also validate our student representation, we do so on the task of predicting dropout. This is because the emphasis on course completion belies a binary narrative of success, tied to adhering to the instructor's expectations [33], which is at odds with the manifest diversity of students' motivations.

Students' interaction patterns are not static, but are likely to change with time. For instance, as assignment deadlines approach, students may interact more with the learning environment. The explicit modeling of these dynamics not only enables us to provide richly descriptive models that are interpretable, but also captures salient information regarding learning outcomes. To achieve this, we utilise a general model for approaching problems in sequential data, namely, a hidden Markov model (HMM). The HMM consists of a set of latent states, which represent discrete points in time, along with a transition matrix that governs the switching dynamics between them. In the present context, these states represent common interaction patterns, which students transition between over time. By contrast, Coleman et al. [9] utilise Latent Dirichlet Allocation (LDA) which, through making the *bag of words* assumption, notably ignores all temporal structure within the data [6].

Finally, we emphasise the practical utility of our representation by demonstrating its generality. Specifically, we perform the same dropout prediction task, but use models trained on one MOOC and tested on another [18]. We conclude by arguing that the interpretable representations and predictions provided by our model can be used to facilitate future MOOC design objectives, such as the identification of students at-risk of dropping out.

2 Related Work

In reviewing the MOOC literature, we focus on three distinct aspects. First, it has been shown that students have diverse goals and motivations for enrolling, and thus their behaviour is characterised by considerable variety [19]. We contribute to this literature by presenting a machine learning model which learns a simplified representation of these interaction patterns. Secondly, we review the literature on interpretable machine learning, and argue for the importance of interpretable representations in MOOCs. Lastly, to ensure that these representations capture salient information about students' learning outcomes, and to link our research to the existing literature on student dropout, we review the well-studied task of dropout prediction.

2.1 Diverse Motivations; Diverse Outcomes

Academic achievement, in its various guises, is a common preoccupation of the literature, and is often operationalised as accumulated course grade [5, 8, 10, 14, 20, 31]. But such a definition may be viewed as the vestigial influence of formal educational settings, and may not be appropriate for non-formal settings, especially MOOCs. Given the broad diversity of backgrounds and intentions that govern MOOC participants, it would be remiss to make any assumptions as to what academic outcomes are most valuable for student learning [33]. But this leaves the researcher in a quandary: in the face of students' diverse motivations, what academic outcomes should be promoted? To answer this, and in a bid to circumvent preexisting assumptions about measuring learning outcomes through grades or continuous assessment, Sharma et al. [45] reconceptualised course outcomes to include learners whose intentions diverged from course completion, and instead emphasised the extent of interactions with the course materials as a measured outcome. In the literature, this has often been operationalised as dropout, and an extensive body of research has developed accurate predictive models of this outcome [13, 27, 32, 49]. However, accurate prediction should not be confused with greater understanding of the underlying phenomenon [43]; for this to be the case it is necessary that a model also possess explanatory power.

2.2 Interpreting Models; Illuminating Interactions

Regardless of how learning outcomes are measured, the provision of insight is beset by a number of challenges. In particular, MOOC datasets are typically large, highly dimensional and, in their raw form, opaque. While these properties make MOOC data amenable to analysis using machine learning methods, the decision as to how best to represent students' interactions is far from obvious. Nevertheless, the interpretability of these representations has become a prominent topic in the educational technology literature [43]. In particular, researchers have called for interdisciplinary approaches to develop learner models that provide interpretable and actionable insight which can, for instance, be derived from interpreting the parameters estimates of learned models [43].

Concerns regarding interpretability, however, are not limited to educational research. In the machine learning literature, the impetus for model interpretability is quite clear: as autonomous systems begin making decisions that were once the preserve of humans, it is necessary for these mechanisms to explain themselves [21, 44]. Although a number of researchers have proposed stringent codes of practice, such that a machine learning algorithm should also provide an explanation for its recommendation [12, 52], there are currently no agreed standards as to what constitutes an explanation [21]. Although consensus over such fundamental questions remains elusive, the machine learning literature contains a number of taxonomies for operationalising these con-

structs. In particular, Doshi-Velez et al. [12] provide a tripartite taxonomy for interpretability evaluations, which contains application-grounded evaluations, where model generated explanations of a phenomenon are compared to those of human experts; human-grounded evaluations, where model explanations are evaluated without domain experts; and functionally-grounded applications, which require no human input but rather use a formal definition of interpretability as a proxy for explanation quality [12]. Of these three, functionally-grounded evaluations are particularly relevant to the present study. This taxonomy, however, leaves quality undefined as it depends on the context of each evaluation [12]. In the case of evaluating the interpretability of a model of students’ interactions, we argue that the *dimensionality* of the representation may be used as a proxy for the quality of an explanation: while there is no definite threshold beyond which a model becomes uninterpretable, lower dimensional representations provide more readily digestible information about students’ behaviour to developers and course instructors. In addition, since students’ interactions are temporally grounded, we argue that the manner in which this temporality is described by the model is a useful heuristic for assessing model interpretability.

To understand interpretability, cross-model comparisons are vital. However, there is an extensive literature on analysing students’ interactions, and some approaches may be more suited than others to form the basis of a comparison. Early work in this area by Kizilcec et al. [33] used clustering methods to identify a small set of “prototypical” interaction patterns. Specifically, at each time-step, students were labelled as “on-track”, “behind”, “auditing”, or “out”. These sequences of labels were then clustered, forming four distinct sub-populations: “completing”, “auditing”, “disengaging”, and “sampling”, which were then compared on the basis of demographics, survey responses, and forum activity. Although this approach has much to recommend it – in particular, the ability to capture temporal change in students’ behaviour through transitions between descriptive, expert-defined states – these advantages come at a heavy price in terms of granularity, as a large portion of the statistical information within the original data is lost. Moreover, the statistical properties of this clustering approach were not explicitly evaluated. Later work addressed this shortcoming by proposing more quantitative approaches. For instance, Herskovic et al. [23] proposed that diverse interaction patterns reflect high levels of engagement. To quantify this diversity, the authors used PCA at regular time intervals to project student’s interactions along the top three principal components. These interactions were then clustered into common trajectories. Of particular relevance to the present study, however, is the work of Coleman et al. [9], who rejected the curation of rigidly defined feature sets and instead used Latent Dirichlet Allocation (LDA) to discover behavioural patterns, or “topics”, directly from students’ interactions (in an unsupervised manner). By representing students as a mixture of these latent types, the authors were able to predict student outcomes (specifically, certification) with a high degree of accuracy.

In addition to interpretability, ensuring validity – that is, the extent to which a given metric actually measures what it purports to measure [40] – is a central concern during model development. Following Messick [40], validity may be subdivided into three core types: *structural validity*, or the fidelity of the measurement to the structure of the construct itself; *external validity*, which includes supportive evidence from related constructs [40]; and *generalisability*, or the extent to which a measurement generalises across populations and contexts. The first of these, *structural validity*, is of particular relevance: if we take continued interactions with the course as our measured outcome [45], we can validate that our model measures this behaviour through a dropout prediction task.

2.3 Promoting Interactions; Predicting Dropout

While the combination of descriptive models of students’ interactions with predictive models of learning outcomes has been largely overlooked by the literature, a number of fields come close. For instance, the literature surrounding learning strategies (defined as “any thoughts, behaviours, beliefs or emotions that facilitate the acquisition, understanding, or later transfer of new knowledge and skills” [48, p. 227]), identify common interaction patterns which are associated with learning outcomes [14, 30, 35, 38]. For instance, Fincham et al. [14] used an HMM to identify common student interaction patterns (referred to as “study tactics”) and trajectories across them (referred to as “learning strategies”). While the authors found that students’ choice of learning strategy was associated with their final grade, they did not model this relationship and so were unable to evaluate whether or not students’ choice of learning strategy was predictive of their learning outcomes.

While much of the learning strategies literature has focused on course grade, the present study focuses on student dropout. To promote continued interactions with course materials, it must be possible to reliably identify when students are at risk of *not* interacting, that is, when students are at risk of dropping out. However, the difficulty of identifying dropout predictors that may be acted upon during the course, is that the requisite ground truth labels, required for any supervised learning, are only available at the *end* of the course, when any such intervention is moot [49]. This problem is compounded by the fact that researchers have typically measured test performance on the same dataset that is used for training, which can inflate performance statistics [18, 49]. These two challenges pertain to validity: in predicting dropout, the claim that a model can identify at-risk students is unfounded unless validated against a set of ground truth labels. Furthermore, a model may be valid, but only in certain contexts. For instance, Balakrishnan et al. [3] trained an HMM using hand-crafted features such as course progress and forum activity counts, as well as the percentage of lecture content viewed. While this method classified dropout with an AUC of 0.71 in the course under study, it relied upon discussion forums, the utilisation of which is often a function of course design,

such as tool availability, and external conditions, such as priming [19, 39, 51]. The existence of variation in course design calls the generality of this method into question. To address this potential limitation, it is important to design models that are both valid predictors of dropout, and have sufficient explanatory power to generalise.

3 Research Questions

The research questions detailed in this section investigate how a model can not only provide interpretable representations of students' diverse behaviour, but also be applied to common prediction tasks [43]. In answering these questions, our research addresses the gap in the literature between interpretable but solely descriptive models of students' interactions, and predictive models of students' learning outcomes.

To bridge this gap we employ an HMM, where the latent states correspond to common distributions over a set of course actions. The present study, however, is not the first that has sought to combine these two fields: previous work by Coleman et al. [9] used LDA to address the same problem, and the similarity of these two approaches warrants a direct comparison. Notably, both models claim to provide interpretable representations of student behaviour which can also be used to predict learning outcomes. In particular, we focus on dropout as other metrics, such as course completion, are difficult to reconcile with the diversity of students' motivations [33]. Bridging this gap and, in doing so, comparing the HMM and LDA approaches, leads us to the following research questions.

1. **RQ1:** What do latent states in the HMM and LDA models represent and how can these be interpreted in the context of students interacting with a MOOC?
2. **RQ2:** As two examples of unsupervised learning techniques, do the HMM and LDA models learn representations that are useful in predicting student dropout within MOOCs?
3. **RQ3:** To what extent do the representations learnt by the models exhibit cross-course generality in the dropout prediction task?

3.1 Research Question 1

In assessing the interpretability of the HMM and LDA models, we ground our analysis in the machine learning literature surrounding explainable AI [21]. In particular, we follow the taxonomy for investigating interpretability proposed by Doshi-Velez et al. [12], and argue that the dimensionality of the models' observation space may be used as a proxy for the quality of their explanation. Furthermore, the manner in which time is represented within these spaces is also an important heuristic.

The HMM has a number of beneficial properties that directly relate to our interpretability criteria. For instance, the model not only reduces the behaviour of the underlying process to a finite number of states, but also describes the transition dynamics between these states. The finite size of the state space directly relates to our first criterion of interpretability: accordingly, the HMM represents complicated trajectories as transitions between a limited number of states. While our approach echoes previous work by Kizilcec et al. [33], who used expert knowledge to identify four prototypical patterns of student interactions, our states are learnt directly from the data using unsupervised learning techniques, and may facilitate data-driven approaches to identify students at-risk of dropping out.

Before validating this claim, however, we must first evaluate the interpretability of our model; in particular, how it compares to a baseline, such as the LDA approach presented by Coleman et al. [9]. By treating user activity as a *bag of interactions* – analogous to the topic modeling concept of a *bag of words* – Coleman et al. [9] identified a set of use-cases which provide representations of students’ interactions. These use-cases may be understood as a probability distribution over the unique set of course resources. However, inferring these use-cases from student-interaction data is not well suited to the mixture representation that LDA engenders. For instance, consider two use-cases drawn from Coleman et al. [9]: *disengaging* and *completing*. A model that might describe a student as a mixture of these two is both counter-intuitive and largely uninterpretable (e.g., how can one be both completing and disengaging?). Furthermore, student activity within a course is an inherently temporal process, yet the ability to model sequential data is noticeably absent from the LDA setup. Additional issues arise for LDA when deciding on the number of unknown use-cases. Typically, this choice is settled using measures of held-out log-likelihood (with held-out perplexity [6] being a common measure of accuracy); however, models with optimal perplexity are not guaranteed to find the most interpretable topics within a corpus [7].

3.2 Research Question 2

In addition to arguing that our model provides more interpretable representations of students’ interactions, we contend that our approach facilitates the prediction of student dropout. In doing so, it may offer a balance between the descriptive work of Kizilcec et al. [33] and the literature on predictive models of dropout [13, 49]. While such a hybrid model may be less descriptive and less accurate than the current state-of-the-art in these two fields, the literature of explainable AI asserts that interpretability is the *sine qua non* of a useful prediction [12, 52]. Nevertheless, it is necessary to ascertain the extent to which our model is capable of capturing salient information regarding students’ dropout intentions. In addition to providing partial validation of our student representation, this prediction exercise emphasises the potential utility of our method to a course instructor. In keeping with our first research

question, this dropout prediction task is used to evaluate the performance of both LDA and our HMM-based approach.

3.3 Research Question 3

While there is an extensive literature on predicting dropout, it has been noted that the validity of these results has often gone unexamined [49]. While this is not uncommon in the learning sciences, and many existing approaches do not meet the criteria required for asserting validity [15, 34], in the present context these concerns relate to generality. While there have been some notable exceptions (see [32, 49, 50]), researchers have typically relied on a single course to train and test their methods, leading to a potential inflation of performance statistics [49]. Given these concerns, evaluating the extent to which our HMM approach generalises across course-contexts is essential for ensuring the validity of our method.

4 Methods

In this section we introduce our methodology, which we use to analyse two MOOCs. We detail the setup of both the LDA and HMM models, before describing how these models are used to address our three research questions.

4.1 Data Sources

This study used data from two MOOCs offered by separate institutions and drawn from disparate disciplines. The first course, “Big Data in Education”, was offered by Columbia University in 2013 and delivered on the Coursera platform. In this course, students learned a variety of educational data-mining methods and applied them to research questions related to the design and improvement of interventions within educational software and systems. The course materials consisted of lecture videos, formative in-video quizzes, and 8 weekly assignments. Each weekly assignment was structured as multiple-choice questions or questions requiring a numerical input and, to answer these, students conducted analyses on a provided dataset. In order to receive a grade, students had to complete the assignment within two weeks of its release and, while they were allowed three attempts, only the highest score was counted.

The second course, “Code Yourself”, also delivered on the Coursera platform, was offered by the University of Edinburgh in 2015. The course was designed to introduce teenagers to computer programming, while also covering basic topics in software engineering and computational thinking. The course materials consisted of lectures, videos, formative in-video quizzes, peer-reviewed programming projects, and 5 weekly assignments. Students who scored at least 50% in their coursework were deemed to have passed, while those who scored 75% or more were awarded with a distinction.

Although the two courses had initial enrolments of over 45,000 and 59,900, respectively, a large proportion of these students did not actively participate, and only 18,222 and 26,514 accessed at least a single item in the course; we restrict our study to these students. Of these active students, only 3.5% and 6.0% earned a certificate. This low completion rate can be attributed to a high rate of student attrition – common across online learning environments [32] – which, although significantly attenuated after the first weekly assignment, remained substantial throughout the course. While 17% of students within Code Yourself remained active in the final week of the course, this figure was only 8% for Big Data in Education.

4.2 Model Configuration

The two models compared in this study consist of LDA, following the specification provided by Coleman et al. [9], and an HMM. For the sake of brevity, we omit a comprehensive, technical description of these two models ¹ but, in both cases, a number of details merit discussion.

4.2.1 Representing Interactions with LDA

In its original form, LDA represents a document as a random mixture over latent topics, where each topic is a distribution over a given vocabulary of words [6]. Analogously, Coleman et al. [9] represent a student as a random mixture over *use-cases*, where each use-case is a distribution over the set of course resources. The effort that a student expends on any given resource is represented by their time-on-task, which is calculated by taking the difference, in seconds, between event timestamps; differences over 30 minutes are discarded [9]. Accordingly, each student is represented by a *bag of interactions* where each item represents the total time spent on that resource.

A central feature of the LDA representation is the number of topics, or use-cases, which is a hyper-parameter that must be selected. The standard approach is to use the held-out log-likelihood per interaction [6]. Following the methodology in Coleman et al. [9], a range of use-case models were trained ($K \in [3, 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100]$), and the optimal was selected on the basis of perplexity.

4.2.2 Representing Interactions with HMMs

To account for the temporal dynamics of the data, we model students' interactions on a weekly basis. Specifically, in each week, we assign a student to a discrete latent class which indexes a distribution over the actions that she took. In different weeks, each student could have a different class assignment and thereby have a different distribution that describes her actions. The

¹ For a comprehensive description, we refer interested readers to Appendix A & B.

progression of the state assignment variable then describes how the student interacted with the MOOC platform over time. This addresses the shortcoming of LDA, where user activity is treated as a *bag of interactions*, which ignores the temporal dependence that is inherent to the data.

For instance, suppose we develop a model of student interactions that describes students as being in one of two states: z_1 , which represents a high probability of completing the course materials in the required week; and z_2 , which represents a high probability of not interacting with any materials. Over the duration of the course, the t^{th} student can then be represented as a vector $s^t = [z_1, z_1, z_1, z_2, \dots, z_2]$. Here, the vector describes a student who was actively engaged for the first three weeks of the course, before *transitioning* to a state of disengagement in the fourth week, in which she remained.

To implement such a model, we alter the granularity with which the course resources are indexed. While each resource in the LDA model has a unique identifier, we aggregate them into higher-level classes which represent the three major types of actions that a student can take: watch videos, access in-video quizzes, and access weekly assignments. As we are also interested in the temporal relation of these interactions, each resource class is sub-divided into the resources that were accessed in the correct week, those that were accessed from previous weeks, and those that were accessed from future weeks. Finally, a “no observations” element is added, which accounts for the absence of any action in a given week: if a student interacted with a single resource, this element is set to zero, otherwise it is set to one. Therefore, for each week of a course, a student is represented by a 10 element vector where the first three elements index the three action types in the correct week; the next three elements index the actions where a student accessed resources from previous weeks; the next three elements index the actions where a student accessed resources from future weeks; and finally, the last element corresponds to whether or not the student interacted with the resources at all [1].

Rather than model students’ progression through the weeks of the course with an HMM, we opted for a variant known as a sticky-HMM [17]. The “sticky” assumption in the Markov model is that once a student has adopted a particular state, they persist in that state for as long as possible until a new state is required to describe her actions. Not only does this assumption represent many scenarios in real-world data, where states persist through time [17, 28], but it also helps combat the unrealistically rapid switching dynamics that are present in models without this state-persistence bias [16]. Similar to LDA, where the number of use-cases must be selected, the sticky-HMM requires us to specify the number of states. However, to mitigate the impact this has on our model, we place a non-parametric prior over the state space [17]. The implementation uses a weak-limit approximation to the hierarchical Dirichlet process [46], which approximates the unbounded state space by a truncated representation with L states, where we specify $L = 10$. The prior places diminishing probability mass on infrequent states and thus focuses the majority of the probability mass on a relatively small number of major states

in the model ². For the remainder of this paper, we refer to the sticky-HMM as an HMM.

4.3 Comparing Interpretability

To answer our first research question and evaluate the relative interpretability of the HMM and LDA models, we compare their representations in the light of two criteria: the dimensionality of the representations, and their ability to model the temporal characteristics inherent to the data. Doshi-Velez et al. [12] define “functionally-grounded” evaluations of interpretability as evaluations that requires no humans, but rather use measurable proxies of interpretability. Within the literature, a common proxy is the cardinality of the feature space; for example, one might perform matrix factorization to embed a dataset into a lower-dimensional space which we can then seek to interpret [12]. We therefore compare the dimensionality of the representations that are induced by each model, where lower dimensional representations are assumed to provide more readily digestible information regarding students’ interactions.

Dimensionality, however, is not our sole criterion; we also consider how this observation space spans time. Given that students interact with the MOOC platform over time, the manner with which a model captures these temporal dependencies will affect both the interpretability of the representations, and the validity of the inferences that can be made. Accordingly, we also critique the models on their ability to present information about *when* a student displayed a particular behavioural pattern, not merely *if* they displayed that pattern.

4.4 Predicting Dropout

To address our second research question, we compared the predictive performance of the HMM and LDA models on a dropout prediction task. Following the methodology of Coleman et al. [9], we utilise a SVM for the classification task.

The dropout prediction task was formulated as a binary classification problem, conducted on a weekly basis. For each week, student interactions were taken from the beginning of the course to the end of the given week. In the first of these weekly segments, all students who took no further actions in the course were identified and labeled as the positive class. Then, in each subsequent section of the course, all students from the positive class of the previous section were removed (as these students had already dropped out), and a new positive class was generated, consisting of the students who took their last action in the last week of the current section.

² For a more detailed discussion of the weak-limit approximation and how the approximation becomes exact as L tends to infinity, we refer the reader to Teh et al. [46].

Having computed classification labels for each week of the course, student representations were generated using the previously trained models. In a similar fashion to how labels were generated, interactions were segmented by weeks and, for each week, interactions were taken from the beginning of the course to the end of the given week. Of these, students who had previously dropped-out were removed. For LDA, each student was represented by a vector of the weights associated with their distributions over the use-cases. For the HMM, each student was represented by a vector of the marginal probabilities of being assigned to each state for a given week, where these probabilities for each week were concatenated to form a single vector. The student representations for each of the LDA models and the HMM model were separated into training and test sets, stratified by dropout, which were then used to train and evaluate a set of SVM classifiers; one for each week. This process was conducted separately for each course.

To provide a further benchmark against which to compare the HMM, we compute the accuracy of an omniscient yet naïve classifier: specifically, for each week we take the classification labels and classify all students as belonging to the majority class. We refer to this baseline as the “trivial-baseline”. In evaluating the two models, we calculate a wide range of performance statistics, including accuracy, true negative rate, true positive rate, F_1 -score, AUC, and Cohen’s κ .

4.5 Investigating Generality

To answer our third research question and investigate the extent to which the HMM is capable of generalising across course settings and student populations, we took the trained model from one course and performed the same evaluation described in Section 4.4 on the dropout data from the other course. Specifically, student interaction vectors for the weeks under consideration were converted to their marginal state assignment distributions using the model trained on the opposite course. These distributions were then concatenated into a single vector, and were used to train a SVM (described in Section 4.4), which was evaluated by the same performance metrics.

Problematically, the method detailed in Coleman et al. [9] learns a set of use-cases, or distributions, over the unique resources of a given course, which curtails its cross-course generality. To counter this limitation, we altered the LDA methodology by creating a fixed set of course resources, where the individual resources were first grouped by the week of the course, before being grouped again into the three major action types (videos, in-video quizzes, and assignments). Since our courses were of differing lengths, the fixed vocabulary was created on the larger course (i.e., these resources will have a probability of zero in the shorter course).

We aggregated the resources within our two courses according to this rubric and trained an LDA model on each course. Classification labels and interaction vectors for each week of each course were then generated, as described in

Section 4.4. These interaction vectors for the weeks under consideration were converted to their use-case proportions using the model trained on the opposite course. These proportions were used to train a SVM (described in Section 4.4). We refer to this model as the “baseline”.

5 Results

In this section we review the results of LDA and HMM models with regards to interpretability before comparing them on the task of dropout prediction. Finally, we present the remarkable generality of the HMM representation when used for predicting dropout on a cross-course basis.

5.1 RQ1: Assessing Interpretability

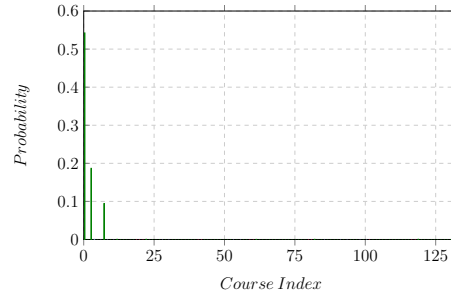
To assess interpretability across the two models under study, we generated plots illustrating their parameters. For the sake of brevity, these plots are only displayed for the larger course under study (Big Data in Education).

Fig. 2 shows the total proportion of all states that any one state comprises. State 1 makes up 50% of all states but states 4, 5, 6, 7 and 9 account for a large portion of the data, making up 43% of all states. While the number of states is a hyper-parameter that we control, we select 10 as the limit. The impact of this decision, however, is mitigated by our model specification: the sticky-HMM assumes that students persist in the current state for as long as possible, while our non-parametric prior places diminishing probability mass on newer states [17]. Since the 6 most common states account for more than 90% of all states adopted in the data, this choice of 10 states is expected to have little effect on our results.

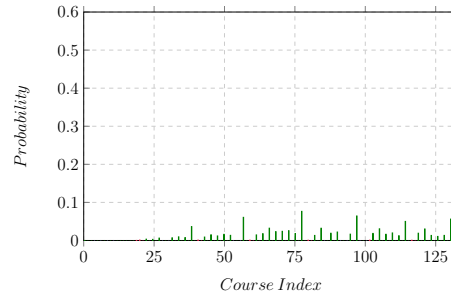
To investigate the interpretability of these state representations individually, we turn to Fig. 4³. Of particular interest are states 6 and 9, which describe participation with resources from the current week and participation with resources from previous weeks, respectively. By contrast, states 1, 4, 5, and 7 describe a complete absence of interactions with a high probability mass on “no observations” and with low probability on the other actions. As such, these states could also be given semantic labels in keeping with those found in the literature: “out” (states 1, 4, 5 and 7), “on-track” (state 6), and “behind” (state 9) [33].

The HMM, however, does not only consist of states, but also the transitions between them, as illustrated by Fig. 3. While it may seem redundant to have four common states describing a complete lack of interaction, the transition matrix reveals that these states represent different trajectories of student behaviour. For instance, state 1 has a negligible probability of transitioning to any other state beyond itself, implying that once a student enters state 1, they are highly likely to never interact with the course again (i.e. dropout). While

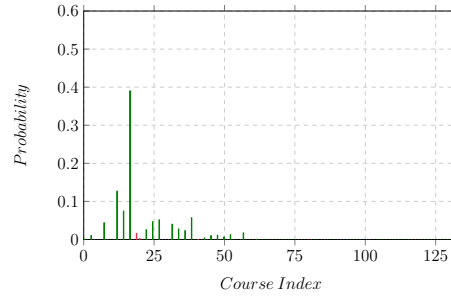
³ Parameters for all states are provided in Appendix C.



(a). Shopping use-case



(b). Completing use-case



(c). Disengaging use-case

Fig. 1. Probability distributions from a 3 use-case model trained on Big Data in Education. Note that videos are represented by green bars, in-video quizzes are represented by blue bars, and assignments are represented by red bars.

state 4 also describes inactivity, it is most likely to transition either to itself, or to state 9. As we saw, state 9 describes interactions with previous weeks' assignments. Accordingly, the transition from a state of inactivity in one week to a state of “catching-up” [33] is an intuitive result.

In the case of LDA, we took a simple model consisting of 3 use-cases and, while this did not fulfil our optimality criterion outlined Section 4.2.1, we nevertheless trained it on the entire set of course interactions. As each use-case defines a distribution over course resources, each plot in Fig. 1 describes the

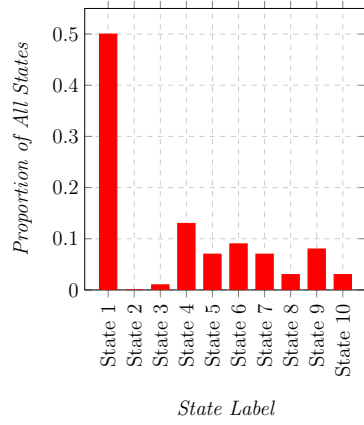


Fig. 2. State count proportions for Big Data in Education.

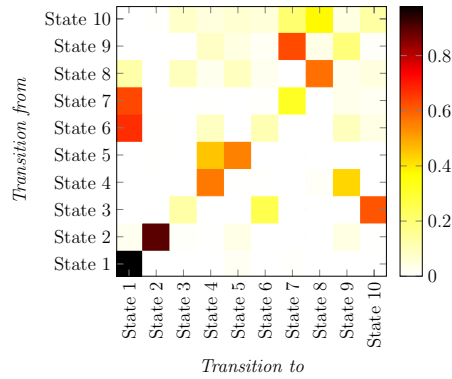


Fig. 3. 10 state HMM transition matrix for Big Data in Education.

probability of interacting with the individual course resources (indexed by the x-axis). The three use-cases document clearly distinct behaviours: in Fig. 1a the probability is concentrated on the first week of lecture content, indicating that a student described by this use-case would likely dropout after this point. In keeping with the previous literature, such students may be described as “sampling” [33] or “shopping” [9]. In the case of Fig. 1b, while there is a limited probability of interacting with any particular item in the course, the consistency of the interactions may be described as “completing” [9, 33]. Finally, Fig. 1c describes when students actively interact with the course, before dropping out or “disengaging” [9, 33].

5.2 RQ2: Dropout Prediction

For each week, a SVM classifier used the student representations generated by the two models to predict dropout in the subsequent week of the course. In

6. TOWARDS INTERPRETABLE INSIGHT

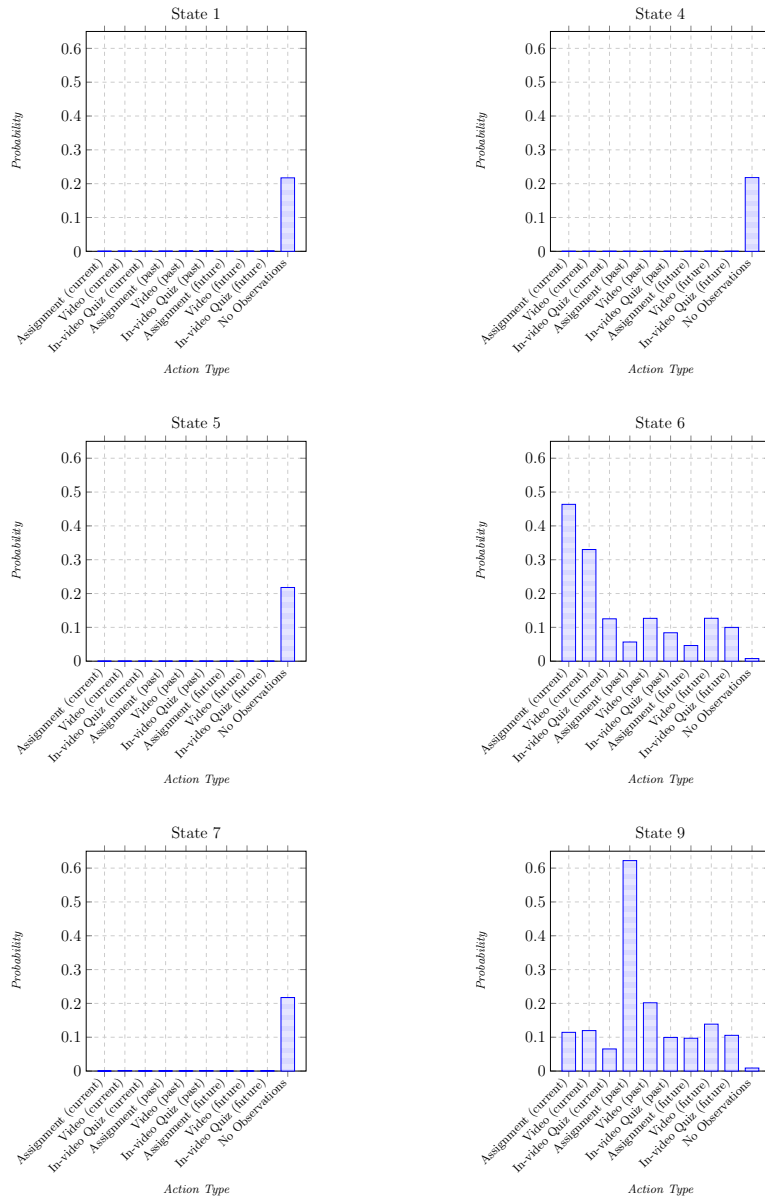
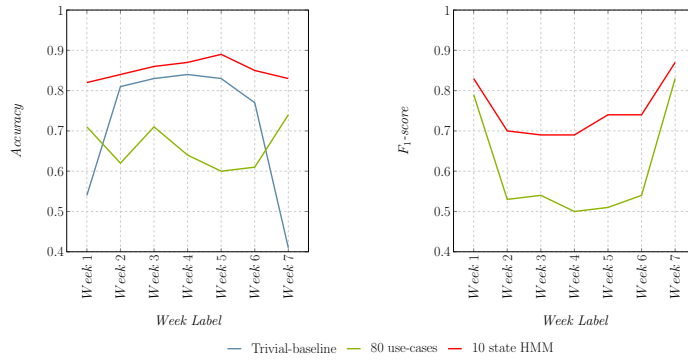
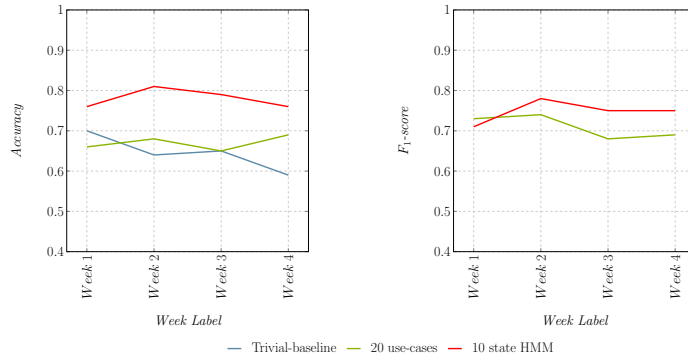


Fig. 4. Probability distributions over action types for the 6 most frequent states within our 10 state HMM, trained on Big Data in Education. Notably, while states 1, 4, 5, and 7 describe a total lack of interaction with course materials, the transition matrix indicates that, of these four, only students in state 1 transition to almost exclusively the same inactive state across time periods.

the case of LDA, the 80 use-case and the 20 use-case models were selected as optimal (following Section 4.2.1) for Big Data in Education and Code Yourself,



(a). Big Data in Education

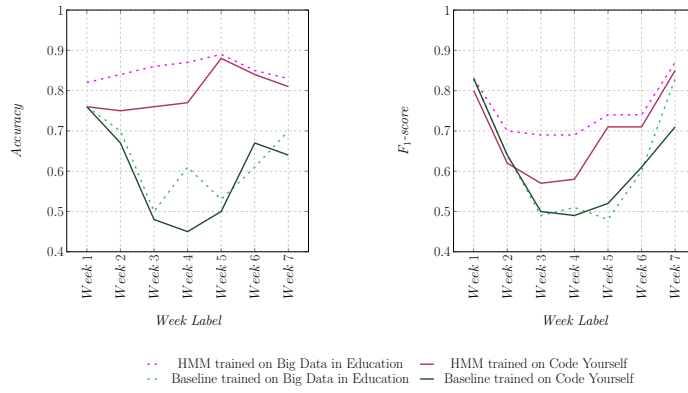


(b). Code Yourself

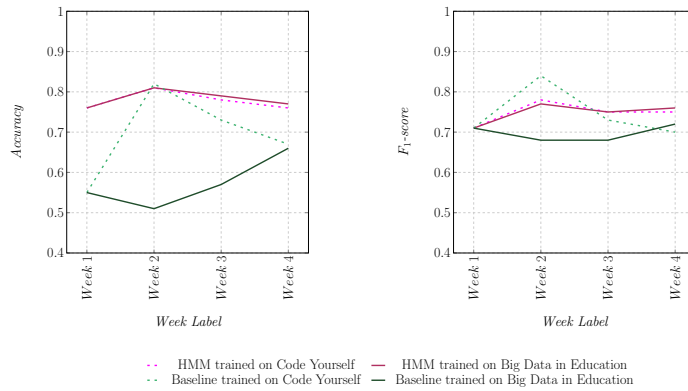
Fig. 5. Comparison of the LDA and HMM models on the task of predicting student disengagement. For each course, two models are trained: an LDA model with the optimal number of use-cases, and a 10 state HMM. In addition, the accuracy plots include the performance of our trivial-baseline classifier.

respectively. Both courses utilised an HMM with a weak limit cut-off of 10 states.

Fig. 5 presents the overall accuracy and F_1 -scores of the two models. In the case of Big Data in Education, the HMM outperforms both the trivial-baseline and the LDA model in all weeks of the course. Notably, the LDA model underperforms the trivial-baseline except for the first and last weeks. A similar result is found in Code Yourself: the HMM outperforms all alternative models (except for week 1), but the LDA model does not compare favourably to the trivial-baseline until the very last week of the course. This echoes the results of Coleman et al. [9], who found that the trivial classifier outperformed the LDA classifier in the prediction of student certification. For tables documenting a wider range of performance statistics, including AUC and Cohen's κ , please refer to Appendix D.



(a). Big Data in Education



(b). Code Yourself

Fig. 6. Assessing the generality of the HMM and baseline methodologies. Models were trained on each of the two courses, before being evaluated on the opposite course. To make the generality comparison explicit, the dotted lines in each plot represent the predictive performance of the test set course, trained and tested on itself (also displayed in Fig. 5)

5.3 RQ3: Model Generality

Cross-course performance was compared to the performance of models trained and tested on the same course (indicated with dotted lines in Fig. 6). In the case of the HMM (purple lines), the model’s cross-course performance suffered somewhat when trained on the shorter course (see Fig. 6a). However, when trained on the longer course the cross-course performance was almost identical (see Fig. 6b). This stands in stark contrast to the cross-course performance of the baseline model, which performed comparably when trained on the shorter course (see Fig. 6a), but substantially underperformed when trained on the longer course (see Fig. 6b). Nevertheless, across all weeks, the HMM either outperformed or was indistinguishable from the baseline model.

For tables documenting a wider range of performance statistics, including AUC and Cohen’s κ , please refer to Appendix E.

6 Discussion

In this section, we review our results in the light of our three research questions and emphasise that not only does our approach provide interpretable representations of student interactions, but also our model captures salient information regarding student dropout, and is capable of generalising across course-contexts.

6.1 RQ1: Comparing Model Interpretability

Our discussion of model interpretability revolves around the two criteria identified in Section 3.1. The first of these, dimensionality, we treat as a direct proxy of explanation quality [12]. The rationale for this is that a lower dimensional representation provides more readily digestible information to an end user. This argument finds considerable support within the machine learning literature, where dimensionality reduction techniques have received ample attention [11, 36]. Accordingly, if we compare the observation space of the two models, we find that for LDA, the dimensionality is equal to the number of unique course resources (135 and 150 for the data used in the present study). While this will vary from course to course, it is highly likely to exceed the 10 dimensions we specify for our HMM.

In addition to comparing the models with respect to their dimensionality, we also consider how they account for time. In the case of the HMM, the model learns a set of general state descriptions that capture global behaviours which a student transitions between for the duration of their course. Time is instead captured by transitions within the state space. By contrast, LDA makes the *bag of interactions* assumption which ignores when actions takes place. Without temporal information, the validity of any inferences we make cannot be assured. For example, consider a student who only interacts with the current week’s resources. In this case, LDA would describe the student by an equal probability placed across the entire set of course resources. The problem with this is that the model is incapable of distinguishing between this student and another who interacts with the same resources, but does so only in the first week of the course. Our HMM, however, would model the first student as being “on-track” whereas the second student would be “ahead” in the first week, and “out” thereafter.

Model interpretability has become increasingly prominent within the educational technology literature, and researchers have argued that machine learning models are not themselves a panacea, but must be combined with interpretable insight into learners behaviour and the learning process [43]. These concerns are echoed in the machine learning literature, where it is broadly accepted that a machine learning algorithm should be able to provide an account

of what decisions it made [12, 21, 44, 52]. For the HMM, the limited state space and the explicit modeling of transitions between these states are important advantages that facilitate this clarity. This is particularly apparent in the context of querying the model. For instance, for a given student we may ask “what is the probability of dropping out *in the next week of interaction?*”. Being able to point to the student’s prior state and the associated transition probabilities governing that state marks an important step towards transparency.

6.2 RQ2: Preempting Student Dropout

In predicting dropout, we are not concerned with hyper-parameter tuning or comparing classifiers, but rather whether or not our student representation captures salient statistical information regarding our construct of interest. That is, when a student’s behaviour is indicative of dropout.

In reviewing the results displayed in Fig. 5, the representation provided by the HMM achieves consistent, strong performance across all weeks, particularly in the case of Code Yourself (see Fig. 5a). In the case of Big Data in Education, although the F_1 -scores drop after the first week, the loss in performance is not severe and is recuperated over the remainder of the course. The consistency of these results provides reasonable evidence to state that our HMM generates valid representations of student dropout.

By contrast, the representations generated by the LDA models result in divergent performance across the two courses. As these two models were trained using different numbers of use-cases, a direct comparison is not possible. What is notable, however, is that even with an 80-dimensional student representation (Big Data in Education), the predictive performance is considerably weaker than the 10 state HMM. Furthermore, the inconsistency of these results, whereby the F_1 -scores are high in the first and last weeks, but collapse in all others, calls into question the validity of this predictive model.

The underperformance of the HMM in week one of Code Yourself (see Fig. 5b) could be attributed to differences in the cardinality of the state space. The dimensionality of the LDA representation is twice that of the HMM and, since it is only the first week, the temporal advantages of the HMM do not come into play.

6.3 RQ3: Identifying Global Behaviours

The cross-course performance, displayed in Fig. 6, provides substantial evidence that the states captured by the HMM describe global behaviours that generalise across course-contexts. When trained on the shorter course and tested on the longer (as in Fig. 6a), there is some performance loss in the first five weeks, compared to the model trained and tested on the longer course. In the opposite case, however, when we train on the longer course and test on the shorter (as in Fig. 6b), the generality is remarkable and the performance

is indistinguishable from the original model trained and tested on the shorter course. Although the courses contain substantially different content, the extent of the generality that our HMM achieves warrants further investigation on a larger sample of courses. For the present, however, our results indicate that the states described by the HMM are both accurate and general predictors of student dropout across contexts.

This conclusion is especially clear when comparing the performance of the HMM to the LDA baseline model. In particular, when trained on Big Data in Education and tested on Code Yourself (see Fig. 6b), the baseline performance is substantially curtailed compared to the model trained and tested on Code Yourself. Furthermore, apart from the first few weeks of the larger course, the performance of the baseline is considerably lower than the HMM.

7 Limitations and Future Work

In arguing for the interpretability of our HMM representation, we follow the taxonomy proposed by Doshi-Velez et al. [12]. In doing so, however, we rely on functionally-grounded evaluations, which use commonly accepted proxies for interpretability such as lower dimensionality of the observation space. While this is appropriate for the present, machine learning-focused study, in future work it is essential to engage domain experts such as course instructors.

While we find that our model provides valid, general representations of student dropout that may, in future work, be used to identify students at-risk, our analysis is only conducted on two courses. Although these courses represent distinct disciplines and pedagogies, further validation is required to ensure that factors such as course design do not unduly bias our state parameters. We therefore propose a larger study that would include more courses from a wider variety of disciplines.

In this paper we present an argument for a temporal state-based representation of student behaviour. We propose that a student can transition between a small number of individually interpretable states. While the state space presents a high level overview of the general activities that characterise students' activity, the individual state transitions permit a description of the diverse motivations that are prevalent within MOOCs. Although we chose to implement this representation with an HMM, this is not the only model that could facilitate such a design. In particular, the Beta process (Indian Buffet process) describes an approach where the state transitions are coupled [47]. This means the model explicitly learns popular trajectories through time, and not merely popular states that the students might persist in [42]. In future work, we intend to compare and contrast this approach with that which is presented in this study.

Finally, our focus on validating the state representations caused us to overlook certain details relating to the HMM. In particular, we did not engage with hyper-parameter tuning to select the best HMM from a set of candidate models. Instead, we used Bayesian non-parametrics to offset the cost of not

exploring the cardinality of the state space. Nevertheless, these models still have parameters that would require tuning to both maximise the likelihood of the model, and to maximize the interpretability of the representations that are induced. There is a trade-off between interpretability and the accuracy of the representations that are learnt, and this trade-off needs exploring in the context of MOOC data [24]. While these limitations highlight important shortcomings of the present work, it is worth noting that they largely relate to important future directions for this research programme which are beyond the purview of this initial, methodological, study.

8 Conclusions

In developing an interpretable model of student behaviour, the present study makes several contributions towards the literature, and identifies a number of interesting avenues of future research. In particular, we propose a model that learns interpretable representations of students' behaviour. These representations not only provide a high level descriptions of typical interaction patterns, but also accommodate for the diversity of students' motivations. We evaluate the model on a common dropout prediction task and argue that, compared to the previous literature, our approach provides more interpretable insight into *how* students dropped out. While the prediction task demonstrates the structural validity of these representations, we also demonstrate their generality through conducting the same analysis on a cross-course basis.

Following on these results, we propose that future research should work with MOOC instructors to present common student trajectories, derived from the model, which gather insight about how students engage with the course. This can be done by (1) using the HMM to learn a set of latent states about the students' interactions; (2) sourcing expert insight (MOOC instructors) to apply semantic labels to these states (which may require additional visualisations of the state representations); and (3) presenting common trajectories through this state space to the instructors, in order to identify insightful behaviours present in the students' activity.

This is an important result for the educational literature, as it demonstrates how learner models can move beyond predictive accuracy to additionally offer interpretable insight into learners' behaviour and the learning process [43]. Our results also work towards the transferability challenge [2], wherein we apply the model learnt from one course to represent student behaviours in an entirely different course. The generality of this result suggests that global interaction behaviours are present across different disciplines and can be harnessed by models to provide interpretable insight even in lieu of course specific training data.

Acknowledgements We would like to thank Ryan Baker for his invaluable advice and guidance in the development of this research.

Conflict of interest

The authors declare that they have no conflict of interest.

References

1. Anderson, A., Huttenlocher, D., Kleinberg, J., Leskovec, J.: Steering user behavior with badges. In: Proceedings of the 22nd international conference on World Wide Web, pp. 95–106. ACM (2013)
2. Baker, R.S.: Challenges for the future of educational data mining: The baker learning analytics prizes. *JEDM— Journal of Educational Data Mining* **11**(1), 1–17 (2019)
3. Balakrishnan, G.: Predicting student retention in massive open online courses using hidden markov models. Master’s thesis, EECS Department, University of California, Berkeley (2013). URL <http://www2.eecs.berkeley.edu/Pubs/TechRpts/2013/EECS-2013-109.html>
4. Beal, M.J., Ghahramani, Z., Rasmussen, C.E.: The infinite hidden markov model. In: Proceedings of the 14th International Conference on Neural Information Processing Systems: Natural and Synthetic, NIPS’01, pp. 577–584. MIT Press, Cambridge, MA, USA (2001). URL <http://dl.acm.org/citation.cfm?id=2980539.2980615>
5. Bergner, Y., Kerr, D., Pritchard, D.E.: Methodological challenges in the analysis of mooc data for exploring the relationship between discussion forum views and learning outcomes. In: Paper presented at the proceedings of the 8th International Conference on Educational Data Mining (EDM’15). International Educational Data Mining Society (2015). URL http://www.educationaldatamining.org/EDM2015/uploads/papers/paper_61.pdf
6. Blei, D.M., Ng, A.Y., Jordan, M.I.: Latent dirichlet allocation. *Journal of Machine Learning Research* **3**, 993–1022 (2003). URL <http://dl.acm.org/citation.cfm?id=944919.944937>
7. Chang, J., Boyd-Graber, J., Gerrish, S., Wang, C., Blei, D.M.: Reading tea leaves: How humans interpret topic models. In: Proceedings of the 22nd International Conference on Neural Information Processing Systems, NIPS’09, pp. 288–296. Curran Associates Inc., USA (2009). URL <http://dl.acm.org/citation.cfm?id=2984093.2984126>
8. Coffrin, C., Corrin, L., de Barba, P., Kennedy, G.: Visualizing patterns of student engagement and performance in moocs. In: Proceedings of the Fourth International Conference on Learning Analytics And Knowledge, LAK ’14, pp. 83–92. ACM, New York, NY, USA (2014). DOI 10.1145/2567574.2567586. URL <http://doi.acm.org/10.1145/2567574.2567586>
9. Coleman, C.A., Seaton, D.T., Chuang, I.: Probabilistic use cases: Discovering behavioral patterns for predicting certification. In: Proceedings of the Second (2015) ACM Conference on Learning @ Scale, L@S ’15, pp. 141–148. ACM, New York, NY, USA (2015). DOI 10.1145/2724660.2724662. URL <http://doi.acm.org/10.1145/2724660.2724662>
10. Crossley, S., McNamara, D.S., Baker, R., Wang, Y., Paquette, L., Barnes, T., Bergner, Y.: Language to completion: Success in an educational data mining massive open online class. International Educational Data Mining Society (2015)
11. Cui, P., Wang, X., Pei, J., Zhu, W.: A survey on network embedding. *IEEE Transactions on Knowledge and Data Engineering* (2017). DOI 10.1109/TKDE.2018.2849727. URL <https://dx.doi.org/10.1109/TKDE.2018.2849727>
12. Doshi-Velez, F., Kim, B.: Towards a rigorous science of interpretable machine learning. arXiv preprint arXiv:1702.08608 (2017). URL <https://arxiv.org/abs/1702.08608>
13. Fei, M., Yeung, D.Y.: Temporal models for predicting student dropout in massive open online courses. In: Proceedings of the 2015 IEEE International Conference on Data Mining Workshop (ICDMW), ICDMW ’15, pp. 256–263. IEEE Computer Society (2015). DOI 10.1109/ICDMW.2015.174. URL <http://dx.doi.org/10.1109/ICDMW.2015.174>
14. Fincham, E., Gašević, D., Jovanović, J., Pardo, A.: From Study Tactics to Learning Strategies: An Analytical Method for Extracting Interpretable Representations. *IEEE Transactions on Learning Technologies* **12**(1), 59–72 (2019). DOI 10.1109/TLT.2018.2823317. URL <https://ieeexplore.ieee.org/document/8331946>

15. Fincham, E., Whitelock-Wainwright, A., Kovanović, V., Joksimović, S., van Staalduinen, J.P., Gašević, D.: Counting clicks is not enough: Validating a theorized model of engagement in learning analytics. In: Proceedings of the 9th International Conference on Learning Analytics & Knowledge, LAK19, pp. 501–510. ACM, New York, NY, USA (2019). DOI 10.1145/3303772.3303775. URL <http://doi.acm.org/10.1145/3303772.3303775>
16. Fox, E.B., Sudderth, E.B., Jordan, M.I., Willsky, A.S.: An hdp-hmm for systems with state persistence. In: Proceedings of the 25th international conference on Machine learning, pp. 312–319. ACM (2008)
17. Fox, E.B., Sudderth, E.B., Jordan, M.I., Willsky, A.S., et al.: A sticky hdp-hmm with application to speaker diarization. *The Annals of Applied Statistics* **5**(2A), 1020–1056 (2011)
18. Gardner, J., Brooks, C.: Evaluating predictive models of student success: Closing the methodological gap. arXiv preprint arXiv:1801.08494 (2018). URL <https://arxiv.org/abs/1801.08494>
19. Gašević, D., Dawson, S., Rogers, T., Gašević, D.: Learning analytics should not promote one size fits all: The effects of instructional conditions in predicting academic success. *The Internet and Higher Education* **28**, 68–84 (2016)
20. Gillani, N., Eynon, R.: Communication patterns in massively open online courses. *The Internet and Higher Education* **23**, 18 – 26 (2014). DOI <https://doi.org/10.1016/j.iheduc.2014.05.004>. URL <http://www.sciencedirect.com/science/article/pii/S1096751614000438>
21. Gilpin, L., Bau, D., Z. Yuan, B., Bajwa, A., Specter, M., Kagal, L.: Explaining explanations: An overview of interpretability of machine learning. In: Proceedings of the 5th IEEE International Conference on Data Science and Advanced Analytics (DSAA), pp. 80–89 (2018). DOI 10.1109/DSAA.2018.00018
22. Gütl, C., Rizzardini Rocael Hernández and Chang, V., Morales, M.: Attrition in mooc: Lessons learned from drop-out students. In: L. Uden, J. Sinclair, Y.H. Tao, D. Liberona (eds.) *Learning Technology for Education in Cloud, MOOC and Big Data*, pp. 37–48. Springer International Publishing, Cham (2014)
23. Herskovic, H., Vilenchik, D., Gal, Y.: Modeling engagement in self-directed learning systems using principal component analysis. *IEEE Transactions on Learning Technologies* pp. 1–1 (2019). DOI 10.1109/TLT.2019.2922902
24. Hoernle, N., Gal, K., Grosz, B., Lyons, L., Ren, A., Rubin, A.: Interpretable models of human interaction in immersive simulation settings. arXiv preprint arXiv:1909.11025 (2019). URL <https://arxiv.org/abs/1909.11025>
25. Hoffman, M.D., Blei, D.M., Bach, F.: Online learning for latent dirichlet allocation. In: Proceedings of the 23rd International Conference on Neural Information Processing Systems - Volume 1, NIPS'10, pp. 856–864. Curran Associates Inc., USA (2010). URL <http://dl.acm.org/citation.cfm?id=2997189.2997285>
26. Ishwaran, H., Zarepour, M.: Markov chain monte carlo in approximate dirichlet and beta two-parameter process hierarchical models. *Biometrika* **87**(2), 371–390 (2000)
27. Jiang, S., Williams, A.E., Schenke, K., Warschauer, M., O'Dowd, D.K.: Predicting mooc performance with week 1 behavior. In: Proceedings of the 7th International Conference on Educational Data Mining (EDM'14), pp. 273–275. International Educational Data Mining Society (2014)
28. Johnson, M.J., Willsky, A.S.: Bayesian nonparametric hidden semi-markov models. *Journal of Machine Learning Research* **14**(1), 673–701 (2013). URL <http://dl.acm.org/citation.cfm?id=2567709.2502602>
29. Joksimović, S., Poquet, O., Kovanović, V., Dowell, N., Mills, C., Gašević, D., Dawson, S., Graesser, A.C., Brooks, C.: How do we model learning at scale? a systematic review of research on moocs. *Review of Educational Research* **88**(1), 43–86 (2018). DOI 10.3102/0034654317740335
30. Jovanović, J., Gašević, D., Dawson, S., Pardo, A., Mirriahi, N.: Learning analytics to unveil learning strategies in a flipped classroom. *The Internet and Higher Education* **33**, 74–85 (2017)
31. Kennedy, G., Coffrin, C., de Barba, P., Corrin, L.: Predicting success: How learners' prior knowledge, skills and activities predict mooc performance. In: Proceedings of

- the Fifth International Conference on Learning Analytics And Knowledge, LAK '15, pp. 136–140. ACM, New York, NY, USA (2015). DOI 10.1145/2723576.2723593. URL <http://doi.acm.org/10.1145/2723576.2723593>
32. Kizilcec, R., Halawa, S.: Attrition and achievement gaps in online learning. In: Proceedings of the Second (2015) ACM Conference on Learning @ Scale, L@S '15, pp. 57–66. ACM (2015). DOI 10.1145/2724660.2724680. URL <http://doi.acm.org/10.1145/2724660.2724680>
 33. Kizilcec, R., Piech, C., Schneider, E.: Deconstructing disengagement: Analyzing learner subpopulations in massive open online courses. In: Proceedings of the Third International Conference on Learning Analytics & Knowledge, LAK '13, pp. 170–179. ACM (2013)
 34. Kovanović, V., Gašević, D., Dawson, S., Joksimović, S., Baker, R.: Does time-on-task estimation matter? implications on validity of learning analytics findings. *Journal of Learning Analytics* **2**(3), 81–110 (2016). DOI 10.18608/jla.2015.23.6. URL <http://www.learning-analytics.info/journals/index.php/JLA/article/view/4501>
 35. Kovanović, V., Joksimović, S., Poquet, O., Hennis, T., de Vries, P., Hatala, M., Dawson, S., Siemens, G., Gašević, D.: Examining communities of inquiry in massive open online courses: The role of study strategies. *The Internet and Higher Education* **40**, 20 – 43 (2019). DOI <https://doi.org/10.1016/j.iheduc.2018.09.001>
 36. Lever, J., Krzywinski, M., Altman, N.: Points of significance: Principal component analysis. *Nature Methods* (2017). DOI 10.1038/nmeth.4346. URL <https://doi.org/10.1038/nmeth.4346>
 37. Loevinger, J.: Objective tests as instruments of psychological theory. *Psychological Reports* **3**(3), 635–694 (1957). DOI 10.2466/pr0.1957.3.3.635. URL <https://doi.org/10.2466/pr0.1957.3.3.635>
 38. Matcha, W., Gašević, D., Uzir, N.A., Jovanović, J., Pardo, A.: Analytics of learning strategies: Associations with academic performance and feedback. In: Proceedings of the 9th International Conference on Learning Analytics & Knowledge, LAK19, pp. 461–470. ACM (2019). DOI 10.1145/3303772.3303787. URL <http://doi.acm.org/10.1145/3303772.3303787>
 39. Matcha, W., Gašević, D., Uzir, N.A., Jovanović, J., Pardo, A., Lim, L., Maldonado-Mahauad, J., Gentili, S., Perez-Sanagustín, M., Tsai, Y.S.: Analytics of learning strategies: Role of course design and delivery modality. *Journal of Learning Analytics* (in press)
 40. Messick, S.: Standards of validity and the validity of standards in performance assessment. *Educational Measurement: Issues and Practice* pp. 5–8 (1995)
 41. Řehůřek, R., Sojka, P.: Software Framework for Topic Modelling with Large Corpora. In: Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks, pp. 45–50. ELRA (2010). <http://is.muni.cz/publication/884893/en>
 42. Ren, Y., Fox, E.B., Bruce, A., et al.: Clustering correlated, sparse data streams to estimate a localized housing price index. *The Annals of Applied Statistics* **11**(2), 808–839 (2017)
 43. Rosé, C., McLaughlin, E., Liu, R., Koedinger, K.: Explanatory learner models: Why machine learning (alone) is not the answer: Explanatory learner models. *British Journal of Educational Technology* (2019). DOI 10.1111/bjet.12858
 44. Rosenfeld, A., Richardson, A.: Explainability in human-agent systems. *Autonomous Agents and Multi-Agent Systems* pp. 1–33 (2019)
 45. Sharma, K., Jermann, P., Dillenbourg, P.: Identifying styles and path towards success in moocs. In: Proceedings of the 8th International Conference on Educational Data Mining (EDM'15), pp. 408–411. International Educational Data Mining Society (2015)
 46. Teh, Y.W., Jordan, M.I., Beal, M.J., Blei, D.M.: Sharing clusters among related groups: Hierarchical dirichlet processes. In: *Advances in neural information processing systems*, pp. 1385–1392 (2005)
 47. Thibaux, R., Jordan, M.I.: Hierarchical beta processes and the indian buffet process. In: *Artificial Intelligence and Statistics*, pp. 564–571 (2007)
 48. Weinstein, C.E., Husman, J., Dierking, D.: Chapter 22. Self-Regulation Interventions with a Focus on Learning Strategies. *Handbook of Self-regulation* (2012)

49. Whitehill, J., Mohan, K., Seaton, D., Rosen, Y., Tingley, D.: Mooc dropout prediction: How to measure accuracy? In: Proceedings of the Fourth (2017) ACM Conference on Learning @ Scale, L@S '17, pp. 161–164. ACM, New York, NY, USA (2017). DOI 10.1145/3051457.3053974. URL <http://doi.acm.org/10.1145/3051457.3053974>
50. Whitehill, J., Williams, J., Lopez, G., Austun Coleman, C., Reich, J.: Beyond prediction: First steps toward automatic intervention in mooc student stopout. In: Proceedings of the 8th International Conference on Educational Data Mining (EDM'15), pp. 171–178 (2015)
51. Winne, P., Hadwin, A.: Studying as self-regulated learning. *Metacognition in Educational Theory and Practice* **93**, 27–30 (1998)
52. Wu, M., Hughes, M.C., Parbhoo, S., Zazzi, M., Roth, V., Doshi-Velez, F.: Beyond sparsity: Tree regularization of deep models for interpretability. In: Thirty-Second AAAI Conference on Artificial Intelligence (2018)

A LDA Technical Description

The data⁴ consist of T registered students who each access a fixed set of course resources. A *course resource* is the basic unit of discrete data and is defined as an item from a set of C possible course resources that are available where $m_i, i \in [1, \dots, C]$ is the i^{th} resource in the course. While each student $s^t, t \in [1, \dots, T]$, can access the same set of course resources, they may do so at different times and in different proportions. For example, student s^i might access a particular video resource, m_k , on numerous occasions, whereas student s^j may instead choose to frequent a particular set of lecture notes, m_l . Of course, there may also be some mixture of these two extremes. We represent the t^{th} student, s^t , as a sequence⁵ of n^t course resources where m_i^t denotes that student s^t accessed resource m_i : $s^t = (m_1^t, \dots, m_{n^t}^t)$. \mathcal{S} is the set of all student sequences.

LDA represents a document as a random mixture over latent topics, where each topic is a distribution over the vocabulary of words [6]. Similarly, Coleman et al. [9] represent a student as a random mixture over *use-cases*, where each use-case is a distribution over the set of course resources. To represent the relative amount of effort that a student spends accessing a certain resource, the resource access count is calculated by the inferred number of seconds that the student spends interacting with the specific resource. The time is calculated by taking the difference, in seconds, between event timestamps. Differences over 30 minutes are discarded. Accordingly, each student is represented by a bag of resource interactions where each item represents the total time spent on that resource.

Finally, we introduce the parameters that control the use-case distributions over course resources and the student-specific distributions over use-cases. Student s^t has the probability over use-cases defined by ϕ^t and we let Φ denote the set of all T student – use-case probability vectors. Similarly, given a total of K use-cases, $\theta_k, k \in [1, \dots, K]$ is the k^{th} use-case probability over course resources⁶. Θ is the $K \times C$ matrix of all θ_k vectors.

The joint distribution of interest is the posterior of the parameters Θ and Φ given the observed course interactions for all students, as described in Equation (1).

$$P(\Theta, \Phi | \mathcal{S}) \quad (1)$$

To enable inference over the latent parameters in the model, LDA makes a number of independence assumptions. First, each student’s use-case proportions, ϕ^t , are independent given a global shared prior β . Second, each access of a course resource, m_i^t , is assigned a use-case assignment, z_i^t , that is an independent categorical draw from the student’s use-case distribution, given proportions ϕ^t . The course resource access is then itself a categorical draw from the distribution with proportions θ_k indexed by the value of $z_i^t = k$. The generative description for the data, and the implied conditional independencies, is shown by the graphical model in Fig. 7 and is delineated below:

1. For $k \in [1, \dots, K]$, sample the use-case specific proportion over course resources: $\phi_k \sim \text{Dirichlet}(\beta)$.
2. For $t \in [1, \dots, T]$, sample the student specific proportion over use-cases: $\theta^t \sim \text{Dirichlet}(\alpha)$.
3. For each course resource access belonging to a specific student, sample the use-case assignment: $z_i^t \sim \text{Categorical}(\theta^t)$.
4. For each course resource access belonging to a specific student, given the use-case assignment, sample the course-resource access from the corresponding use-case proportion: $m_i^t \sim \text{Categorical}(\phi_{z_i^t})$

The structure of the model, and the independence assumptions enable us to factorise the joint probability described in Equation (1) into a form that, whilst remaining intractable,

⁴ We follow the same notation used by Coleman et al. [9]

⁵ We deviate from the n_t notation of [9] for consistency with s^t and m^t referring to the t^{th} student and her resources.

⁶ Coleman et al. [9] refer to this θ_k parameter as $\beta_{\hat{a}}$. We again deviate from their notation in conforming with the more widely adopted notation of θ_k , leaving β for a hyper-parameter in the model [6].

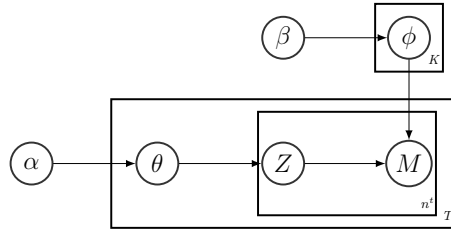


Fig. 7. Probabilistic graphical model of LDA

is amenable to approximate posterior inference using Gibbs sampling or variational Bayes. Following [9], we use the online variational Bayes algorithm outlined by Hoffman et al. [25] and implemented by the python package gensim [41]. This factored representation of the joint probability is described by Equation (2).

$$p(\Theta, \Phi | S, \alpha, \beta) \propto \prod_{t=1}^T p(\theta^t | \alpha) \times \prod_{k=1}^K p(\phi_k | \beta) \times \prod_{t=1}^T \prod_{i=1}^{n^t} p(m_i^t | \theta^t, \Phi) \quad (2)$$

Note that the model includes a hyper-parameter K that dictates the number of use-cases in the model. The standard approach for model selection is to use the log held-out perplexity per interaction [6]. This metric is, the negative log likelihood of a held-out test corpus, divided by the number of interactions within that corpus ($\sum_{corpus} n^t$), as in Equation (3).

$$\log\{perplexity(corpus)\} = \frac{-\log\{P(corpus | \alpha, \beta)\}}{\sum_{corpus} n^t} \quad (3)$$

B HMM Technical Description

Each student’s weekly progression through the course was modeled with a sticky hidden Markov model [17]. In the hidden Markov model, the sticky assumption is that once a student has adopted a state, the student persists in that state for as long as is possible until a new state is needed to describe zero actions. Not only does this assumption represent many scenarios in real-world data where states persist through time [17, 28], but it also helps with the identifiability of the temporal dynamics of the model by combating the unrealistically rapid switching dynamics that are present in models without this state-persistence bias [4, 46].

Moreover, as the number of states is latent and unknown, we use the hierarchical Dirichlet process (HDP) hidden Markov model (HMM). Here, we use the weak-limit approximation to the HDP [16, 28] that approximates the infinite state-space with a truncation of L states. While the limit of the number of states still needs to be chosen, as in LDA, the effect of this choice is reduced, as the HDP prior places diminishing probability mass on newer states and therefore they are adopted infrequently. Rather, a smaller number of more important states dominate the student-to-state assignments and it is these frequently accessed states whose interpretability we assess. Theoretically, as L is increased, the truncated approximation to the true HDP posterior is guaranteed to become exact [26].

More formally, a student consists of a latent state sequence, $z = [z_1, \dots, z_T]$, where T is the number of weeks in the course in our data. Each $z_t \in \{1, \dots, L\}$ is a discrete state that indexes a particular emission distribution over the observed actions (analogous to the state assignment variable, z , in Fig. 7 which indexes a distribution over course resource items). Since, for a particular week within the course, a student is in a particular state and the state indexes the probability distribution over the actions that a student may take, we can interpret the emission distribution of the HMM as a *policy* for that student in that week. A student may transition from state z_i to state z_j with probability at row i and element j in the transition matrix $\pi = (\pi_{ij})_{i,j=1}^L$; thus: $\pi_{ij} = p(z_{t+1} = j | z_t = i)$. The policy (or

emission distribution) for the actions in state i is therefore $p(y_t|z_t, \theta_i)$ where θ_i refers to the parameters that control this emission distribution. In the sticky-HMM model, when a student adopts a state, the sticky parameter inflates the self-transition probability and therefore the student persists in that state where possible.

The graphical model for the hidden Markov model is shown in Fig. 8. The state variable indexes a distribution over actions for an entire week rather than making the assumption that each course-resource access may be assigned a state independently. We highlight the significant reduction in parameters in the model: the transition matrix from the hidden Markov model requires L^2 parameters and a student is assigned to one of L states for each week of work. This is in comparison to the student – use-case matrix (Θ) which requires $T \times C \gg L \times T$ parameters. As the distribution over course resource items has now been generalized to include the type of action rather than the specific index, the state space of the use-case distribution has been reduced from C to L . The result of this reduction in latent space will aid in the interpretability of the model while the temporal evolution of each student through the state space allows the model to maintain the complexity that is needed to enable predictive accuracy on the tasks proposed in our research questions.

It is important to note that the gains in reduction of model parameters are countered by the increased difficulty in inference. While the factored representation of LDA allows efficient batch computation, the HMM approach requires execution of the forward-backward algorithm to evaluate the likelihood of trajectories across the entire course. Our inference implementation uses the `pyhsmm`⁷ python package which follows the collapsed Gibbs sampling technique outlined by Fox et al. [16]. The python implementation, created by the authors of [28], uses an efficient routine for the forward-backward recursion, written in C.

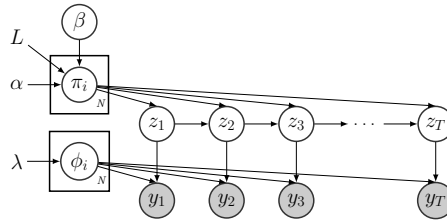


Fig. 8. Hidden Markov model showing the transition probability over adopting a state (π), the state-specific distribution over course resources (ϕ), the state assignments (z) and the observations where students access course resource items

⁷ <https://github.com/mattjj/pyhsmm>

C HMM State Parameters

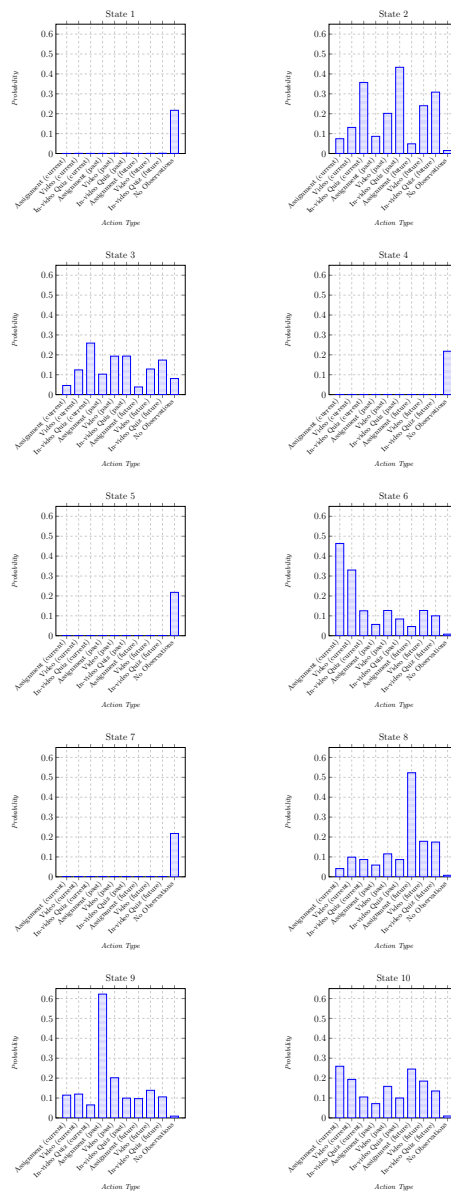


Fig. 9. Probability distributions over action types for each state in the HMM.

D Dropout Prediction Results**Table 1**

Performance of the HMM at Predicting Dropout Week (Big Data in Education)

Weeks	10 state model					Cohen's κ
	ACC	TNR	TPR	F1	AUC	
1-1	0.83	0.73	0.94	0.83	0.83	0.65
1-2	0.82	0.84	0.74	0.62	0.88	0.60
1-3	0.84	0.86	0.71	0.59	0.89	0.61
1-4	0.85	0.86	0.75	0.61	0.89	0.62
1-5	0.86	0.90	0.69	0.64	0.89	0.67
1-6	0.84	0.87	0.73	0.68	0.88	0.65
1-7	0.78	0.73	0.82	0.81	0.81	0.65

Table 2

Performance of the HMM at Predicting Dropout Week (Code Yourself)

Weeks	10 state model					Cohen's κ
	ACC	TNR	TPR	F1	AUC	
1-1	0.76	0.66	0.10	0.71	0.84	0.59
1-2	0.81	0.75	0.92	0.78	0.83	0.61
1-3	0.78	0.71	0.92	0.75	0.82	0.60
1-4	0.76	0.67	0.88	0.75	0.79	0.56

Table 3

Performance of LDA at Predicting Dropout Week (Big Data in Education)

Weeks	80 use-case model					Cohen's κ
	ACC	TNR	TPR	F1	AUC	
1-1	0.71	0.51	0.82	0.79	0.69	0.42
1-2	0.62	0.60	0.65	0.53	0.72	0.37
1-3	0.71	0.77	0.58	0.54	0.58	0.11
1-4	0.64	0.64	0.64	0.50	0.61	0.19
1-5	0.60	0.56	0.72	0.51	0.67	0.30
1-6	0.61	0.58	0.67	0.54	0.65	0.26
1-7	0.74	0.27	0.94	0.83	0.53	0.09

Table 4
Performance of LDA at Predicting Dropout Week (Code Yourself)

Weeks	20 use-case model					Cohen's κ
	ACC	TNR	TPR	F1	AUC	
1-1	0.55	0.00	1.00	0.71	0.50	0.00
1-2	0.71	0.51	0.89	0.76	0.70	0.41
1-3	0.70	0.70	0.72	0.70	0.71	0.42
1-4	0.72	0.85	0.61	0.69	0.73	0.46

E Generality Results

Table 5
Performance of the Code Yourself Trained
HMM at Predicting Dropout for Big Data in Education

Weeks	10 state model					Cohen's κ
	ACC	TNR	TPR	F1	AUC	
1-1	0.76	0.56	1.00	0.80	0.78	0.54
1-2	0.75	0.69	1.00	0.62	0.85	0.47
1-3	0.76	0.71	1.00	0.57	0.86	0.45
1-4	0.77	0.72	1.00	0.58	0.86	0.45
1-5	0.88	0.89	0.85	0.71	0.87	0.64
1-6	0.84	0.83	0.88	0.71	0.86	0.61
1-7	0.81	0.56	0.99	0.85	0.78	0.58

Table 6
Performance of the Code Yourself Trained
Baseline-LDA at Predicting Dropout for Big Data in Education

Weeks	20 use-case model					Cohen's κ
	ACC	TNR	TPR	F1	AUC	
1-1	0.75	0.28	0.95	0.83	0.67	0.38
1-2	0.63	0.50	0.89	0.61	0.69	0.32
1-3	0.42	0.21	0.92	0.48	0.42	0.09
1-4	0.58	0.53	0.70	0.48	0.62	0.18
1-5	0.59	0.52	0.76	0.51	0.64	0.21
1-6	0.61	0.48	0.86	0.60	0.67	0.28
1-7	0.66	0.66	0.66	0.73	0.66	0.28

Table 7
Performance of the Big Data in Education
Trained HMM at Predicting Dropout for Code Yourself

Weeks	10 state model					Cohen's κ
	ACC	TNR	TPR	F1	AUC	
1-1	0.76	0.66	1.00	0.71	0.83	0.54
1-2	0.80	0.75	0.90	0.76	0.83	0.60
1-3	0.78	0.71	0.91	0.74	0.81	0.56
1-4	0.77	0.70	0.86	0.75	0.78	0.54

Table 8
 Performance of Big Data in Education Trained
 Baseline-LDA at Predicting Dropout for Code Yourself

Weeks	80 use-case model					
	ACC	TNR	TPR	F1	AUC	Cohen's κ
1-1	0.55	0.00	1.00	0.71	0.50	0.00
1-2	0.70	0.62	0.79	0.73	0.70	0.41
1-3	0.55	0.15	0.99	0.68	0.57	0.14
1-4	0.66	0.46	0.85	0.71	0.66	0.31

6.3 Summary

The work presented in this chapter completes the arc of our investigation into students' interactions with the academic domain. In keeping with the central predictions of Tinto's (1975) model, we find that students' interactions are not only associated with their academic performance (as discussed in Chapter 4 and Chapter 5), but also with their dropout decisions.

In the paper we proposed a novel methodology that provides temporal representations of students' interactions with online courses. In keeping with our research questions, we also sought to provide a model that is interpretable. Our understanding of interpretability was grounded in the literature surrounding interpretable AI, from which we extracted two criteria against which the interpretability of our model was assessed, namely, dimensionality and temporality. The student representation that we develop not only captures the timing of students' interactions with course resources relative to their prescribed week, but also an absence of participation. Comparing our HMM representation to the LDA method proposed by Coleman et al. (2015), we find that the reduced dimensionality of our representation's observation space, and the explicit modeling of temporal dynamics engendered by the HMM, not only make our approach more interpretable, but also allow us to understand patterns of student interaction that would be ambiguous under the LDA representation.

We also found that our representation captures salient information that may be used to preempt student dropout. In validating this claim, we evaluated the HMM and LDA models on a weekly dropout prediction task, and found that our approach substantially outperformed. Finally, we investigated the extent to which the states described by our HMM capture global behaviours that generalise across course settings. By conducting the same dropout prediction task using models trained and tested on separate courses, we found that our model achieves a remarkable degree of generality.

While our research is concerned with interpretability, the student representations generated by our model are derived from a set of model parameters, and thus require expert knowledge to interpret. Nevertheless, the present study demonstrates that our proposed method captures the salient features of students' interactions better than existing approaches. It thus provides an essential foundation for future work to explore how these salient features may be conveyed to stakeholders not schooled in statistical analysis, in particular, course instructors. One promising approach would be to use the model parameters to visualise how students move between interaction patterns, as well as flagging any students who are at high risk of moving to a pattern of no activity.

Given the validity, interpretability, and generality of our proposed method, we contend that, with the appropriate visualisations, it could substantially influence the research and practice of learning analytics by providing interpretable predictions of students at-risk, even in lieu of course-specific training data.

7

Conclusions

You do things and do things and nobody really has a clue.

— John Updike, *Rabbit, Run*

While the work presented in this thesis takes inspiration from Tinto's (1975) model of student interactions, the implications of our research go far beyond this model, and address key questions facing the field of learning analytics. In particular, given the widespread availability of student data, it is essential that any analytics provide valid, theoretically-grounded, and interpretable insight for stakeholders. Furthermore, the scope and scale of the data that is available to researchers provide ample opportunity for the development of novel methodologies.

In this chapter, we briefly summarise the main findings, contributions, and implications of the work presented in this thesis. In doing so, we focus on how our work has addressed the four research questions outlined in Section 1.1. Given the influence that learning analytics can have on educational policy and practice, we pay particular attention to the impact the present work may have on future research and practice. Finally, we conclude with a short overview of the thesis and a summary of its key contributions.

7.1 Impact of the Present Work

7.1.1 RQ 1: Ensuring Validity and Generality

Concerns surrounding the validity of methods used within learning analytics serve as a foundation for the work presented in this thesis. While validity is a multi-faceted construct that may be subdivided into a number of core types, we are particularly concerned with two: construct validity and generalisability. The importance of these two is hard to overstate as, without them, the relevance of any results to the research and practice of learning and teaching cannot be assured.

Following the division of students' interactions into the social and academic domains (Tinto, 1975), the thesis began by assessing social interactions. While SNA has been a predominant approach within learning analytics, limited research has considered the validity of these techniques (Joksimović et al., 2016; Dawson, Gašević, Siemens, & Joksimović, 2014). Accordingly, in Chapter 2

we assessed the influence that social tie definitions have on the structural and statistical properties of networks derived from discussion forum interactions. Our analysis demonstrated that the choice of tie definition substantially alters the structural and statistical properties of a network, and illustrates the need for researchers to justify their choice of tie definition. While this study was framed by Tinto's (1975) model, the implications are compatible with alternative theoretical perspectives such as GSCL, which often employs SNA to examine how the quantity and quality of ties influences interactions (Hurme, 2006).

Having considered validity within the study of social networks, we turned our attention to an empirical investigation of the social domain, as defined by Tinto (1975). In doing so, we created vast co-enrolment networks which were analysed using graph-embedding techniques. Given the novelty of this approach within learning analytics, it was essential to ensure that our analysis was valid. This was achieved using two distinct approaches. On the one hand, we assessed how the performance on grade prediction and dropout classification tasks was influenced by changes in the parameterisation of our embedding techniques; on the other, we performed a corruption procedure (which pruned up to 50% of the co-enrolment ties) before evaluating how this corruption influenced performance on the same two tasks. Our results found that the embedding techniques we employed were robust both to changes in their parameterisation and to corruption in the underlying networks, indicating a degree of construct validity.

Our focus on validity persisted into our investigation of students' interactions with the academic domain. For instance, in Chapter 5 we identified that while the term "engagement" is frequently employed in the learning analytics literature, there is little consensus regarding the very definition of the construct. This raises serious concerns regarding validity, which we addressed by grounding our understanding of engagement within the theoretical literature; specifically, the conceptualisation of engagement provided by Joksimović et al. (2018). In implementing this model, we used metrics common to the learning analytics literature and employed factor analysis to identify a latent structure within the data. In doing so, we sought to identify a structure that generalised across course contexts. This was assessed using the latent variable modeling concept of measurement invariance: that is, whether or not, under different conditions, measurements yield measures of the same attributes (Horn & Mcardle, 1992). While our analysis was limited to an assessment of factor means, we found that these did not vary significantly across the three courses under analysis, providing preliminary evidence of our engagement model's generality.

Finally, in Chapter 6 we provided a foundation for developing a representation of students' interactions with course resources that can be interpreted by stakeholders such as course instructors. In doing so, we explicitly modeled the temporal dynamics of these interactions, and argued that this facilitates the prediction of student dropout. To validate this claim, we conducted a dropout prediction task on both a single and a cross-course basis. Our results not only validate that our model captures our construct of interest – dropout – but also suggest that our model describes

global behaviours that generalise across course settings.

7.1.2 RQ 2: Theoretically-Grounded Analytics

In addition to validity, the role of theory is increasingly recognised as essential for informing not only the choice of questions asked, but also the hypotheses tested (Rogers et al., 2016; Wise & Shaffer, 2015). Broadly speaking, the argument goes that since learning analytics are about learning, any computational analyses should be grounded within existing educational theory and research (Gašević et al., 2015). As such, the work presented in this thesis is broadly framed by Tinto's (1975) model of student integration. In Chapter 3 we focused on the social domain and investigated two key predictions of the model: namely, that students' social interactions are associated with not only their academic performance, but also their persistence in their studies.

While Tinto's (1975) model frames the overall thesis, our investigation of the academic domain also draws on the theoretical literature to identify generalisable models of students' engagement with academic resources. For instance, in Chapter 4 we grounded our analysis in the literature surrounding "learning strategies" (Weinstein et al., 2012) and found that differences in students' choice of strategy was associated with their academic performance. In Chapter 5 we addressed our concerns regarding the generality of learning strategies by operationalising and empirically validating a theoretical model of student engagement. Although this was originally developed for formal educational settings (Reschly & Christenson, 2012), the model was subsequently adapted for non-formal learning environments by Joksimović et al. (2018). While conducted on a limited set of three courses, our analysis provided partial validation of the engagement construct posited by Joksimović et al.'s (2018) model.

7.1.3 RQ 3: Towards Interpretable Models

While the preceding research questions identify central issues within learning analytics, if the field is to have widespread impact, it is necessary that any insights are effectively conveyed to the relevant stakeholders. These concerns relate to what the machine learning literature refers to as "interpretability"; that is, "the ability to explain or present [the results of any analysis] in understandable terms to a human" (Doshi-Velez & Kim, 2017, p. 2). We first addressed this problem in Chapter 4, where we developed a taxonomy of students' interactions with course materials. Specifically, drawing on the literature, we postulated that students' learning strategies could be decomposed into shorter-term "tactics" (Alexander et al., 1998; Kirby, 1988). These tactics were described by the states of an HMM, while learning strategies were identified as common sequences of these tactics. We argued that the reduction of students' complex interaction patterns into a limited set of descriptive states marks an important step towards interpretability.

In Chapter 6 we proposed a model that provides representations of student interactions. Importantly, the parameters of this model are more interpretable to those with expert statistical knowl-

edge than competing approaches in the literature. To demonstrate this, we defined interpretability in terms of the dimensionality of the observation space, and the manner in which temporal relations are described by the model. Our representation consists of 9 features describing not only the extent of students' interactions with different course elements, but also the timing of these interactions in relation to the course design. In addition, we add a 10th "no interaction" feature capturing a total absence of activity. Once again, student interactions were modeled with an HMM and, to assess the interpretability of this approach, we compared it to the LDA method proposed by Coleman et al. (2015). Our results showed that not only was our HMM approach more interpretable, but we were also able to explain patterns of interaction that were ambiguous under the LDA representation. This work, however, is but the first step towards providing stakeholders with interpretable, and thus actionable, insight.

7.1.4 RQ 4: Novel Methodologies

Learning is increasingly recognised as a profoundly complex process (Jacobson et al., 2016). To account for this, it is essential that the field of learning analytics not only provide valid, theoretically-grounded and interpretable models, but also develop methods that can cope with this complexity. Accordingly, a major contribution of the work presented in this thesis has been the development and assessment of novel methodologies. For instance, in Chapter 3 we presented a set of graph-embedding techniques and assessed the influence that students' social networks had on their academic performance and dropout decisions. This analysis was conducted at an unprecedented scale within the literature. Furthermore, it was theoretically-grounded in Tinto's (1975) model of student integration, and the robustness of these results indicated a degree of construct validity.

Finally, in Chapter 6, we employed a particular variant of the HMM. Specifically, we used a sticky-HMM with a hierarchical Dirichlet process (HDP) prior to model students' interactions with course resources. While the "sticky" assumption biases the model towards self-transitions (Fox, Sudderth, Jordan, Willsky, et al., 2011), the HDP prior places diminishing probability mass on infrequent states and thus focuses the majority of the probability mass on a relatively small number of major states in the model (Fox, Sudderth, Jordan, & Willsky, 2008). This parsimony of the state-space contributed to the model's ability to readily predict student dropout and facilitated the identification of global behaviours (i.e. states) which generalised to a remarkable extent across course settings.

7.2 Implications for Research and Practice

By addressing a number of central problems within learning analytics, this thesis has important implications for research and practice, and reveals a number of promising directions for future study. In particular, given that the majority of the studies presented are focused on the development and assessment of novel methods, these implications are predominantly methodological in nature.

The thesis began by assessing the influence that an oft-overlooked methodological detail has on

the results on any SNA analysis. Specifically, we investigated the construct validity of a number of commonly used social tie definitions in the analysis of discussion forums. We found that, across two diverse learning environments, the choice of tie definition profoundly influences the properties of the derived network. While this is perhaps an unsurprisingly result – researchers such as Gruzd and Haythornthwaite (2008) have long since noted that each definition makes specific assumptions about the nature of social interactions – the importance of this decision is not reflected in the literature. Accordingly, our results emphasise the need for not only transparency in researchers' choice of tie definition, but also an explicit justification for this choice. Furthermore, we advise that practitioners investigate a range of definitions to ascertain the extent to which such methodological decisions can bias their results.

Having emphasised the importance of validity in the study of social networks, in Chapter 3 we presented a methodology, novel within learning analytics, for assessing the influence that students' social networks have on their academic performance and persistence. In doing so, our study assessed co-enrolment networks at an unprecedented scale and offers SNA practitioners a viable alternative to a number of common centrality metrics which are limited by their high computational complexity (Cui et al., 2017). Furthermore, the robustness of this method to changes in the parameterisation permits cheap computation with limited cost to downstream model performance. From a research perspective, the use of graph-embedding techniques reveals a number of interesting directions for future research, such as the identification of learning communities, or assessing the temporal development of students' social networks and how this influences their academic outcomes and persistence.

In investigating students' interactions within the academic domain, we developed a number of novel methodologies. The first of these, presented in Chapter 4, sought to reduce the manifest diversity of students' interactions into a set of theoretically-grounded learning strategies. In contrast to the existing literature, which has traditionally relied upon self-reported data (Bannert et al., 2014; P Winne, 2013), our computational approach permits the analysis of learning strategies at scale. Furthermore, rather than capturing students' *perceptions* of their learning, our trace-based approach captures their *actual* use of study tactics and learning strategies (P Winne & Jamieson-Noel, 2002; Hadwin et al., 2007). That said, the literature emphasises the need for these representations to be supplemented with information about students' dispositions in order to account for *why* students display certain behavioural patterns (Tempelaar et al., 2017). Combining such data with our approach would permit the real-time identification of different study tactics and learning strategies, and could be used not only to inform the provision of feedback, but also to evaluate the impact such feedback has on students' choice of learning strategy.

In an effort to develop a more general model of student interactions, we turned to the theoretical literature, and focused on the research surrounding student engagement. Having identified a theoretical model, developed by Reschly and Christenson (2012) and reoperationalised for non-

formal online settings by Joksimović et al. (2018), we sought to validate the model empirically using trace data drawn from three diverse MOOCs. In doing so, we demonstrated that theories drawn from the learning analytics literature may be subjected to empirical validation. This is an important contribution and suggests that framing analyses within theory can not only inform the questions asked (Rogers et al., 2016; Wise & Shaffer, 2015), but can also lead us to re-evaluate the predictions made by a theory in the light of empirical results. This approach to the modeling of student engagement has already informed practice: informal conversations with researchers at an Australian university has found that this model provides the foundation for theory-informed design for reporting tools addressing student engagement.

Finally, in Chapter 6, we proposed a novel methodology that provides interpretable representations of students' interactions with course materials. This interpretability, however, is limited to individuals with expert statistical knowledge. For such individuals, the model parameters are richly descriptive and outline a set of states which can be supplemented with semantic content. When considered across time, the trajectories of students through these interpretable states are not only predictive of dropout, but are also able to provide *explanations* for these predictions. This feature of the model has considerable potential and merits a user study to evaluate its utility for course instructors. To achieve this, however, will require substantial future research. In particular, the model requires further validation on additional courses as well as a thorough investigation of how best to convey the insights to stakeholders unschooled in statistical analysis. One promising approach is the use of dashboard visualisations, such as sankey diagrams, to illustrate how students transition between the states in the model. These states will also need semantic descriptions, which could be automatically generated as the states are probability distributions over semantically meaningful labels (for instance, the probability of interacting with assignments from previous weeks). Finally, students who are likely to transition to states associated with dropout will need to be highlighted to the instructor, in order to inform their intervention. Evaluating how the insights derived from our model may be conveyed to stakeholders is an important research question, and will require a set of user studies to assess the comparative interpretability of a number of visualisations. In doing so, the extensive literature surrounding the assessment and evaluation of student dashboards will be an essential guide.

The work presented in this thesis, however, is not without its limitations. In particular, our investigation of the academic domain was constrained by the data to which we had access. That is, trace logs of students' interactions. While the literature has emphasised the value of such logs, which provide information about students' *actual* (rather than *perceived*) behaviour (P. Winne & Jamieson-Noel, 2002; Hadwin et al., 2007), this is not to say that supplemental data would not lead to improved models or insight. In particular, survey and questionnaire data could shed light on students' dispositions, which could help answer *why* students display certain behaviours (Tempelaar et al., 2017). Understanding students' motivations in this way could lead to vast improvements

in the timeliness and appropriateness of interventions. Data access, however, was not the only limitation. Given the context of student dropout at higher education institutions that gave rise to Tinto's (1975) model, it is not clear how suitable this approach is for MOOCs, where students' interactions are fundamentally different to campus environments. However, this is a limitation of the existing theoretical literature regarding MOOCs, which is beyond the scope of the current thesis. Nevertheless, there is considerable potential for further theoretical work regarding student retention in MOOCs.

7.3 Conclusions

Learning analytics was founded for the purpose of harnessing the vast quantities of data that institutions generate about the learning activities of students (Gašević, Kovanović, & Joksimović, 2017). Since its inception, the field has attracted considerable attention, and has prompted educational institutions to harvest the thousands of digital traces that students generate. However, this data introduces a number of challenges, not least its inherent ambiguity (Siemens, 2013). Given the uncertain quality of this data, it is unsurprising that the methodological validity of some research has been called into question (Caulfield, 2013; Dawson et al., 2017). To address this, the literature has emphasised the importance of validity (Gašević et al., 2017), theoretically-grounded research (Rogers et al., 2016), and interpretable models (Pardo et al., 2016). These challenges to learning analytics research form the basis of this thesis, and are mirrored in our research questions.

We have presented a number of methods and techniques that address these challenges and facilitate the analysis of the vast quantities of digital data that higher educational institutions generate. This work has been framed by Tinto's (1975) model of student integration and presents methods that have the potential to identify students at-risk on the basis of social and academic interactions. This is particularly important in online learning, where large student numbers make it challenging for instructors to monitor and assist students. Finally, this thesis lays the foundation for real-world experiments. It provides the means to not only identify where and when to intervene, but also how the impact of an intervention may be assessed.

Bibliography

- Ahmed, N. K., Rossi, R., Lee, J. B., Kong, X., Willke, T. L., Zhou, R., & Eldardiry, H. (2018). Learning role-based graph embeddings. *arXiv preprint arXiv:1802.02896*. Retrieved from <https://arxiv.org/pdf/1802.02896.pdf>. (Cited on p. 43)
- Alexander, P. A., Graham, S., & Harris, K. R. (1998). A perspective on strategy research: Progress and prospects. *Educational Psychology Review*, 10(2), 129–154. (Cited on pp. 9, 63, 134).
- Altbach, P. G. & Reisberg, L. (2018). Global trends and future uncertainties. *Change: The Magazine of Higher Learning*, 50(3-4), 63–67. doi:10.1080/00091383.2018.1509601. (Cited on p. 1)
- Amaury Nora, A. M., L. C. Attinasi. (1990). Testing qualitative indicators of precollege factors in tinto's attrition model: A community college student population. *The Review of Higher Education*, 13(3), 337–355. doi:10.1353/rhe.1990.0021. (Cited on p. 42)
- Bannert, M., Reimann, P., & Sonnenberg, C. (2014). Process mining techniques for analysing patterns and strategies in students' self-regulated learning. *Metacognition and Learning*, 9(2), 161–185. doi:10.1007/s11409-013-9107-6. (Cited on pp. 63, 136)
- Braxton, J. M., Sullivan, A. V. S., & Johnson, R. M. (1997). Appraising tinto's theory of college student departure. In J. Smart (Ed.), *Higher education: Handbook of theory and research 12* (pp. 107–164). Agathon Press. New York. (Cited on p. 3).
- Calderon, A. (2018). *Massification of higher education revisited*. RMIT University. [Online; accessed 1-August-2019]. (Cited on p. 1).
- Caulfield, M. (2013). What the course signals "kerfuffle" is about, and what it means to you. <https://er.educause.edu/blogs/2013/11/what-the-course-signals-kerfuffle-is-about-and-what-it-means-to-you>. Accessed: 2019-08-01. (Cited on pp. 4, 138).
- Coleman, C. A., Seaton, D. T., & Chuang, I. (2015). Probabilistic use cases: Discovering behavioral patterns for predicting certification. In *Proceedings of the second (2015) acm conference on learning @ scale* (pp. 141–148). L@S '15. Vancouver, BC, Canada: ACM. doi:10.1145/2724660.2724662. (Cited on pp. 95, 131, 135)
- Corrin, L. & de Barba, P. (2014). Exploring students' interpretation of feedback delivered through learning analytics dashboards. In B. Hegarty, J. McDonald, & S.-K. Loke (Eds.), *Proceedings of the ascilite 2014 conference* (pp. 629–633). Dunedin, New Zealand. (Cited on p. 4).

- Cui, P., Wang, X., Pei, J., & Zhu, W. (2017). A survey on network embedding. *IEEE Transactions on Knowledge and Data Engineering*. doi:10.1109/TKDE.2018.2849727. (Cited on pp. 42, 136)
- Dawson, S., Gašević, D., Siemens, G., & Joksimović, S. (2014). Current state and future trends: A citation network analysis of the learning analytics field. In *Proceedings of the fourth international conference on learning analytics & knowledge* (pp. 231–240). LAK '14. ACM. (Cited on p. 132).
- Dawson, S., Jovanovic, J., Gašević, D., & Pardo, A. (2017). From prediction to impact: Evaluation of a learning analytics retention program. In *Proceedings of the seventh international learning analytics & knowledge conference* (pp. 474–478). LAK '17. New York, NY, USA: ACM. doi:10.1145/3027385.3027405. (Cited on pp. 4, 138)
- den Bossche, P. V., Gijssels, W. H., Segers, M., & Kirschner, P. A. (2006). Social and cognitive factors driving teamwork in collaborative learning environments: Team learning beliefs and behaviors. *Small Group Research*, 37(5), 490–521. doi:10.1177/1046496406292938. (Cited on p. 2)
- Doshi-Velez, F. & Kim, B. (2017). Towards a rigorous science of interpretable machine learning. *arXiv preprint arXiv:1702.08608*. (Cited on pp. 94, 134).
- Durkheim, É. (1961). *Suicide (j. spauling & g. simpson, trans)*. Glencoe: The Free Press. (Cited on p. 2).
- Eckles, J. E. & Stradley, E. G. (2012). A social network analysis of student retention using archival data. *Social Psychology of Education*, 15(2), 165–180. doi:10.1007/s11218-011-9173-z. (Cited on pp. 3, 62)
- Fairchild, A. J., Horst, S. J., Finney, S. J., & Barron, K. E. (2005). Evaluating existing and new validity evidence for the academic motivation scale. *Contemporary Educational Psychology*, 30(3), 331–358. doi:https://doi.org/10.1016/j.cedpsych.2004.11.001. (Cited on p. 2)
- Fincham, E., Gašević, D., Jovanović, J., & Pardo, A. (2019). From study tactics to learning strategies: An analytical method for extracting interpretable representations. *IEEE Transactions on Learning Technologies*, 12(1), 59–72. doi:10.1109/TLT.2018.2823317. (Cited on pp. viii, 9, 65)
- Fincham, E., Gašević, D., & Pardo, A. (2018). From social ties to network processes: Do tie definitions matter? *Journal of Learning Analytics*, 5(2), 9–28. doi:10.18608/jla.2018.52.2. (Cited on pp. viii, 7, 14)
- Fincham, E., Hoernle, N., Gal, K., & Gašević, D. (2019). The road not taken: Preempting dropout in moocs. *Manuscript submitted for publication*. (Cited on pp. viii, 11, 95).
- Fincham, E., Rózemberczki, B., Kovanović, V., Joksimović, S., & Gašević, D. (2019). Persistence and performance in co-enrolment network embeddings. *Manuscript submitted for publication*. (Cited on pp. viii, 8, 44).

- Fincham, E., Whitelock-Wainwright, A., Kovanović, V., Joksimović, S., van Staalduinen, J.-P., & Gašević, D. (2019). Counting clicks is not enough: Validating a theorized model of engagement in learning analytics. In *Proceedings of the 9th international conference on learning analytics & knowledge* (pp. 501–510). LAK19. Tempe, AZ, USA: ACM. doi:10.1145/3303772.3303775. (Cited on pp. viii, 10, 82)
- Fox, E. B., Sudderth, E. B., Jordan, M. I., & Willsky, A. S. (2008). An hdp-hmm for systems with state persistence. In *Proceedings of the 25th international conference on machine learning* (pp. 312–319). ACM. (Cited on p. 135).
- Fox, E. B., Sudderth, E. B., Jordan, M. I., Willsky, A. S., et al. (2011). A sticky hdp-hmm with application to speaker diarization. *The Annals of Applied Statistics*, 5(2A), 1020–1056. (Cited on p. 135).
- Gašević, D., Dawson, S., & Siemens, G. (2015). Let's not forget: Learning analytics are about learning. *TechTrends*, 59(1), 64–71. doi:http://doi.org/10.1007/s11528-014-0822-x. (Cited on pp. 1, 4, 134)
- Gašević, D., Kovanović, V., & Joksimović, S. (2017). Piecing the learning analytics puzzle: A consolidated model of a field of research and practice. *Learning: Research and Practice*, 3(1), 63–78. doi:10.1080/23735082.2017.1286142. (Cited on p. 138)
- Gilpin, L., Bau, D., Z. Yuan, B., Bajwa, A., Specter, M., & Kagal, L. (2018). Explaining explanations: An overview of interpretability of machine learning. In *Proceedings of the 5th ieee international conference on data science and advanced analytics (dsaa)* (pp. 80–89). doi:10.1109/DSAA.2018.00018. (Cited on p. 94)
- Gruzd, A. & Haythornthwaite, C. A. (2008). Automated discovery and analysis of social networks from threaded discussions. In *Proceedings of the 18th international sunbelt social network conference (insna 2008)*. (Cited on p. 136).
- Gunn, A. (2018). Metrics and methodologies for measuring teaching quality in higher education: Developing the teaching excellence framework (tef). *Educational Review*, 70(2), 129–148. doi:10.1080/00131911.2017.1410106. (Cited on p. 1)
- Hadwin, A. F., Nesbit, J. C., Jamieson-Noel, D., Code, J., & Winne, P. H. (2007). Examining trace data to explore self-regulated learning. *Metacognition and Learning*, 2(2), 107–124. doi:10.1007/s11409-007-9016-7. (Cited on pp. 6, 63, 64, 80, 136, 137)
- Halevy, A., Norvig, P., & Pereira, F. (2009). The unreasonable effectiveness of data. *IEEE Intelligent Systems*, 24, 8–12. Retrieved from http://www.computer.org/portal/cms_docs_intelligent/intelligent/homepage/2009/x2exp.pdf. (Cited on p. 1)
- Horn, J. & Mcardle, J. (1992). A practical and theoretical guide to measurement invariance in aging research. *Experimental Aging Research*, 18(3), 117–144. doi:10.1080/03610739208253916. (Cited on p. 133)

- Hurme, T.-R. (2006). Metacognition in joint discussions: An analysis of the patterns of interaction and the metacognitive content of the networked discussions in mathematics. *Metacognition and Learning*, 1(2), 181–200. doi:10.1007/s11409-006-9792-5. (Cited on p. 133)
- Jacobson, M. J., Kapur, M., & Reimann, P. (2016). Conceptualizing debates in learning and educational research: Toward a complex systems conceptual framework of learning. *Educational Psychologist*, 51(2), 210–218. doi:10.1080/00461520.2016.1166963. (Cited on pp. 5, 135)
- Jayaprakash, S. M., Moody, E. W., Lauría, E. J., Regan, J. R., & Baron, J. D. (2014). Early alert of academically at-risk students: An open source analytics initiative. *Journal of Learning Analytics*, 1(1), 6–47. doi:10.18608/jla.2014.11.3. (Cited on p. 4)
- Jeong, H., Biswas, G., Johnson, J., & Howard, L. (2010). Analysis of productive learning behaviors in a structured inquiry cycle using hidden markov models. In *Proceedings of the 3rd international conference of educational data mining* (pp. 81–90). Pittsburgh, PA. (Cited on p. 64).
- Joksimović, S., Manataki, A., Gašević, D., Dawson, S., Kovanović, V., & Friss De Kereki, I. (2016). Translating network position into performance: Importance of centrality in different network configurations. In *Proceedings of the 6th international conference on learning analytics & knowledge (lak 2016)* (pp. 314–323). ACM Press. doi:10.1145/2883851.2883928. (Cited on pp. 62, 132)
- Joksimović, S., Poquet, O., Kovanović, V., Dowell, N., Mills, C., Gašević, D., Dawson, S., Graesser, A. C., & Brooks, C. (2018). How do we model learning at scale? a systematic review of research on moocs. *Review of Educational Research*, 88(1), 43–86. doi:10.3102/0034654317740335. (Cited on pp. 9, 10, 82, 93, 133, 134, 137)
- Jovanović, J., Gašević, 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. doi:10.1016/j.iheduc.2017.02.001. (Cited on pp. 64, 80)
- Kandiko, C. B. & Mawer, M. (2013). *Student expectations and perceptions of higher education*. London: Kings Institute. (Cited on p. 1).
- Kelly, G. (2012). Lecture attendance rates at university and related factors. *Journal of Further and Higher Education*, 36, 17–40. doi:10.1080/0309877X.2011.596196. (Cited on p. 43)
- Kirby, J. R. (1988). Style, strategy, and skill in reading. In R. R. Schmeck (Ed.), *Learning strategies and learning styles* (pp. 229–274). Boston, MA: Springer US. (Cited on pp. 9, 63, 64, 134).
- Kovanović, V., Gašević, D., Joksimović, S., Hatala, M., & Adesope, O. (2015). Analytics of communities of inquiry: Effects of learning technology use on cognitive presence in asynchronous online discussions. *The Internet and Higher Education*, 27, 74–89. (Cited on p. 64).
- Kovanović, V., Joksimović, S., Poquet, O., Hennis, T., de Vries, P., Hatala, M., Dawson, S., Siemens, G., & Gašević, D. (2019). Examining communities of inquiry in massive open online courses: The role of study strategies. *The Internet and Higher Education*, 40, 20–43. doi:https://doi.org/10.1016/j.iheduc.2018.09.001. (Cited on p. 64)

- Lesnik-Oberstein, K. (2015). Let uk universities do what they do best – teaching and research. *The Guardian, Letters*. Retrieved from <https://www.theguardian.com/education/2015/jul/06/let-uk-universities-do-what-they-do-best-teaching-and-research>. (Cited on p. 1)
- Matcha, W., Ahmad Uzir, N., Gašević, D., & Pardo, A. (2019). A systematic review of empirical studies on learning analytics dashboards: A self-regulated learning perspective. *IEEE Transactions on Learning Technologies*, 1–1. doi:10.1109/TLT.2019.2916802. (Cited on p. 4)
- Matcha, W., Gašević, D., Uzir, N. A., Jovanović, J., & Pardo, A. (2019). Analytics of learning strategies: Associations with academic performance and feedback. In *Proceedings of the 9th international conference on learning analytics & knowledge* (pp. 461–470). LAK19. Tempe, AZ, USA: ACM. doi:10.1145/3303772.3303787. (Cited on p. 64)
- McCowan, T. (2017). Higher education, unbundling, and the end of the university as we know it. *Oxford Review of Education*, 43(6), 733–748. doi:10.1080/03054985.2017.1343712. (Cited on p. 1)
- Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American psychologist*, 50(9), 741. (Cited on p. 13).
- Moore, T. & Morton, J. (2017). The myth of job readiness? written communication, employability, and the 'skills gap' in higher education. *Studies in Higher Education*, 42(3), 591–609. doi:10.1080/03075079.2015.1067602. (Cited on p. 1)
- Munro, B. H. (1981). Dropouts from higher education: Path analysis of a national sample. *American Educational Research Journal*, 18(2), 133–141. doi:10.3102/00028312018002133. (Cited on p. 42)
- Nora, A. & Rendon, L. I. (1990). Determinants of predisposition to transfer among community college students: A structural model. *Research in Higher Education*, 31(3), 235–255. Retrieved from <http://www.jstor.org/stable/40195941>. (Cited on pp. 3, 42)
- Pardo, A., Mirriahi, N., Martinez-Maldonado, R., Jovanovic, J., Dawson, S., & Gašević, D. (2016). Generating actionable predictive models of academic performance. In *Proceedings of the sixth international conference on learning analytics & knowledge* (pp. 474–478). LAK '16. Edinburgh, United Kingdom: ACM. doi:10.1145/2883851.2883870. (Cited on pp. 5, 138)
- Pascarella, E. T. & Terenzini, P. T. (1980). Predicting freshman persistence and voluntary dropout decisions from a theoretical model. *The Journal of Higher Education*, 51(1), 60–75. Retrieved from <http://www.jstor.org/stable/1981125>. (Cited on p. 42)
- Perozzi, B., Al-Rfou, R., & Skiena, S. (2014). Deepwalk: Online learning of social representations. In *Proceedings of the 20th acm sigkdd international conference on knowledge discovery and data mining* (pp. 701–710). KDD '14. New York, New York, USA. doi:10.1145/2623330.2623732. (Cited on p. 43)

- Reimann, P. (2016). Connecting learning analytics with learning research: The role of design-based research. *Learning: Research and Practice*, 2(2), 130–142. doi:10.1080/23735082.2016.1210198. (Cited on pp. 4, 5)
- Reschly, A. L. & Christenson, S. L. (2012). Jingle, jangle, and conceptual haziness: Evolution and future directions of the engagement construct. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 3–19). Boston, MA: Springer US. doi:10.1007/978-1-4614-2018-7_1. (Cited on pp. 9, 81, 82, 134, 136)
- Rienties, B., Lewis, T., McFarlane, R., Nguyen, Q., & Toeteneel, L. (2018). Analytics in online and offline language learning environments: The role of learning design to understand student online engagement. *Computer Assisted Language Learning*, 31(3), 273–293. doi:10.1080/09588221.2017.1401548. (Cited on p. 80)
- Rienties, B., Tempelaar, D., den Bossche, P. V., Gijsselaers, W., & Segers, M. (2009). The role of academic motivation in computer-supported collaborative learning. *Computers in Human Behavior*, 25(6), 1195–1206. doi:https://doi.org/10.1016/j.chb.2009.05.012. (Cited on p. 2)
- Rogers, T., Dawson, S., & Gašević, D. (2016). Learning analytics and the imperative for theory-driven research. In C. Haythornthwaite, R. Andrews, J. Fransman, & E. Meyers (Eds.), *The sage handbook of e-learning research* (pp. 232–250). SAGE Publications. (Cited on pp. 4, 81, 134, 137, 138).
- Rolfe, H. (2001). The effect of tuition fees on students' demands and expectations: Evidence from case studies. *National Institute of Economic and Social Research, NIESR Discussion Papers*. (Cited on p. 1).
- Roschelle, J. & Teasley, S. D. (1995). The construction of shared knowledge in collaborative problem solving. In C. O'Malley (Ed.), *Computer supported collaborative learning* (pp. 69–97). (Cited on p. 2).
- Rosenfeld, A. & Richardson, A. (2019). Explainability in human-agent systems. *Autonomous Agents and Multi-Agent Systems*, 1–33. (Cited on p. 94).
- Rourke, L., Anderson, T., Garrison, D., & Archer, W. (1999). Assessing social presence in asynchronous text-based computer conferencing. *Journal of Distance Education*, 14(2), 50–71. Retrieved from <http://www.ijede.ca/index.php/jde/article/view/153>. (Cited on p. 2)
- Ryan, R. M. & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology*, 25(1), 54–67. doi:https://doi.org/10.1006/ceps.1999.1020. (Cited on p. 2)
- Schreurs, B., Teplovs, C., Ferguson, R., de Laat, M., & Buckingham Shum, S. (2013). Visualizing social learning ties by type and topic: Rationale and concept demonstrator. In *Proceedings of the third international conference on learning analytics and knowledge* (pp. 33–37). LAK '13. Leuven, Belgium: ACM. doi:10.1145/2460296.2460305. (Cited on p. 62)

- Shea, P. & Bidjerano, T. (2010). Learning presence: Towards a theory of self-efficacy, self-regulation, and the development of a communities of inquiry in online and blended learning environments. *Computers and Education*, 55(4), 1721–1731. doi:<https://doi.org/10.1016/j.compedu.2010.07.017>. (Cited on p. 2)
- Siemens, G. (2013). Learning analytics: The emergence of a discipline. *American Behavioral Scientist*, 57(10), 1380–1400. doi:[10.1177/0002764213498851](https://doi.org/10.1177/0002764213498851). (Cited on pp. 1, 2, 138)
- Skrypnyk, O., Joksimović, S., Kovanović, V., Gašević, D., & Dawson, S. (2015). Roles of course facilitators, learners, and technology in the flow of information of a cmooc. *The International Review of Research in Open and Distributed Learning*, 16(3). doi:[10.19173/irrodl.v16i3.2170](https://doi.org/10.19173/irrodl.v16i3.2170). (Cited on p. 62)
- Spady, W. G. (1970). Dropouts from higher education: An interdisciplinary review and synthesis. *Interchange*, 1(1), 64–85. doi:[10.1007/BF02214313](https://doi.org/10.1007/BF02214313). (Cited on p. 2)
- Strayhorn, T. L. (2018). *College students' sense of belonging: A key to educational success for all students*. Routledge. (Cited on p. 2).
- Tempelaar, D. T., Rienties, B., & Nguyen, Q. (2017). Towards actionable learning analytics using dispositions. *IEEE Transactions on Learning Technologies*, 10(01), 6–16. doi:[10.1109/TLT.2017.2662679](https://doi.org/10.1109/TLT.2017.2662679). (Cited on pp. 6, 80, 136, 137)
- Terenzini, P. T. & Pascarella, E. T. (1977). Voluntary freshman attrition and patterns of social and academic integration in a university: A test of a conceptual model. *Research in Higher Education*, 6(1), 25–43. doi:[10.1007/BF00992014](https://doi.org/10.1007/BF00992014). (Cited on p. 42)
- Thomas, S. L. (2000). Ties that bind: A social network approach to understanding student integration and persistence. *The Journal of Higher Education*, 71(5), 591–615. Retrieved from <http://www.jstor.org/stable/2649261>. (Cited on pp. 2, 3)
- Tinto, V. (1975). Dropout from higher education: A theoretical synthesis of recent research. *Review of Educational Research*, 45(1), 89–125. Retrieved from <http://www.jstor.org/stable/1170024>. (Cited on pp. 2–6, 8, 10, 13, 41–43, 62, 93, 94, 131–135, 138)
- Tinto, V. (1993). *Leaving college: Rethinking the causes and cures of student attrition (2st ed.)* University of Chicago Press. (Cited on p. 2).
- Tinto, V. (2012). *Completing college: Rethinking institutional action*. University of Chicago Press. (Cited on p. 2).
- Tomlinson, M. (2017). Student perceptions of themselves as ‘consumers’ of higher education. *British Journal of Sociology of Education*, 38(4), 450–467. doi:[10.1080/01425692.2015.1113856](https://doi.org/10.1080/01425692.2015.1113856). (Cited on p. 1)
- Weinstein, C. E., Husman, J., & Dierking, D. (2012). Chapter 22. self-regulation interventions with a focus on learning strategies. *Handbook of Self-regulation*. (Cited on pp. 8, 63, 134).

7. BIBLIOGRAPHY

- Winne, P. (2013). Learning strategies, study skills, and self-regulated learning in postsecondary education. In M. B. Paulsen (Ed.), *Higher education: Handbook of theory and research: Volume 28* (pp. 377–403). Dordrecht: Springer Netherlands. (Cited on pp. 63, 136).
- Winne, P. & Jamieson-Noel, D. (2002). Exploring students' calibration of self reports about study tactics and achievement. *Contemporary Educational Psychology*, 27(4), 551–572. doi:10.1016/S0361-476X(02)00006-1. (Cited on pp. 6, 63, 80, 136, 137)
- Wise, A. F. & Shaffer, D. (2015). Why theory matters more than ever in the age of big data. 2, 5–13. doi:10.18608/jla.2015.22.2. (Cited on pp. 4, 81, 134, 137)
- Zhou, M. & Winne, P. H. (2012). Modeling academic achievement by self-reported versus traced goal orientation. *Learning and Instruction*, 22(6), 413–419. doi:https://doi.org/10.1016/j.learninstruc.2012.03.004. (Cited on p. 64)