Characterizing university students' self-regulated learning behavior using dispositional learning analytics

Louise Ainscough, Richard Leung, Kay Colthorpe, Tracey Langfield
School of Biomedical Sciences, The University of Queensland, Australia.

Abstract

Learning analytics can be used in conjunction with learner dispositions to identify at-risk students and provide personalized guidance on how to improve. Participants in the current study were students (n=192) studying a first year anatomy and physiology course. A two-step cluster analysis was performed using learning analytics data from the learning management system and self-regulated learning behavior from meta-learning assessment tasks. Three clusters of students were identified – high, medium and low self-regulated learners. High self-regulated learners were engaged with the meta-learning tasks, reported the most self-regulated learning strategies and used new strategies during semester. They also had the highest academic achievement. Compared to low self-regulated leaners, medium self-regulated learners were more engaged in the meta-learning tasks and used more learning strategies during semester, including new strategies; however, both medium and low self-regulated learners had similar levels of academic achievement. It is possible that the medium self-regulated learners represent students who were attempting to improve their learning, but had not yet found strategies that were right for them. Future evaluation of academic performance may determine whether the attempts to improve learning by medium self-regulated learners distinguishes them from low self-regulated learners in the later years of their study.

Keywords: *Self-regulated learning; learning analytics; academic achievement; cluster analysis.*

1. Introduction

Every time a student interacts with a website - such as logging onto a student portal or submitting assignments - a digital trail is left behind. Learning analytics involves the collection, measurement and analysis of this trace data to improve learning and teaching (Tempelaar et al., 2018). Previous studies using learning analytics have identified online factors that correlated with course outcomes, such as log-in frequency, and engagement with the learning management system (LMS), such as submitting assessment, viewing assessment feedback and contributing to discussion boards (Colthorpe et al., 2015; Macfadyen & Dawson, 2010; Smith et al., 2012). These findings demonstrate the potential for learning analytics to identify indicators of online engagement that may influence student progress within a course and their academic performance. Such information can potentially allow course coordinators to intervene by identifying at-risk students early.

Learning analytics may help to identify disengaged students who are academically at-risk; however, it does not explain why these students are at-risk, nor how they can improve. More recently, learning analytics data has been combined with learner dispositions to better understand how students engage with learning activities, so that more targeted feedback can be provided to at-risk students (Shum & Crick, 2012; Tempelaar et al., 2017). The combined analysis of learning data and learner data has been termed 'dispositional learning analytics' (Tempelaar et al., 2017). Dispositional learning analytics researchers have augmented traditional trace data with a variety of student dispositions, including students' expectancies and values, motivations, approaches to learning, emotions, goals or lifelong learning traits (Gašević et al., 2017; Shum & Crick, 2012; Tempelaar et al., 2018; Tempelaar et al., 2017). These studies have demonstrated that collecting self-reported data on learner dispositions increased the predictive power of learning analytics models (Tempelaar et al., 2017).

Although a broad range of dispositions have been incorporated into learning analytics studies, the specific types of self-regulated learning strategies students use while they learn have been largely ignored. Past research has demonstrated that students can use a variety of different learning strategies (Colthorpe et al., 2017; Hattie & Donoghue, 2016). Certain types of strategies are more effective than others. For example, strategies where students organize or transform information are associated with high academic achievement (Nota et al., 2004), whereas more passive learning approaches, such as managing the environment, are less effective (Hattie & Donoghue, 2016). In addition, higher achieving students use more learning strategies than lower achieving students (Kitsantas, 2002). The current study expands the field of dispositional learning analytics by combining learning analytics data with students' goal orientation, the type and number of learning strategies used during exam preparation, and their strategy adaptability. This data was used to group students using two-step cluster analysis.

2. Methods

Participants were students studying a first-year anatomy and physiology course at the University of Queensland, Australia. As part of course assessment, students completed four online meta-learning tasks (12%) consisting of open-ended questions designed to encourage students to reflect on their learning (Colthorpe et al., 2018). Learning analytics data was collected from the tasks, including average word count per question, submission time and task completion rate. Other course assessment included a mid-semester exam (20%), end of semester exam (50%) and group assignment (18%).

This study received ethics approval from the University of Queensland Human Ethics Committee. To ensure consenting students (n=192; 74%) were academically representative of the whole cohort, a t-test was used to compare their final exam results. There was no significant difference (p>0.05) between consenting participants (65.8% ± 1.1) and the whole cohort (64.9% ± 1).

2.1. Self-regulated learning behaviors

Responses to meta-learning questions were thematically analyzed to characterize students' self-regulated learning behaviors. Early in semester, students described their personal goals for the course. Responses to this question were deductively coded (Braun & Clarke, 2006) into mastery or performance goals (Hattie & Donoghue, 2016). Half-way through semester, students described the study strategies they used when preparing for the mid-semester exam. Responses to this question were deductively coded using the learning strategy categories identified in the self-regulated learning interview schedule (Nota et al., 2004) to determine both the type and number of strategies students used. In a meta-learning question asked at the end of semester, students identified new strategies they had used during the course. This question was coded to determine how many new strategies students used, which indicated their adaptability.

To confirm the reliability of the coding schemes, a second naïve researcher coded 25% of the data for both goals and strategies, with inter-rater reliability scores of 87% and 89% respectively. These reliability scores are acceptable for coding self-regulated learning strategies (Nota et al., 2004).

2.2. Cluster analysis

A two-step cluster analysis was performed in SPSS^{TM} (IBM Statistics, NY, USA) using the learning analytics data (meta-learning task completion rate and time of submission, and the average number of words written for each question) and self-regulated learning behavior (type and number of learning strategies used, goals, and strategy adaptability). The two-step clustering procedure (Tkaczynski, 2017) involves the formation of pre-clusters, where

students within the same pre-cluster are treated as a single entity. In the second stage of the procedure, the pre-clusters are merged using a hierarchical clustering algorithm. The procedure generated a 3-cluster model when using either the Schwarz's Bayesian Criterion and Akaike's Information Criterion to determine optimal cluster number. The variables with the most influence over cluster formation included the type of goals students set, the number of planning strategies and the total number of learning strategies students reported using for the mid-semester exam (Figure 1).

The clustering procedure was validated by splitting the sample into two random groups and then repeating the procedure to ensure that the same number of clusters with similar characteristics were generated compared to the original cluster solution.
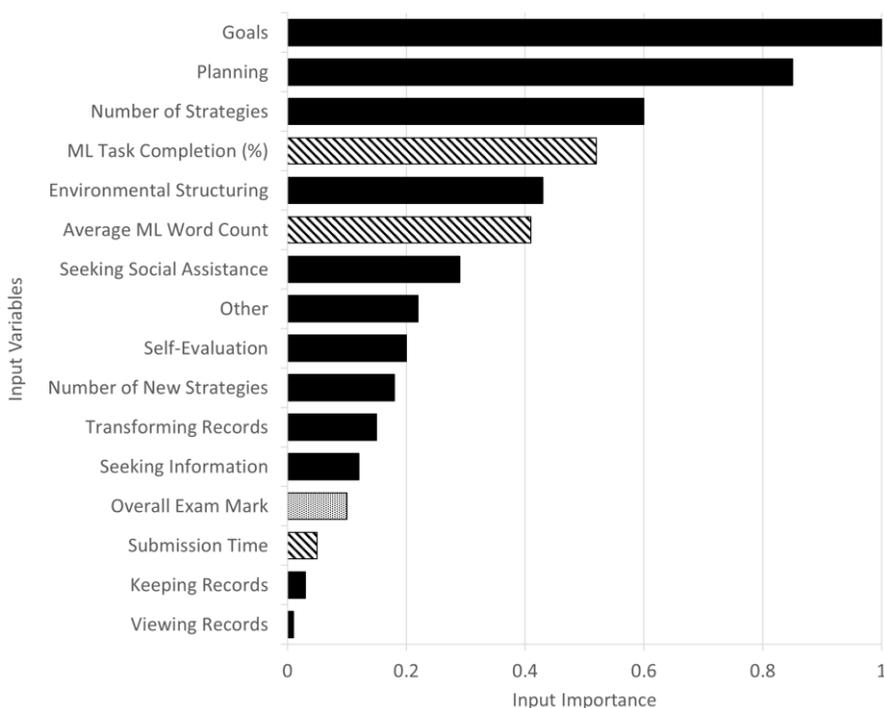

Figure 1. Predictor importance of variables in generating clusters. Solid bars represent self-regulated learning strategies. Bars with diagonal lines represent meta-learning (ML) task analytics. The dotted bar represents the students' exam mark.

3. Results

The two-step cluster analysis separated students into high, medium and low self-regulated learners (Table 1). High self-regulated learners (n=43) reported using the most learning strategies for exam preparation, and the most new strategies. They had the highest academic

performance (71.4% ± 2.3), wrote the most detailed responses for the meta-learning tasks, and had a high task completion rate. Low self-regulated learners (n=55) reported the lowest number of learning strategies, were the least likely to try new strategies and relied on mastery goals. These students were also the least engaged with the meta-learning tasks with the lowest word count for their responses, and the lowest completion rate. Medium self-regulated learners (n=72) set a combination of mastery and performance goals, used more strategies than low self-regulated learners and were willing to try new strategies. They had higher engagement with meta-learning tasks than low self-regulated learners, but had similar exam results (medium = 64.6% ± 1.7; low = 64.4% ± 2). High and medium self-regulated learners relied most heavily on self-evaluation and transforming records strategies, whereas low self-regulated learners relied on reviewing records (Figure 2).

Table 1: Characteristics of students clustered into high, medium, and low self-regulated learners. SEM = standard error of the mean, ML = Meta-learning task, M = mastery goals, P = performance goals, n.s. = non-significant.

	Low (1)	Medium (2)	High (3)	Significance
Overall exam mark (mean ± SEM)	64.4 ± 2	64.6 ± 1.7	71.4 ± 2.3	3>1,2*
ML word count (mean ± SEM)	62.9 ± 3.4	81.7 ± 3.6	95.8 ± 4.9	3>1, 2; 2>1*
ML submission time (mean hrs ± SEM)	60.1 ± 7	73.1 ± 5.9	79.5 ± 8.8	n.s.
Students completing all ML tasks	65%	100%	98%	2,3>1+
Goal type	75% M only	1% M only	60% M only	M only = 1,3>2; M+P = 2>1,3+
	11% P only	29% P only	12% P only	
	14% M+P	70% M+P	28% M+P	
Strategies for exam (mean # ± SEM)	2.1 ± 0.1	2.9 ± 0.1	3.8 ± 0.2	3 >1,2; 2>1*
New strategies (mean # ± SEM)	1.1 ± 0.1	1.6 ± 0.1	1.5 ± 0.1	2,3>1*

Statistical signifcance between low (1) medium (2) and high (3) self-regulated learners was determined using *ANOVA with Tukey's post hoc test or + chi-square test.

4. Discussion

In this study, a two-step cluster analysis was used to categories students based on their engagement with online meta-learning tasks, their self-regulated learning behavior and academic achievement. Three groups of students were identified – low, medium and high

self-regulated learners. The LMS data used to measure engagement with the meta-learning tasks included average word count for each meta-learning question, submission time for the meta-learning tasks and completion rate of the tasks. Word count and completion rate varied between clusters; however, submission time did not. Only 65% of low self-regulated learners completed all meta-learning tasks, whereas almost all medium and high self-regulated learners had a 100% completion rate. In addition, high self-regulated learners wrote more words on average for each question than medium self-regulated learners, who in turn wrote more than low self-regulated learners. As medium and low self-regulated learners had similar exam scores, these results suggest that engagement with the online meta-learning tasks does not relate to academic performance. In contrast, other studies have shown a correlation between LMS data, such as task submission time (Colthorpe et al., 2015) and assignment completion rate (Macfadyen & Dawson, 2010), with academic performance. As the meta-learning tasks in the current study were designed to prompt students to think about their learning, the high and medium self-regulated learners may have been more interested in the tasks, and completed them more thoroughly compared to low self-regulated learners. Future studies could include additional LMS data, such as completion of formative assessment (Tempelaar et al., 2018) or discussion board use (Macfadyen & Dawson, 2010) to facilitate the identification of non-engaged and low performing students.
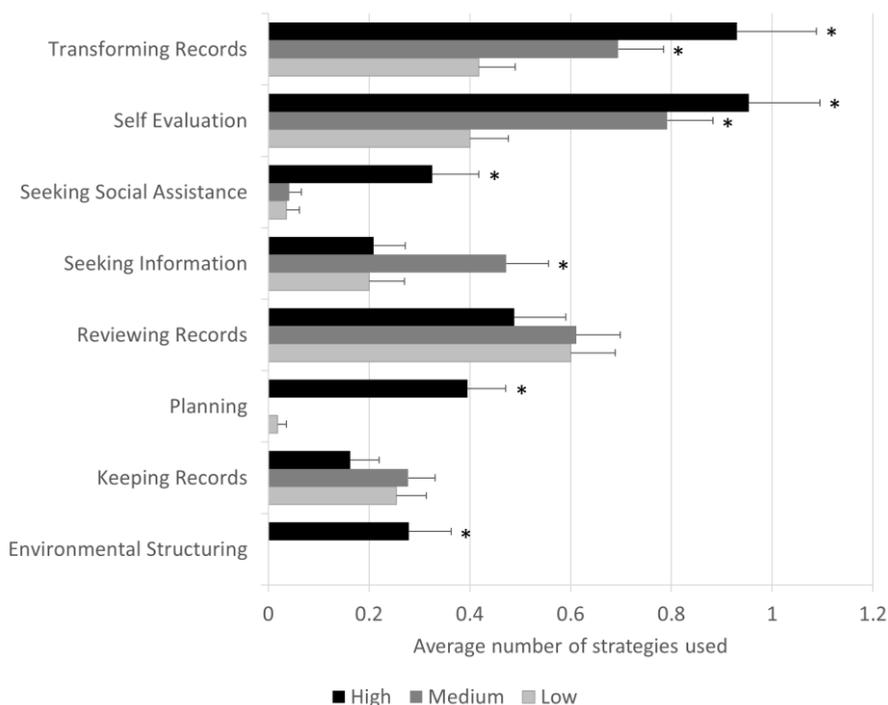
Figure 2: Self-regulated learning strategies used by students clustered as high (black bars), medium (dark grey bars) or low (light grey bars) self-regulated learners. Data represents the average number of strategies used by students in each cluster ± SEM. Asterisk indicates significant difference from low self-regulated learners using Welch's ANOVA with Games-Howell post hoc test.

Although low and medium self-regulated learners had similar academic achievement, they differed in their self-regulated learning strategies. High and medium self-regulated learners used more strategies than low self-regulated learners, a trait that has been associated with higher academic performance (Kitsantas, 2002). In particular, high and medium self-regulated learners used more strategies involving transforming records (E.g. Summaries, diagrams and flow charts) and self-evaluation (E.g. Practice tests and group study) compared to low self-regulated learners. These strategy categories have been identified as deep strategies, which are effective for learning (Hattie & Donoghue, 2016; Nota et al., 2004). Both high and medium self-regulated learners were also more likely to seek help either from social sources (seeking social assistance) or non-social sources (seeking information) respectively. High self-regulated learners were more likely to use planning and environmental restructuring than medium or low self-regulated learners. Although planning has been associated with high academic achievement (Hattie & Donoghue, 2016; Kitsantas, 2002), managing the environment is less effective (Hattie & Donoghue, 2016).

If the medium self-regulated learners used more strategies overall, and more deep strategies than low self-regulated learners, why was there no difference in their academic achievement? Perhaps the medium self-regulated learners were actively attempting to improve their learning approach, but had not yet achieved an improvement in grades. In support of this hypothesis, medium self-regulated learners incorporated more new strategies during semester compared to low self-regulated learners, suggesting that they were experimenting with their learning approach. The students in the current study were in their first year at university, and may need time to identify learning strategies that work well for them. In a study by Colthorpe et al. (2018), second year students who adapted learning strategies during semester had lower overall academic performance compared to students who did not adapt, but also demonstrated an improvement in grades over time. Together, these results suggest that incorporating new strategies can improve academic performance, but it might take multiple semesters for students to find approaches that work best for them. In the current study, medium self-regulated learners also set both mastery and performance goals, suggesting that they were attempting to improve both their understanding of the course content, and their course grades.

Learning analytics can be a powerful tool for monitoring student success and identifying students in need of support (Macfadyen & Dawson, 2010). In the current study, dispositional learning analytics was used to characterize students based on self-regulated learning behavior; however, the procedure was less successful at identifying students at different levels of academic achievement. Potentially, the medium and low self-regulated learners identified in this study reflect differences between clusters of students who are aware of their learning processes and are actively attempting to improve and students who are not (Ainscough et al. 2018). Examining the performance of these students at later stages of their program may provide evidence of the effectiveness of their approaches.

References

Ainscough, L., Stewart, E., Colthorpe, K., & Zimbardi, K. (2018). Learning hindrances and self-regulated learning strategies reported by undergraduate students: identifying characteristics of resilient students. *Studies in Higher Education, 43*(12), 2194-2209.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77-101.

Colthorpe, K., Ogiji, J., Ainscough, L., Zimbardi, K., & Anderson, S. (2017). Metacognitive assessment tasks provide insight into undergraduate pharmacy students' self-regulated learning behavior. *American Journal of Pharmaceutical Education*, ajpe6646.

Colthorpe, K., Sharifirad, T., Ainscough, L., Anderson, S., & Zimbardi, K. (2018). Prompting undergraduate students' metacognition of learning: implementing 'meta-

learning'assessment tasks in the biomedical sciences. *Assessment and Evaluation in Higher Education, 43*(2), 272-285.

Colthorpe, K., Zimbardi, K., Ainscough, L., & Anderson, S. (2015). Know thy student! Combining learning analytics and critical reflections to increase understanding of students' self-regulated learning in an authentic setting. *Journal of Learning Analytics, 2*(1), 134-155.

Gašević, D., Jovanović, J., Pardo, A., & Dawson, S. (2017). Detecting learning strategies with analytics: Links with self-reported measures and academic performance. *Journal of Learning Analytics, 4*(2), 113-128.

Hattie, J. A., & Donoghue, G. M. (2016). Learning strategies: a synthesis and conceptual model. *npj Science of Learning, 1*, 16013.

Kitsantas, A. (2002). Test preparation and performance: A self-regulatory analysis. *The Journal of Experimental Education, 70*(2), 101-113.

Macfadyen, L. P., & Dawson, S. (2010). Mining LMS data to develop an "early warning system" for educators. *Computers and Education, 54*(2), 588-599.

Nota, L., Soresi, S., & Zimmerman, B. J. (2004). Self-regulation and academic achievement and resilience: A longitudinal study. *International Journal of Educational Research, 41*(3), 198-215.

Shum, S. B., & Crick, R. D. (2012, 29 Apr - 02 May 2012). *Learning dispositions and transferable competencies: pedagogy, modelling and learning analytics.* Paper presented at the Proceedings of the 2nd International Conference on Learning Analytics and Knowledge, Vancouver, British Columbia, Canada.

Smith, V. C., Lange, A., & Huston, D. R. (2012). Predictive modeling to forecast student outcomes and drive effective interventions in online community college courses. *Journal of Asynchronous Learning Networks, 16*(3), 51-61.

Tempelaar, D., Rienties, B., Mittelmeier, J., & Nguyen, Q. (2018). Student profiling in a dispositional learning analytics application using formative assessment. *Computers in Human Behavior, 78*, 408-420.

Tempelaar, D. T., Rienties, B., & Nguyen, Q. (2017). Towards actionable learning analytics using dispositions. *IEEE Transactions on Learning Technologies, 10*(1), 6-16.

Tkaczynski, A. (2017). Segmentation using two-step cluster analysis. In T. Dietrich, S. Rundle-Thiele, & K. Kubacki (Eds.), *Segmentation in Social Marketing: Process, Methods and Application* (pp. 109-125). Singapore: Springer.