Received: 22 November 2022 | Accepted: 9 May 2023

DOI: 10.1111/bjet.13340

ORIGINAL ARTICLE

British Journal of Educational Technology BERA

Leveraging complexity frameworks to refine theories of engagement: Advancing self-regulated learning in the age of artificial intelligence

Jonathan C. Hilpert[1] | **Jeffrey A. Greene**[2] | **Matthew Bernacki**[2]

[1]Department of Educational Psychology Leadership and Higher Education, University of Nevada Las Vegas, Las Vegas, Nevada, USA

[2]Learning Sciences and Psychological Studies, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA

Correspondence
Jonathan C. Hilpert, Associate Professor and Director, Department of Educational Psychology Leadership and Higher Education, Office of Learning Analytics, University of Nevada Las Vegas, 4505 S. Maryland Parkway, Las Vegas, NV 89154, USA.
Email: jonathan.hilpert@unlv.edu

Funding information
National Science Foundation 1821601

Abstract

Capturing evidence for dynamic changes in self-regulated learning (SRL) behaviours resulting from interventions is challenging for researchers. In the current study, we identified students who were likely to do poorly in a biology course and those who were likely to do well. Then, we randomly assigned a portion of the students predicted to perform poorly to a science of learning to learn intervention where they were taught SRL study strategies. Learning outcome and log data (257 K events) were collected from $n = 226$ students. We used a complex systems framework to model the differences in SRL including the amount, interrelatedness, density and regularity of engagement captured in digital trace data (ie, logs). Differences were compared between students who were predicted to (1) perform poorly (control, $n = 48$), (2) perform poorly and received intervention (treatment, $n = 95$) and (3) perform well (not flagged, $n = 83$). Results indicated that the regularity of students' engagement was predictive of course grade, and that the intervention group exhibited increased regularity in engagement over the control group immediately after the intervention and maintained that increase over the course of the semester. We discuss the implications of these findings in relation to the future of artificial intelligence and potential uses for monitoring student learning in online environments.

KEYWORDS

learning analytics, network analysis, self-regulation

Practitioner notes

What is already known about this topic

- Self-regulated learning (SRL) knowledge and skills are strong predictors of post-secondary STEM student success.
- SRL is a dynamic, temporal process that leads to purposeful student engagement.
- Methods and metrics for measuring dynamic SRL behaviours in learning contexts are needed.

What this paper adds

- A Markov process for measuring dynamic SRL processes using log data.
- Evidence that dynamic, interaction-dominant aspects of SRL predict student achievement.
- Evidence that SRL processes can be meaningfully impacted through educational intervention.

Implications for theory and practice

- Complexity approaches inform theory and measurement of dynamic SRL processes.
- Static representations of dynamic SRL processes are promising learning analytics metrics.
- Engineered features of LMS usage are valuable contributions to AI models.

INTRODUCTION

Modern reform in postsecondary science, technology, engineering and mathematics (STEM) education has prioritized teaching not only declarative and conceptual knowledge, but also authentic STEM practices, dispositions and norms (Andrews et al., 2022). Such modern STEM education foci require moving beyond lecture-based instruction to embracing high-structure course designs with active learning pedagogies (Freeman et al., 2014; Lombardi et al., 2021; Theobald et al., 2020). Such designs and pedagogies can be far more beneficial than lecture-based instruction, but often students struggle to take advantage of these innovations (Miller & Bernacki, 2019) because they lack the self-regulated learning (SRL; Greene, 2018) knowledge and skill to thoughtfully use them (Greene et al., 2019). SRL knowledge and skills, involving the active pursuit of academic goals via metacognitive planning, monitoring, control and evaluation of the cognitive, motivational, behavioural and affect aspects of learning (Schunk & Greene, 2018), are strong predictors of postsecondary STEM student success (Theobald, 2021). SRL knowledge and skills can be taught (Bernacki et al., 2019), but existing evidence has focused mostly on such interventions' effects on learning outcomes (Cogliano et al., 2021) with less attention to how students' actual online engagement (Martin & Borup, 2022) changes. Such changes can be difficult to observe using traditional frequentist data gathering and modelling approaches (eg, Greene et al., 2019) because SRL is a dynamic, contingent, temporal process (Ben-Eliyahu & Bernacki, 2015) where increasing sophistication often leads to more organized and purposeful engagement (Zimmerman, 2013). To continue to iteratively evolve SRL theory and interventions (Greene, 2022), what is needed are methods of capturing SRL behavioural data over time and analytic tools that can measure how comprehensive, adaptive and organized students' SRL behaviours become as a result of intervention and growing experience (Hilpert & Marchand, 2018).

In this study, we utilized digital trace data (Bernacki, 2018) to capture, understand, intervene upon and model students' SRL knowledge and skills to better understand not just whether they engaged more after intervention, but also whether that engagement was in fact more organized and adaptive, and how those changes were associated with course performance. We used data from the first 2 weeks in the semester of a large introductory STEM course to predict student success and then deliver a targeted SRL intervention to students in need of support. Then, we continued to gather digital trace data throughout the course, which allowed us to model the dynamic and temporal aspects of SRL, which in turn proved to be a statistically significant and positive predictor of student success in the course and a difference in the outcome of intervention.

LEARNING ANALYTICS AND STUDENT SUCCESS

Although many forms of evidence can be gleaned from an LMS, one form that can be particularly useful is student log data (Cicchinelli et al., 2018; Lu et al., 2017). Log data provide an electronic history of digital traces of student interactions (ie, clicks) with course content (Bernacki, 2018). Log data can be used for various analytic applications such as prediction modelling (Cogliano et al., 2022), recommender systems (Fazeli et al., 2017) and data dashboards (Klerkx et al., 2017) that provide instructors and students with information about learners' engagement with digital content. Within the area of prediction modelling, logs of student click events within an LMS are often used to examine frequency of access to particular pieces of course content or elements of the LMS. These event data are typically organized into features of student learning behaviour including counts and patterns of usage of single items, or classes of items, merged with outcome variables and historical achievement data such as GPA or pretest scores, and used to train and test prediction models (Arizmendi et al., 2022; Baker et al., 2015; Brooks & Thompson, 2017). These models can be useful for identifying students likely to be at risk of a poor educational outcome, but are often not aligned with educational principles or theory in a meaningful or explainable way (Turek, 2018).

Making meaning of learning analytics

Knight and Shum (2017) described one of the foundational concepts of learning analytics as going from clicks to constructs. What they meant is that learning analytics frameworks necessitate epistemological assumptions about the use of a tool or data. Analytics can be used in ways that are theory-free and rely purely on machine derived patterns and predictions, or learning analytics can be aligned with educational principles so end users can interpret them in a meaningful way. For example, interpreting learning analytics from a SRL perspective has become a popular way to align LMS log data with an established theoretical construct. Cicchinelli et al. (2018) described a method for aligning analytics with SRL constructs by triangulating self-assessments with LMS log data. Winne and colleagues have crafted a careful body of work aligning log data within digital learning platforms with self-regulated engagement (eg, Winne, 2017a; Perry & Winne, 2006). Bernacki and colleagues have also accumulated a body of work aligning various features of technologically enhanced learning environments with SRL principles and models (Ben-Eliyahu & Bernacki, 2015; Bernacki, 2017; Bernacki et al., 2011). Other researchers have also used various types of network analysis to align SRL micro-processes with click events, including the use of transition networks (Siadaty et al., 2016) and the use of sequence mining (Saint et al., 2020).

These approaches have largely focused on studying specific combinations of click events or graph systems.

Hoppe (2017) described a trinity of network analytic methods for analysing LMS data that can be used to examine collaboration patterns, dynamic resource access and the evolution of communities and pathways through a learning platform. These types of network structures are what Hilpert and Marchand (2018) described as *interaction-dominant* ways to represent educational or learning phenomena (c.f. Lerner's notion of strong interactionist research). As opposed to a *component-dominant* approach such as sequence mining, which prioritizes operationalizing and measuring specific combinations or a priori psychological constructs, an interaction-dominant approach prioritizes the examination of dynamic relationships among components of a learning or educational system and interpreting them from a theoretical stance. Rooted in a complex systems epistemological tradition (Jacobson et al., 2019; Koopmans, 2020) interaction-dominant approaches seek to understand the emergent behaviour of complex combinations of elements within a system. Psychological phenomena are the product of large, complex systems composed of components that interact with one another. Research can concern itself with understanding the components or the interactions among the components and related structures that emerge from those interactions. In a component-dominant system, outcomes are the summed activity of the constituent parts. In an interaction-dominant system, outcomes are emergent and parts arrange themselves according to the changing demands of context (Dixon et al., 2012). Complexity frameworks can be used to examine interaction-dominant behaviours of a system. Complex systems are operationalized as networked interactions among agents in a system such as dynamic interactions among children in educational settings (see Jacobson et al., 2019 for a review of examples), or within-person phenomena such as the Web of complex interactions that underscore psychological constructs or dynamic patterns of behaviour (eg, Beymer et al., 2022; Epskamp et al., 2018; Fried & Robinaugh, 2020). Below, we describe our interaction-dominant approach to studying student engagement using click stream logs collected by an LMS system.

An interaction-dominant approach to student self-regulation in LMS

Interaction-dominant forms of student log data can provide rich information about SRL. We differentiate two different ways of making meaning of events that are recorded as student logs. The first is a component-dominant approach where the log of a click, or event, is interpreted in a defined way. For example, if a student clicks on the course calendar, that may be interpreted as a cognitive construct, such as planning for an upcoming event in class. In this example, it may be useful to count the number of times the student plans (or clicks on the calendar) to determine whether it is useful in some way, such as to predict student outcomes. This example is component-dominant because the log itself is the object of study. This has been a predominant approach in well understood learning contexts where individual objects are provided to students by an instructor and researcher during co-design (Lockyer et al., 2013), and inferences can be made that a certain click represents a *learning event* that can be classified based on the way the target object affords the learner an opportunity to engage in a cognitive or metacognitive process (Bernacki, 2018).

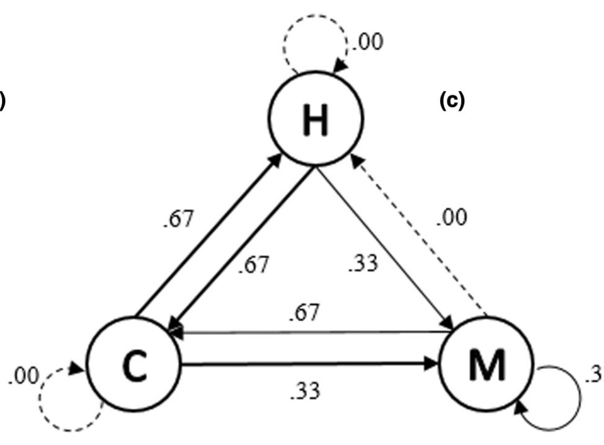
The second is an interaction-dominant approach, where the relationships among the clicks are interpreted in a defined way. For example, if a student clicks on the course calendar (ie, to plan one's study session), then on the course homepage (ie, to find resources that enable them to enact their plan) and then on a practice quiz (ie, to engage in retrieval practice), a triadic structure emerges among the clicks. This is an example of an interaction-dominant approach because the relationships between the logs are the object of study. Although we use these three example clicks to illustrate how one possible structure may emerge, the specific components that form the structure are not fixed in our analytics approach and thus not component-dominant. Component-dominant approaches treat a click as having a pre-determined meaning and view variance in the meaning of the same behaviour as problematic. This provides little recourse for establishing the validity of the inference beyond the a priori affordances for the behaviour. In contrast, our interaction-dominant approach considers a click as part of a system that accounts for contextually different meanings of the same click when it appears in different emergent structures and different relations to other clicks. This flexibility allows for a more valid inference about its meaning to student self-regulation.

Interaction-dominant objects of study are often examined using intensive (ie, many observations closely spaced in proximity) forms of time series and network data (Wallot & Kelty-Stephen, 2018). Intensive forms of data provide unique information about micro-genetic change, or studying change as it occurs, that cross-sectional or repeated measures data cannot (Hilpert & Marchand, 2018). In the case of network data, nodes describe the components of the LMS landscape (eg, clicking on the calendar) and edges describe the relationships between the components (eg, navigating from the calendar to the home page). A time series of student logs can be transformed into a network using a Markov approach where the node describes the components of the LMS (eg, calendar, homepage, etc) and the edges describe the probability of clicking from one component to the next (eg, the likelihood of clicking from the calendar to the home page).

Figure 1 provides a simplified graphical representation of how a time series of discrete LMS click events for an individual student can be transformed into a probability transition matrix and then represented as a network. The figure provides a streamlined example of three LMS components, a homepage, a calendar and a module. In this type of network, the *number of nodes* for a given student describes the components of the LMS they

TimeStamp	LMS_Component
1:00:00 PM	Homepage
1:01:00 PM	Calendar
1:02:00 PM	Module
1:03:00 PM	Module
1:04:00 PM	Calendar
1:05:00 PM	Homepage
1:06:00 PM	Module
1:07:00 PM	Calendar
1:08:00 PM	Homepage
1:09:00 PM	Calendar

(a)



(c)

	Calendar	Homepage	Module
Calendar	0.00	0.67	0.33
Homepage	0.67	0.00	0.33
Module	0.67	0.00	0.33

(b)

FIGURE 1 Graphical representation of LMS click event network model. Panel (a) contains example time series of click events for an individual student. Panel (b) contains the first-order probability transition matrix for the time series (ie, the probability of transitioning from one event to another based on the observations). Panel (c) contains the graphical representation of the probability matrix.

have clicked on (nodes = 3 for Figure 1). The *number of edges* describes the number of transitions between the components (edges = 6 for Figure 1). The *density* of the network describes the portion of the potential transitions that were actualized (density = 0.83 for Figure 1). And, the *transitivity* of the network describes the overall probability of having adjacent nodes interconnected (ie, in triangular form; transitivity = 0.75 for Figure 1). We provide a technical description of calculating these types of networks in the Methods section.

For this study, we posited that an interaction-dominant approach to digital trace data would capture changes in students' SRL processing as a result of intervention and over time. Specifically, SRL theory does not suggest a single, prescribed set of useful or not-useful digital tools, nor does it dictate a single sequence or pattern of engagements with those tools (Winne & Hadwin, 1998). Rather, SRL processing is necessarily the result of context-, person- and task-specific factors that result in often idiosyncratic patterns and types of engagement (Efklides, 2011; Greene & Azevedo, 2009). As but one example, students with little time to study due to economic pressures to work (ie, context-specific factor), who lack relevant prior knowledge (ie, person-specific factor), may have to engage with summative digital tools (eg, downloaded PowerPoint slides from lectures) to best optimize their task engagement, whereas students in different circumstances might have more time and prior preparation to engage with more substantive tools (eg, practices quizzes). Thus, the optimal choice of digital tool and the optimal nature of engagement with that tool can and often should vary from student to student (Efklides, 2011), resulting in different patterns of engagement with any particular tool across students, which can obfuscate relations with course outcomes. As such, rather than focusing on which tools were used how often (ie, component-dominant approach), we examined the number, types and regularities of engagement across these components (ie, interaction-dominant approach). Specifically, we posited that the number of nodes might represent how *comprehensive* students were in engaging with the suite of digital tools available, the number of connections between tools might capture the degree to which students *dynamically* engaged with tools, and the density of the network might indicate how *actively* students were in engaging with the suite of digital tools available, and transitivity might indicate the degree to which students engaged in a *cyclical, intentional* manner with digital tools. These metrics, if demonstrated to have predictive validity with expected outcomes (Dent & Koenka, 2016), such as course grade, would argue for the utility of using network statistics to capture not just what students engage with in LMSs, but also the nature of that engagement (ie, comprehensive, dynamic, active, cyclical, intentional), which is a fundamental aspect of SRL theory (Ben-Eliyahu & Bernacki, 2015; Efklides, 2011).

Purpose of the current study

For the current study, we transformed intensive time series log data taken from individual students enrolled in a postsecondary biology course into first-order probability transition networks using a Markov approach. Network statistics were calculated for each student, including the number of nodes, the number of edges, network density and network transitivity. Our research questions were: (1) does regularity in click events account for a significant amount of variance in students' final grade in the course?, (2) does regularity in students' click events increase over the course of the semester? and (3) did students who received treatment (ie, a science of learning to learn intervention) and who were predicted to perform well in the course show greater regularity of click events after intervention, and greater increases in regularity of click events over the semester, than students who were predicted to perform poorly and did not receive treatment?

METHOD

Sample

The sample contained 257K logged events produced by 226 students who were enrolled in sections of an introductory biology course and who consented to share their data and receive correspondence from the research team. Students ranged in age from 18 to 72, with 82% of the sample falling between the ages of 18 to 21 years old. According to the university classification system, 73% of the students were female, 25% were Asian, 13.3% were Black, 30% were Hispanic, 11% were Multiracial, 1% were Pacific Islander, and 19% were white. 10% of the students received a failing grade in the course, 15% received a D, 33% received a C, 33% received a B, and 9% received an A. Students were classified[1] into control ($n = 48$), treatment ($n = 95$) and not flagged ($n = 83$) using a combination of a pretest score and counts of resources accessed in the course which were submitted to a forward selection logistic regression with a 10-fold cross-validation process. Full description of the prediction model approach is outside the scope of the current study and but can be found elsewhere (Cogliano et al., 2022).

Measures

Event logs

Student log data were extracted from the university data lake that houses records of all student LMS data using SQL queries written in SQL developer. The events were extracted as system-driven, web-application controller descriptions of each student click event. The logs were classified into 19 different categories. These categories included various online learning behaviours including viewing the home page; viewing course notifications; viewing and submitting assignments; viewing and submitting quizzes; viewing, submitting and responding to discussions; viewing and downloading course files; viewing course resources such as the gradebook, calendar and syllabus; and using external course tools.

Final course grade

Students' final course grade was extracted from the electronic gradebook housed in the course LMS. Students' final grade represented the culmination of their effort in the course, including one final examination, four semester examinations, weekly quizzes and regular homework. The first two unit-examinations consisted of approximately 45 multiple-choice questions plus five short-answer questions. The latter two unit examples consisted of approximately 25 multiple-choice questions plus five short-answer questions. All the examination questions were aligned with learning objectives reflected in the course readings and instructional materials, indicating good content validity.

Self-efficacy

A self-efficacy scale was used to measured students' beliefs about their ability to be successful in their current STEM course using five items drawn from the Patterns of Adaptive Learning Scale (Midgley et al., 2000). The items were measured on a six-point Likert-type scale from strongly disagree to strongly agree (sample item: 'I'm certain I can master the skills taught in this course'). The self-efficacy scale was administered twice during the semester, once during the first week of class and then again during the tenth week of class. Items at both time points exhibited good evidence of internal reliability with alphas >0.80.

Markov process calculation

For each individual student, a Markov process was calculated from the time series of their click events. A *Markov process*$\{X_t\}$is a stochastic process, whereby the values of $\{X_{t+1}\}$ are only influenced by the present value. The probability of any future behaviour of the process when the current state is known is not influenced by any additional knowledge concerning its past behaviour. A discrete chain Markov process contains a knowable or finite set of states whose time index set is $T = (0, 1, 2, \ldots)$. In formal terms, the Markov property is that

$$\Pr\{X_{n+1} = j \mid X_0 = i_0, \ldots, X_{n-1} = i_{n-1}, X_{n=\mathrm{i}}\} \\ = \Pr\{X_{n+1} = j \mid X_{n=i}\}$$

For all time points n and all states $i_0, \ldots, i_{n-1}, i, j$.

The probability of $X_{n=1}$ being in state j given that X_n is in state i is called the first-order transition probability and is denoted by $P_{ij}^{n,n+1} = \Pr\{X_{n+1} = _j | X_{n=i}\}$ where P_{ij} is the conditional probability of undergoing a transition from i to j in one step (Pinsky & Karlin, 2010). The probability calculations are typically presented in a matrix form and converted into network diagrams for visualization, as illustrated in Figure 1.

For each student event log, the click sequence was transformed into a first-order probability transition matrix using function markovchainFit from the R package markovchain (Spedicato et al., 2021). This function can be used to calculate a first-order probability transition matrix from a sequence of states in a time series. The first-order probability matrix is a square adjacency matrix that contains the probability of transitioning from one LMS component to another. The probability transition matrix for each student was transformed into a network object using the network function from the R package statnet (Handcock et al., 2019). For each student network, we calculated the number of nodes (using the network. size function), the number of edges (using the network.edgecount function), the network density (using the gden function) and the network transitivity (using the gtrans function). The network statistics were then windowed into four time points spanning the semester, separating the semester into four equal portions (excluding final examination week). The resulting data set contained network statistics describing the probability transition matrix for each student (overall and at four time points) as well as their final course grades and measures of self-efficacy.

Statistical analysis

Descriptive statistics and bivariate correlations were examined and the evidence suggested that statistical assumptions were met. To answer research question 1, a multiple linear regression (ie, regressing final course grade on the overall network statistics) was run using the psych package in R (Revelle, 2019). Further assumption checking was conducted using the results of the regression. To answer research questions 2 and 3, a latent growth curve model was analysed using the lavaan package in R, using the growth function designed for analysing mean structures underlying linear growth. Growth in network transitivity was examined across four time points over the course of the semester. The data were analysed using full information maximum likelihood (FIML) to maintain the representativeness of the sample (Beaujean, 2014). Goodness-of-fit indices were evaluated using criteria recommended by Hu and Bentler (1999), including the comparative fit index (CFI > 0.95), the root-mean-square error of approximation (RMSEA < 0.06) and the standardized root-mean-square residual (SRMR < 0.08). We expected the three conditions to differ in their final course grade such that, on average, the control group would have the lowest scores, the treatment group would have the next highest scores, and the not flagged group would have the highest scores. In our latent growth model, our intervention variable was dummy coded such that the control group was coded 00, the treatment group was coded 01 and the not flagged group was coded 10, creating two dummy variables that could be entered into the model to compare group differences in the intercept and slope of linear increase in transitivity. Self-efficacy was included in the model as a time varying covariate to account for individual differences in variation in click events not predicted by the intercept or the slope. Figure 2 contains a graphical depiction of the latent growth curve model we tested.

RESULTS

Descriptive statistics are provided in Table 1. The descriptive statistics provide basic engagement information about student log events. Students navigated to an average of 14.7 out of the 19 LMS components in the course. There was an average of 75.9 edges connecting the course components, and an average of 34.7% of the possible relationships between the components navigated, and an average of 52% of the possible triadic relationships between

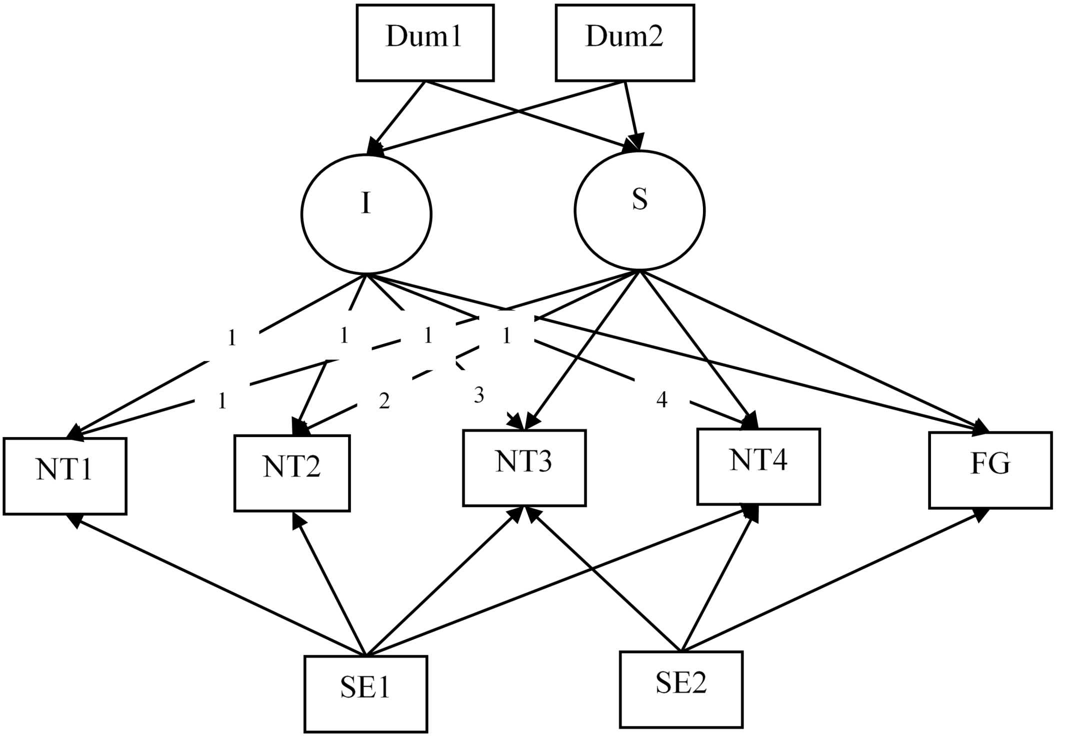


FIGURE 2 Graphical representation of hypothesized latent growth curve model. Dum1 = dummy variable with not flagged group coded as 1, Dum2 = dummy variable with treatment group coded as 1. NT = network transitivity as measured by transitivity (NetTrans) in Markov networks at four time points for latent growth. I = intercept. S = Slope. SE = self-efficacy at two time points entered as a time varying covariate. FG = final grade.

TABLE 1 Summary of descriptive statistics for model variables.

Variable	*N*	Missing	*M*	*SD*	Min	Max
NetSize	226	0	14.74	1.63	3.00	19.00
NetEdge	226	0	75.85	18.48	4.00	141.00
Netden	226	0	0.35	0.06	0.20	0.52
NetTrans	226	0	0.53	0.09	0.00	0.70
NetTrans.1	226	0	0.31	0.14	0.00	1.00
NetTrans.2	225	1	0.35	0.11	0.00	0.57
NetTrans.3	221	5	0.34	0.12	0.00	1.00
NetTrans.4	213	13	0.37	0.13	0.00	0.66
T1SelfEff	195	31	5.11	0.79	1.00	6.00
T2SelfEff	203	23	4.86	0.80	1.00	6.00
FinalGrade	198	28	74.67	16.17	0.90	96.10

the components navigated. The average final grade in the course was 74.6%. Table 1 provides a breakdown of descriptive statistics for all variables included in the statistical models.

Bivariate correlations are provided in Table 2. Analysis of statistical significance patterns revealed that final grade was positively correlated with network size, number of edges and network transitivity. There were also statistically significant and positive correlations among the network statistics, none with a magnitude that would suggest collinearity. Self-efficacy at time 2 was positively correlated with network transitivity at times 3 and 4. When interpreted as evidence of interaction patterns reflecting self-regulated learning activity, students who engage with greater proportions of the content (network size), and did so in ordered patterns involving multiple accesses of unique objects in sequence (edges, transitivity) were more likely to perform well in the course.

A multiple regression was conducted to evaluate how well the network statistics predicted final course grade. The predictors were the four network statistics. The linear combination of the predictors was statistically significantly related to final course grade ($R^2 = 19.00$, $F(4, 226) = 11.308$, $p < 0.001$). Approximately 19% of the variation in final course grade in the sample was accounted for by the predictors. See Table 3 for a summary of estimates. Network transitivity was a statistically significant and positive predictor of final grade over and above the other network measures.

A latent growth curve model was calculated to examine the effects of the intervention, as mediated by Unit Exam Achievement. The model included measures of network transitivity at four time points (ie, NT1-NT4, see Figure 2). Latent growth curve modelling uses these measures as indicators of both a latent intercept factor (ie, an estimate of the mean network transitivity score at time 1, after disaggregating error) and a latent slope factor (ie, an estimate of linear change in network transitivity scores over the four time points). These two latent factors were regressed on dummy-coded group variables, to investigate potential differences in the intercept and/or slope across the three groups. Self-efficacy scores, at two time points, were used as control variables. Finally, final grade scores were regressed on self-efficacy, as a control, and the intercept and slope latent factors, to investigate whether network transitivity scores predicted academic performance.

The chi-squared test for the model was not statistically significant $\chi^2(15, 226) = 24.27$, $p = 0.061$, indicating the model was consistent with the covariance structure of the variables. Examination of fit indices suggested the model demonstrated evidence of good fit: CFI = 0.92, RMSEA = 0.06, SRMR = 0.05. Table 4 provides a summary of the coefficients from the model. Results suggested that variation in the intercept of transitivity was

TABLE 2 Bivariate correlations among all variables included in statistical models.

Variable	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.
NetSize	–										
NetEdge	0.713***	–									
NetDen	−0.163*	0.557***	–								
NetTrans	0.400***	0.747***	0.658***	–							
NetTrans.1	0.233***	0.341***	0.205**	0.347***	–						
NetTrans.2	0.304***	0.455***	0.335***	0.567***	0.133*	–					
NetTrans.3	0.086	0.235***	0.224***	0.296***	−0.033	0.234***	–				
NetTrans.4	0.192**	0.356***	0.308***	0.472***	0.127	0.393***	0.362***	–			
T1SelfEff	0.105	0.058	−0.02	0.026	0.105	0.071	−0.12	0.054	–		
T2SelfEff	−0.005	0.089	0.139*	0.103	0.055	−0.007	0.178*	0.166*	0.325***	–	
FinalGrade	0.288***	0.296***	0.12	0.386***	0.225***	0.074	0.229**	0.413***	0.075	0.233**	–

*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

TABLE 3 Summary of linear regression results for final grade.

Model		Unstandardized	Standard error	Standardized	*t*	*p*
H_0	(Intercept)	74.671	1.149		64.964	<0.001
H_1	(Intercept)	−64.644	61.171		−1.057	0.292
	NetSize	7.749	4.252	0.767	1.823	0.07
	NetEdge	−0.672	0.408	−0.727	−1.648	0.101
	NetDen	127.284	105.317	0.453	1.209	0.228
	NetTrans	59.892	23.385	0.336	2.561	0.011

TABLE 4 Summary of latent growth curve parameter estimates.

						95% confidence interval		Standardized		
Predictor	Outcome	Estimate	Std. error	*z*-value	*p*	Lower	Upper	All	LV	Endo
DUM1NF	i	0.187	0.059	3.173	0.002	0.072	0.303	4.861	9.84	9.84
DUM2T	i	0.13	0.057	2.29	0.022	0.019	0.241	3.324	6.82	6.82
i	FinalGrade	132.363	35.259	3.754	<0.001	63.256	201.47	0.2	2.521	0.2
s	FinalGrade	161.939	42.527	3.808	<0.001	78.587	245.291	0.861	10.869	0.861
T2SelfEff	FinalGrade	3.782	1.048	3.61	<0.001	1.728	5.835	0.239	3.782	0.3
T1SelfEff	NetTrans.1	0.019	0.014	1.39	0.165	−0.008	0.046	0.125	0.019	0.158
	NetTrans.2	0.013	0.008	1.673	0.094	−0.002	0.029	0.106	0.013	0.133
	NetTrans.3	−0.021	0.009	−2.18	0.029	−0.039	−0.002	−0.144	−0.021	−0.181
T2SelfEff	NetTrans.3	0.031	0.01	3.2	0.001	0.012	0.05	0.218	0.031	0.274
T1SelfEff	NetTrans.4	0.003	0.015	0.228	0.819	−0.026	0.032	0.023	0.003	0.029
T2SelfEff	NetTrans.4	0.022	0.011	1.92	0.055	−4.48E-04	0.044	0.149	0.022	0.187
DUM1NF	s	−0.093	0.043	−2.194	0.028	−0.177	−0.01	−0.686	−1.39	−1.39
DUM2T	s	−0.063	0.042	−1.482	0.138	−0.146	0.02	−0.455	−0.934	−0.934

statistically significant and positively related to both dummy variables, suggesting that (most notably) the treatment group had higher initial regularized click events than the control group immediately after intervention. Results also suggested that dummy variable 1 was statistically significant and negatively related to the slope of transitivity, indicating that the regularity of clicks increased at a lower rate in the not flagged group, compared with the control group. The intercept and slope of transitivity were statistically significant and positively related to final grade, as was self-efficacy at time 2.

DISCUSSION

In this study, we responded to members of the learning analytics and SRL community's challenge to adopt more complex learning analytics approaches to move beyond the component-dominant modelling that relies on single accesses of digital tools to reflect traces of the self-regulated learning process, for example Winne (2017a, 2017b, 2020). Rather than rely on the single learning events using digital objects that might be inferred to afford information appropriate to planning, strategy use or monitoring judgements, we adopted a more global, interaction-dominant view of learners' interactions across components and relied on complexity approaches and network statistics to reflect students' patterned behaviours that might reflect their ability to self-regulate over a longer period where the digital objects that are relevant to the immediate task shift over the many weeks of a university course.

Network statistics including the edges that indicate use of ordered pairs of content and the transitivity statistic that demonstrates an organized access of three unique objects provide traces of increasingly complex patterns that do not reflect specific SRL phases but rather engagement in the more general SRL process that affords looseness in its sequence and order to accommodate the task conditions and conditions that the learners' own affordances and constraints impose on their task engagement.

The results of our analysis of these network statistics and their relationship to the academic achievement known to correlate to engagement in SRL (eg, Dent & Koenka, 2016) provide initial evidence that this novel approach to tracing SRL through an interaction-dominant paradigm supports theoretical assumptions about SRL. Network-based approaches to deriving features from log data were useful for predicting student outcomes. Moreover, the metrics associated with the networks can be interpreted from a self-regulation perspective that has practical implications for artificial intelligence efforts in learning analytics, as well as for interrogating how the digital tools designed to improve SRL and found to benefit academic achievement might achieve such effects (eg, Bernacki et al., 2020; Cogliano et al., 2022). Static representations of dynamic processes derived from the markov approaches utilized in this study (ie, transitivity calculated for probability transitions), or other statistical approaches that rely upon time intensive data to capture dynamic learning processes (Asparouhov et al., 2018), are extremely useful for advancing theory, predicting student success and examining the outcome of interventions. There have been recent calls in the education literature for increased study of motivation and engagement dynamics (Pekrun & Marsh, 2022); however, responding to these calls requires methods for aggregating intensive data into meaningful metrics that capture dynamic psychological processes with fidelity and in ways that are scalable.

Implications for prediction modelling: Network statistics and their potential for feature engineering

Network statistics predicted variance in student achievement and can serve as an additional class of engineered features that capture a unique aspect of the student learning experience that can explain performance. What these patterns reflect is a more abstract construct compared with the learning events that are the focus of most learning analytics-driven SRL research (Bernacki, 2018). Unlike other multi-event methods that tend to be agnostic to learning theory and simply combine events via sequence-, pattern-, process- or association-mining approaches, these network statistics are agnostic to the components that comprise such sequences or patterns, and instead describe the act of engaging in an ordered pattern of behaviour. This ordering is itself a trace of something that reflects a regularized order process rather than specific individual sequences. In many cases, data-driven solutions identify many sequences or patterns involving components that contribute to variance explained in an outcome, but they are difficult for researchers to interpret, or even determine whether they should be interpreted. This decision to explain the AI solution would require that researchers put forth the effort to collect some form of corroborating evidence of a learner's process, and that those efforts be brought to bear to validate assumptions about what such patterns of resource use might reflect vis-a-vis SRL theory (Winne, 2020). Such work is tremendously time intensive, and even when researchers do generate data to corroborate learning events and event sequences, they are not necessarily captured in sufficiently rich detail or temporal proximity to fully validate inferences about the digital event. Opportunities to validate sequences of events are even more uncommon, as the uniqueness of a 3-component pattern of events happens so infrequently that it is nearly impossible to collect enough concurrent corroborating data to even consider whether a valid inference can be drawn. Because counts of edges and transitivity are interaction-dominant methods and not component-dominant ones, they serve as more general traces of one's tendency to engage with multiple digital objects meant to afford different SRL processes. Because SRL frameworks acknowledge that the most appropriate pattern of SRL processes is dictated by a complex combination of learner and task affordances, progress towards a learning goal and implications of events that have preceded that exact moment, there is no specific SRL process that is the correct one for all learners in all moments. Thus, instead of reflecting a non-existent 'right SRL move', network statistics reflect one's tendency to engage in a 'next SRL move' by engaging with more content, and one's tendency to do so over the learning task. As such, network statistics that capture use of digital objects meant to afford specific SRL processes are themselves theory-aligned traces that reflect a unique layer of the SRL process and should predict achievement over and above the specific processes captured by event-based predictors that typically comprise feature sets in prediction models (eg, Arizmendi et al., 2022). In addition to explaining unique variance over and above these theory-aligned event-based traces, the occurrence of these theory-aligned interaction-based traces can also serve as phenomena that can explain how digital tools meant to support SRL obtain their effects on performance outcomes.

Implications for self-regulated learning support

Bernacki and colleagues (2020) and Cogliano and colleagues (2022) deployed digital science of learning to learn training programmes to students and found that students who completed them performed better on subsequent examinations in their respective biology courses. Using a component-dominant learning analytics approach, Bernacki and colleagues (2020) identified increases in students' use of resources for planning in the first week of a subsequent unit, greater initial use of tools for self-assessment in the same period, and greater use of resources designed to support cognitive strategies like retrieval practice. In this study, we re-analysed the sample observed by Cogliano and colleagues (2022) to benefit from a 15-minute version of the science of learning to learn intervention and here found that those who an algorithm classified as unlikely to require learning support and who also received and completed the learning support were more apt to engage in ordered patterns of learning behaviour reflecting the SRL cycle. Importantly, learners who initially behaved like the group who was flagged as likely to obtain a poor course grade based on their use of individual use of learning objects in the first weeks of the course eventually changed something about their learning that led them to outperform their peers on all subsequent examinations and to earn better grades in the course. Whereas the behaviours that earned them their flagged classification were initially the same as the flagged control group, growth curve analyses examining network statistics provided evidence that the loosely sequenced, self-regulatory aspects of their click behaviour (ie, transitivity) were increased immediately after the intervention and remained higher than the control group during the four windows of time we identified over the course of the semester. High initial transitivity for the treatment and non-flagged groups (ie, the intercept in the growth curve model) and linear increase in transitivity over time (ie, the slope) were both significantly and positively related to final grade, even when controlling for student self-efficacy, further suggesting that regularized click events reflective of SRL are likely a valuable engineering feature of AI models meant to support student learning. However, one limitation of the study is that although the within-person analysis of the SRL longitudinal complex system was interaction-dominant, the system was in some ways taken apart to examine treatment effects. In this way, the study needed to balance feasibility with complete adherence to a complex systems point of view.

Despite this limitation, the results suggest the use of network statistics affords a new opportunity to not only look at the course-specific, time-bound objects that students should use in particular weeks and months after intervention to plan and study in a next unit of course content, but also provides an enduringly relevant metric that describes a pattern of behaviour reflecting ordered engagement in a complex learning environment. In this way, effects on a course performance metric might be explained by this more sophisticated learning behaviour, and the sustained engagement in such behaviour may become a critical outcome of a digital learning skill training itself. If students can be taught to engage in ordered learning behaviours and to flexibly consider an evolving learning task and how they should engage in it, this may become an important dependent variable by which the value of a science of learning to learn tool can be appraised. Transitive relationships can be interpreted as interaction-dominant forms of SRL, showcasing the ways in which a complexity approach can be combined with SRL theory to produce useful analytics metrics that have practical value for modelling and student support.

FUNDING INFORMATION

This material is based upon work supported by the National Science Foundation under grant no. 1821601.

CONFLICT OF INTEREST STATEMENT

The authors have no known conflict of interest to disclose.

DATA AVAILABILITY STATEMENT

Data will be provided to interested parties via correspondence with the first author until the funding and dissemination efforts for the project are complete, at which point the data will be hosted for public access.

ETHICS STATEMENT

All data collected as a part of the project were gathered under IRB approval and in accordance with federal ethics guidelines described by OHRP and the common rule.

ORCID

Jeffrey A. Greene https://orcid.org/0000-0003-4145-1847
Matthew Bernacki https://orcid.org/0000-0003-1279-2829

ENDNOTE

[1] In order to examine the effects of interventions, and the ways students engaged with them, we needed to recruit a sufficient number of participants into the study that they could be observed to engage in a pattern of behaviour that, when modelled in our algorithm, predicted they would succeed and should not be flagged, or should be flagged as potentially likely to benefit from intervention. There were two levels of consent included in our recruitment. Students could choose to share data only, or share data and receive feedback and support if their data suggested they might benefit. For those who chose only to share data, they were given a biology learning activity that covered biology topics pertinent to the first course unit and examination. This is true for both those who shared data and whose data indicated they should be flagged or non-flagged. This is also the same activity that those who consented to receive feedback and support would receive if they were not flagged, or if they were randomly assigned into the control condition. Because we knew that there would be individuals who would be willing to share data but were not inter-

ested in participating in the experiment in the study, we also knew that these students' data could be considered as additional cases that did not differ from those flagged and randomly assigned to the control condition where they complete the biology activity. We thus oversampled those who consented to the second part of the study into the experimental condition where we offered some students a digital training focused on improving their learning skills and intervention. Those eligible for such modelling include all who consented to share their data. This larger sample size could power not only main effects when compared to the control group (comprising flagged+randomly assigned Level 2 consenters and flagged+Level 1 consenters), but also within this experimental treatment group to examine interactions where different learner groups may benefit differently.

REFERENCES

Andrews, T. C., Speer, N. M., & Shultz, G. V. (2022). Building bridges: A review and synthesis of research on teaching knowledge for undergraduate instruction in science, engineering, and mathematics. *International Journal of STEM Education*, *9*(1), 1–21.

Arizmendi, C. J., Bernacki, M. L., Raković, M., Plumley, R. D., Urban, C. J., Panter, A. T., Greene, J. A., & Gates, K. M. (2022). Predicting student outcomes using digital logs of learning behaviors: Review, current standards, and suggestions for future work. *Behavior Research Methods*, 1–29.

Asparouhov, T., Hamaker, E. L., & Muthén, B. (2018). Dynamic structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal*, *25*(3), 359–388.

Baker, R. S., Lindrum, D., Lindrum, M. J., & Perkowski, D. (2015). *Analyzing early at-risk factors in higher education E-learning courses*. International Educational Data Mining Society.

Beaujean, A. A. (2014). *Latent variable modeling using R: A step-by-step guide*. Routledge.

Ben-Eliyahu, A., & Bernacki, M. L. (2015). Addressing complexities in self-regulated learning: A focus on contextual factors, contingencies, and dynamic relations. *Metacognition and Learning*, *10*(1), 1–13.

Bernacki, M. L. (2017). Examining the cyclical, loosely sequenced, and contingent features of self-regulated learning: Trace data and their analysis. In *Handbook of self-regulation of learning and performance* (pp. 370–387). Routledge.

Bernacki, M. L. (2018). Examining the cyclical, loosely sequenced, and contingent features of self-regulated learning: Trace data and their analysis. In D. H. Schunk & J. A. Greene (Eds.), *Handbook of self-regulated learning and performance* (pp. 370–387). Routledge.

Bernacki, M. L., Aguilar, A. C., & Byrnes, J. P. (2011). Self-regulated learning and technology-enhanced learning environments: An opportunity-propensity analysis. In *Fostering self-regulated learning through ICT* (pp. 1–26). IGI Global.

Bernacki, M. L., Greene, J. A., & Crompton, H. (2020). Mobile technology, learning, and achievement: Advances in understanding and measuring the role of mobile technology in education. *Contemporary Educational Psychology*, *60*, 101827.

Bernacki, M. L., Vosicka, L., & Utz, J. C. (2019). Can a brief, digital skill training intervention help undergraduates "learn to learn" and improve their STEM achievement? *Journal of Educational Psychology*, *112*(4), 765–781. https://doi.org/10.1037/edu0000405

Beymer, P. N., Benden, D. K., & Sachisthal, M. S. (2022). Exploring the dynamics of situated expectancy-value theory: A panel network analysis. *Learning and Individual Differences*, *100*, 102233.

Brooks, C., & Thompson, C. (2017). Predictive modelling in teaching and learning. In C. Lang, G. Siemens, A. Wise, & D. Gasevic (Eds.), *Handbook of learning analytics* (pp. 61–68). Society for Learning Analytics and Research (SOLAR).

Cicchinelli, A., Veas, E., Pardo, A., Pammer-Schindler, V., Fessl, A., Barreiros, C., & Lindstädt, S. (2018, March). Finding traces of self-regulated learning in activity streams. In *Proceedings of the 8th International Conference on Learning Analytics and Knowledge* (pp. 191–200). ACM Digital Library.

Cogliano, M., Bernacki, M. L., & Kardash, C. M. (2021). A metacognitive retrieval practice intervention to improve undergraduates' monitoring and control processes and use of performance feedback for classroom learning. *Journal of Educational Psychology*, *113*(7), 1421.

Cogliano, M., Bernacki, M. L., Hilpert, J. C., & Strong, C. L. (2022). A self-regulated learning analytics prediction-and-intervention design: Detecting and supporting struggling biology students. *Journal of Educational Psychology*, *114*(8), 1801–1816. https://doi.org/10.1037/edu0000745

Dent, A. L., & Koenka, A. C. (2016). The relation between self-regulated learning and academic achievement across childhood and adolescence: A meta-analysis. *Educational Psychology Review*, *28*(3), 425–474. https://doi.org/10.1007/s10648-015-9320-8

Dixon, J. A., Holden, J. G., Mirman, D., & Stephen, D. G. (2012). Multifractal dynamics in the emergence of cognitive structure. *Topics in Cognitive Science*, *4*(1), 51–62.

Efklides, A. (2011). Interactions of metacognition with motivation and affect in self-regulated learning: The MASRL model. *Educational Psychologist*, *46*(1), 6–25.

Epskamp, S., Borsboom, D., & Fried, E. I. (2018). Estimating psychological networks and their accuracy: A tutorial paper. *Behavior Research Methods*, *50*, 195–212.

Fazeli, S., Drachsler, H., Bitter-Rijpkema, M., Brouns, F., Van der Vegt, W., & Sloep, P. B. (2017). User-centric evaluation of recommender systems in social learning platforms: accuracy is just the tip of the iceberg. *IEEE Transactions on Learning Technologies*, *11*(3), 294–306.

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences*, *111*(23), 8410–8415. https://doi.org/10.1073/pnas.1319030111

Fried, E. I., & Robinaugh, D. J. (2020). Systems all the way down: Embracing complexity in mental health research. *BMC Medicine*, *18*(1), 1–4.

Greene, J. A. (2018). *Self-regulation in education*. Routledge.

Greene, J. A. (2022). What can Educational Psychology learn from, and contribute to, theory development scholarship? *Educational Psychology Review*, *34*, 1–25.

Greene, J. A., & Azevedo, R. (2009). A macro-level analysis of SRL processes and their relations to the acquisition of a sophisticated mental model of a complex system. *Contemporary Educational Psychology*, *34*(1), 18–29.

Greene, J. A., Plumley, R. D., Urban, C. J., Bernacki, M. L., Gates, K. M., Hogan, K. A., Demetriou, C., & Panter, T. A. (2019). Modeling temporal self-regulatory processing in a higher education biology course. *Learning and Instruction*, *72*, Article 101201. https://doi.org/10.1016/j.learninstruc.2019.04.002

Handcock, M. S., Hunter, D. R., Butts, C. T., Goodreau, S. M., Krivitsky, P. N., Bender-deMoll, S., & Morris, M. (2019). Package 'statnet'.

Hilpert, J. C., & Marchand, G. C. (2018). Complex systems research in educational psychology: Aligning theory and method. *Educational Psychologist*, *53*(3), 185–202.

Hoppe, H. U. (2017). Computational methods for the analysis of learning and knowledge building communities. In C. Lang, G. Siemens, A. Wise, & D. Gasevic (Eds.), *Handbook of learning analytics* (pp. 23–33). Society for Learning Analytics and Research (SOLAR).

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55.

Jacobson, M. J., Levin, J. A., & Kapur, M. (2019). Education as a complex system: Conceptual and methodological implications. *Educational Researcher*, *48*(2), 112–119.

Klerkx, J., Verbert, K., Duval, E., Lang, C., Siemens, G., Wise, A., & Gasevic, D. (2017). Learning analytics dashboards. In C. Lang, G. Siemens, A. Wise, & D. Gasevic (Eds.), *Handbook of learning analytics* (pp. 143–150). Society for Learning Analytics and Research (SOLAR).

Knight, S., & Shum, S. B. (2017). Theory and learning analytics. In C. Lang, G. Siemens, A. Wise, & D. Gasevic (Eds.), *Handbook of learning analytics* (pp. 17–22). Society for Learning Analytics and Research (SOLAR).

Koopmans, M. (2020). Education is a complex dynamical system: Challenges for research. *The Journal of Experimental Education*, *88*(3), 358–374.

Lockyer, L., Heathcote, E., & Dawson, S. (2013). Informing pedagogical action: Aligning learning analytics with learning design. *American Behavioral Scientist*, *57*(10), 1439–1459.

Lombardi, D., Shipley, T. F., & Astronomy Team, Biology Team, Chemistry Team, Engineering Team, Geography Team, Geoscience Team, and Physics Team. (2021). The curious construct of active learning. *Psychological Science in the Public Interest*, *22*(1), 8–43. https://doi.org/10.1177/1529100620973974

Lu, O. H., Huang, J. C., Huang, A. Y., & Yang, S. J. (2017). Applying learning analytics for improving students engagement and learning outcomes in an MOOCs enabled collaborative programming course. *Interactive Learning Environments*, *25*(2), 220–234.

Martin, F., & Borup, J. (2022). Online learner engagement: Conceptual definitions, research themes, and supportive practices. *Educational Psychologist*, *57*(3), 162–177.

Midgley, C., Maehr, M. L., Hruda, L. Z., Anderman, E., Anderman, L., Freeman, K. E., & Urdan, T. (2000). *Manual for the patterns of adaptive learning scales*. University of Michigan.

Miller, C. J., & Bernacki, M. L. (2019). Training preparatory mathematics students to be high ability self-regulators: Comparative and case-study analyses of impact on learning behavior and achievement. *High Ability Studies*, *30*(1–2), 167–197.

Pekrun, R., & Marsh, H. W. (2022). Research on situated motivation and emotion: Progress and open problems. *Learning and Instruction*, *81*, 101664.

Perry, N. E., & Winne, P. H. (2006). Learning from learning kits: gStudy traces of students' self-regulated engagements with computerized content. *Educational Psychology Review*, *18*(3), 211–228.

Pinsky, M., & Karlin, S. (2010). *An introduction to stochastic modeling*. Academic Press.

Revelle, W. (2019). *psych: Procedures for psychological, psychometric, and personality research* [R package version 1.9.12]. Northwestern University. http://CRAN.R-project.org/package=psych

Saint, J., Gašević, D., Matcha, W., Uzir, N. A. A., & Pardo, A. (2020, March). Combining analytic methods to unlock sequential and temporal patterns of self-regulated learning. In *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge* (pp. 402–411). ACM Digital Library.

Schunk, D. H., & Greene, J. A. (2018). Historical, contemporary, and future perspectives on self-regulated learning and performance. In D. H. Schunk & J. A. Greene (Eds.), *Handbook of self-regulation of learning and performance* (2nd ed., pp. 1–15). Routledge. https://doi.org/10.4324/9781315697048

Siadaty, M., Gašević, D., & Hatala, M. (2016). Measuring the impact of technological scaffolding interventions on micro-level processes of self-regulated workplace learning. *Computers in Human Behavior*, *59*, 469–482.

Spedicato, G. A., Kang, T. S., Yalamanchi, S. B., Yadav, D., & Cordón, I. (2021). *The markovchain package: A package for easily handling discrete Markov chains in R*. ACM Digital Library.

Theobald, E. J., Hill, M. J., Tran, E., Agrawal, S., Arroyo, E. N., Behling, S., Chambwe, N., Cintrón, D. L., Cooper, J. D., Dunster, G., Grummer, J. A., Hennessey, K., Hsiao, J., Iranon, N., Jones, L., Jordt, H., Keller, M., Lacey, M. E., … Freeman, S. (2020). Active learning narrows achievement gaps for underrepresented students in undergraduate science, technology, engineering, and math. *Proceedings of the National Academy of Sciences*, *117*(12), 6476–6483.

Theobald, M. (2021). Self-regulated learning training programs enhance university students' academic performance, self-regulated learning strategies, and motivation: A meta-analysis. *Contemporary Educational Psychology*, *66*, 101976.

Turek, M. (2018). *Explainable artificial intelligence (XAI)*. Defense Advanced Research Projects Agency. https://wwwdarpa.mil/program/explainable-artificial-intelligence

Wallot, S., & Kelty-Stephen, D. G. (2018). Interaction-dominant causation in mind and brain, and its implication for questions of generalization and replication. *Minds and Machines*, *28*(2), 353–374.

Winne, P. H. (2017a). Cognition and metacognition within self-regulated learning. In D. H. Schunk & J. A. Greene (Eds.), *Handbook of self-regulation of learning and performance* (pp. 36–48). Routledge. https://doi.org/10.4324/9781315697048-3

Winne, P. H. (2017b). Learning analytics for self-regulated learning. In C. Lang, G. Siemens, A. Wise, & D. Gasevic (Eds.), *Handbook of learning analytics* (pp. 241–249). Society for Learning Analytics and Research. https://doi.org/10.18608/hla17.021

Winne, P. H. (2020). Construct and consequential validity for learning analytics based on trace data. *Computers in Human Behavior*, *112*, 106457. https://doi.org/10.1016/j.chb.2020.106457

Winne, P. H., & Hadwin, A. E. (1998). Studying as self-regulated learning. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Metacognition in educational therapy and practice* (pp. 291–318). Routledge.

Zimmerman, B. J. (2013). From cognitive modeling to self-regulation: A social cognitive career path. *Educational Psychologist*, *48*(3), 135–147.

How to cite this article: Hilpert, J. C., Greene, J. A., & Bernacki, M. (2023). Leveraging complexity frameworks to refine theories of engagement: Advancing self-regulated learning in the age of artificial intelligence. *British Journal of Educational Technology*, *54*, 1204–1221. https://doi.org/10.1111/bjet.13340