What College Students Say, and What They Do: Aligning Self-Regulated Learning Theory with Behavioral Logs

Joshua Quick
eLearning Research & Practice Lab,
Indiana University, Indiana, United States
jdquick@iu.edu

Benjamin Motz
Department of Psychological and Brain Sciences,
Indiana University, Indiana, United States
bmotz@indiana.edu

Jamie Israel
eLearning Design and Services,
Indiana University, Indiana, United States
jgisrael@iu.edu

Jason Kaetzel
eLearning Design and Services,
Indiana University, Indiana, United States
jkaetzel@iu.edu

Abstract

A central concern in learning analytics specifically and educational research more generally is the alignment of robust, coherent measures to well-developed conceptual and theoretical frameworks. Capturing and representing processes of learning remains an ongoing challenge in all areas of educational inquiry and presents substantive considerations on the nature of learning, knowledge, and assessment & measurement that have been continuously refined in various areas of education and pedagogical practice. Learning analytics as a still developing method of inquiry has yet to substantively navigate the alignment of measurement, capture, and representation of learning to theoretical frameworks despite being used to identify various practical concerns such as at risk students. This study seeks to address these concerns by comparing behavioral measurements from learning management systems to established measurements of components of learning as understood through self-regulated learning frameworks. Using several prominent and robustly supported self-reported survey measures designed to identify dimensions of self-regulated learning, as well as typical behavioral features extracted from a learning management system, we conducted descriptive and exploratory analyses on the relational structures of these data. With the exception of learners' self-reported time management strategies and level of motivation, the current results indicate that behavioral measures were not well correlated with survey measurements. Possibilities and recommendations for learning analytics as measurements for self-regulated learning are discussed.

CCS CONCEPTS

- Applied Computing ~ Learning Management Systems
- Information Systems ~ Web log analysis
- Computing Methodologies ~ Factor Analysis

LAK'20, March 23–27, 2020, Frankfurt, Germany

ACM ISBN 978-1-4503-7712-6/20/03...$15.00
https://doi.org/10.1145/3375462.3375516

KEYWORDS

LMS, self-regulated learning, self-reports, trace data

ACM Reference Format:

Joshua Quick, Benjamin Motz, Jamie Israel, and Jason Kaetzel. 2020. What College Students Say, and What They Do: Aligning Self-Regulated Learning Theory with Behavioral Logs. In *The 10th International Learning Analytics & Knowledge Conference (LAK20), March 2020, Frankfurt, Germany,* 10 pages. https://doi.org/10.1145/3375462.3375516

1 Introduction

Throughout the last decade, learning analytics has developed from a hodgepodge of overlapping interests and methods within the information, computer, and learning sciences to an dynamic, expanding dialogue between the learning and information sciences [49]. A central concern of this dialogue is in exploring the relationship between analytics as measurements, and our theoretical and practical commitments to advancing teaching and learning practices [23, 25, 26, 39]. With the ever-expanding data generated from learning and teaching digital systems, analyses must rely on theory for guidance and structure [57]. The question of connecting our developed measures with extant frameworks of learning are therefore of paramount importance for the advancement of learning analytics as a discipline.

One of the prominent theoretical frameworks of learning within educational research generally, and learning analytics specifically is *self-regulated learning*. Self-regulated learning (SRL) theories describe a family of frameworks that seek to elaborate how cognitive, motivational, and situational or contextual factors influence learning processes [34, 55, 56, 58]. Although differences between these specific frameworks exist, these various aspects represent an overarching category with which to understand and frame learning as an agentic, dynamic, and complex phenomena incorporating motivational, behavioral, cognitive, social, cultural, and situational elements.

SRL represents a broad umbrella of processes with which to understand learning and teaching across a variety of educational contexts. Within online or digitally-mediated education, SRL frameworks have been used to evaluate and understand the relations between learners' perceptions and practices of their regulation to outcomes in MOOCs [30], the relations of SRL processes to performance indicators in online education more generally [6], and the relations between learning behavior and goal attainment in MOOCs [24]. Such investigations exemplify a broad spectrum of practices that use various components and processes of existing SRL frameworks. These reviews and investigations have emphasized the importance of regulative strategies but have also called attention to the fact that regulation is a dynamic complex of skills, processes, and practices that are not necessarily continuously aligned. How contextual elements promote productive SRL practices within both online and blended forms of education has been a central area of investigation and a topic of substantive differences (see [6, 21]).

SRL also has substantive intermingling within learning analytics centered research. Indeed, the first special issue of the second volume of the *Journal of Learning Analytics* was specifically dedicated to the overlap between SRL and learning analytics (see [40]), as was a chapter in *The Handbook of Learning Analytics* (see [54]). To enumerate a few specific examples of this intersection, SRL and learning analytics has been used as a guide to construct measurements for predictive models [39], to detect SRL patterns in exploratory learning environments [45], and examine the relationship between performance and behavioral activity in MOOCS [24].

The alignment between theoretical frameworks of teaching and learning with analytical indices is central to the development of learning analytic methods. Since SRL represents one of the most common conduits connecting learning analytic methodologies with models of learning, this raises questions regarding whether behavioral metrics within learning systems and conventional measures of students' learning strategies measure similar, or even identical constructs. In SRL research traditions, the most common technique for examining the processes of students' learning activities is using self-reported survey instruments. Responses to standardized questionnaires provide data that exhibit learners' reflections on their intentions, strategies and processes in learning content. These approaches have typically been used in SRL analyses to monitor and understand learners' regulative processes across a variety of contexts [4]. A natural question and test for valid interpretations of behavioral analytics informed by SRL or other models of learning is its correspondence to existing measurements of these behaviors. Such approaches are necessary for learning analytics in order to coordinate ***triangulation*** [8] of research methods and designs for reliably and validly interpretable processes within learning analytics.

Several approaches within the learning analytics literature have sought to more firmly align extant measurement processes. Ga ević, Jovanovic, Pardo, and Dawson [15] identified links between deep learning behavioral strategies extracted from trace data of digital learning environments and self-reported deep learning and strategy approaches, but found no significant relations between shallower behavioral strategies and self-reported learning processes. In a more novel approach, Segedy, Kinnebrew, and Biswas [45] developed a process known as *coherence analysis* in order to detect strategies and processes in an open-ended learning environment and compared these constructed metrics with previous results from a similar study. Both of these examples, however, were limited to idiosyncratic educational contexts. Ga ević et al conducted this analysis on a moderate sample of 144 students in an engineering class while Segedy et al's investigation centered on a very specific open learning tool for science disciplinary knowledge.

Broader-scale investigations between the alignments of self-reported SRL strategies and behavioral processes in learning environments have also been conducted in MOOCs. Maldonado-Mahaud and colleagues [31] conducted a process mining approach to identify relationships between learner interaction sequences and existing SRL strategies in the literature. Similarly, Kizilcec et al. [24] found relationships between self-reported SRL strategies and behavioral interactions with resources within MOOCs. These approaches, however, also represent particular contexts and tools that do not necessarily align with more traditional educational learning systems or provide clear predictions out of sample. As was the case in [31], we should expect that a data mining effort to extract patterns in log data corresponding to SRL subtypes will, indeed, identify such patterns. That is, additional validation procedures are needed in the alignment of theoretical frameworks and learning analytic methods to test the generalizability of these inferences. While [24] addresses parts of these concerns through the alignment of SRL self-reports to student success, their approach also raises the need for further refinement and assessment of the metrics available within other digital learning environments, such as non-MOOC contexts, that will reveal more generalizable relationships between SRL strategies, student behaviors, learning outcomes.

Given that a core aim of learning analytic practices is to enhance teaching and learning and promote greater student success [13, 17, 23, 46], the relationship between these measures and established measurement processes in broader institutional contexts requires additional examination and verification. Consequently, this analysis seeks to identify the extent to which SRL derived metrics relate to established measurements for SRL processes, practices, and strategies. The goal of the current study, then, is to use self-reported survey measures developed through SRL frameworks and behavioral metrics internal to learning management systems to serve a triangulation function [14] by providing insights into the differences and similarities observed within these measurement approaches.

Our analytical work is carried out along two converging vectors. Our *top-down* approach starts with analysis of survey responses, and investigates their alignment with behavioral measures observed from the same students. Our *bottom-up* approach starts with analysis of the same behavioral measures, and investigates their alignment with students' survey responses. Along both axes, we use factor analysis to identify a reduced set of values that capture joint variance and then analyze the correlation between factor scores and the comparison measures. At the nexus of these two approaches, we compare both sets of factors (derived from top-down surveys and bottom-up behavioral logs) to evaluate alignment, illustrating limitations and affordances in the relationship between these measurement approaches. Learning analytics, as a developing form of analysis, must address these issues through such convergence in order to provide actionable insights that respond to and impact existing and emerging discussions in educational research and practice.

2 Method

415 undergraduate students at a large Midwestern university in the United States completed the survey in exchange for credit towards their Introductory Psychology experiment participation requirement. Participation in this survey included a consent and FERPA release form for students to give the researchers access to their educational records from Spring 2019 semesters. Of these, 35 students were excluded because they did not complete all questionnaires, and another 188 were excluded because they did not have course enrollments in the previous semester (many of the

Instrument	Subscale	Representative Example Item
ASSIST	Seeking meaning (SM)	I usually set out to understand for myself the meaning of what we have to learn.
	Relating ideas (RI)	I try to relate ideas I come across to those in other topics or other courses whenever possible.
	Use of evidence (UE)	Often I find myself questioning things I hear in lectures or read in books.
	Interest in ideas (II)	Regularly I find myself thinking about ideas from lectures when I'm doing other things.
	Monitoring effectiveness (ME)	I go over the work I've done carefully to check the reasoning and that it makes sense.
	Organized studying (OS)	I usually plan out my week's work in advance, either on paper or in my head.
	Time management (TM)	I work steadily through the term or semester, rather than leave it all until the last minute.
	Achieving (AC)	I put a lot of effort into studying because I'm determined to do well.
	Alertness to demands (AD)	When working on an assignment, I'm keeping in mind how best to impress the marker.
Grit	Consistency of interest (CI)	New ideas and projects sometimes distract me from previous ones.
	Perseverance of Effort (PE)	I finish whatever I begin.
MSLQ	Self-efficacy (SE)	I know that I will be able to learn the material for this class.
	Intrinsic value (IV)	Understanding this subject is important to me.
	Test anxiety (TA)	I have an uneasy, upset feeling when I take a test.
	Cognitive strategy use (CS)	When reading I try to connect the things I am reading about with what I already know.
	Self-regulation (SR)	Before I begin studying I think about the things I will need to do to learn.
LASSI	Anxiety (AX)	When I am taking a test, worrying about doing poorly interferes with my concentration.
	Attitude (AT)	I have a positive attitude about attending my classes.
	Concentration (CN)	If I get distracted during class, I am able to refocus my attention.
	Information Processing (IP)	I try to find relationships between what I am learning and what I already know.
	Motivation (MV)	Even if I am having difficulty in a course, I can motivate myself to complete the work.
	Selecting main ideas (MI)	When I listen to class lectures, I am able to pick out the important information.
	Self-testing (ST)	I review my notes before the next class.
	Test strategies (TS)	I review my answers during essay tests to make sure I have made and supported my main points.
	Time management (TM)	When I decide to study, I set aside a specific length of time and stick to it.
	Using academic resources (AR)	If I find that a course is too difficult for me, I will get help from a tutor.
SBI	Factor 1 ("Carelessness")	When tests are returned, I find that my grade has been lowered because of careless mistakes.
	Factor 2 ("Deprioritization")	I watch too much television and this interferes with my studies.
	Factor 3 ("Self-regulation")	Before attending class, I prepare by reading or studying the assignment.
	Factor 4 ("Organization")	I keep all the notes for each subject together, carefully arranging them in some logical order.
SESRL	[No subscales]	How well can you motivate yourself to do schoolwork?

Table 1: Survey instruments included in the current study and their associated subscales: Approaches and Study Skills Inventory for Students (ASSIST) [48], Grit [11], Motivated Strategies for Learning Questionnaire (MSLQ) [35], Learning and Study Strategies Inventory (LASSI) [53], Study Behaviors Inventory (SBI) [5], and Self-efficacy for Self-regulated Learning (SESRL) [50]. We provide descriptive labels for the SBI subscales in quotes, as no labels were provided in the original article.

Introductory Psychology students were in their first-semester). This removal was due to no behavioral data existing for these students that fell within the inclusion of this study's IRB and consent procedures. After these exclusions, there were 192 participants in the sample that moved to the next stage of analysis.

2.1 Materials

Participants completed a battery that contained 6 questionnaires, implemented in an online survey platform (Qualtrics; Provo, UT). These questionnaires were the Approaches and Study Skills Inventory for Students (ASSIST; [48]), Short Grit Scale (Grit-S; [11]), Motivated Strategies for Learning Questionnaire MSLQ; [35]), Learning and Study Strategies Inventory (LASSI; [53]), Study Behaviors Inventory (SBI; [5]), and Self-efficacy for Self-regulated Learning (SESRL; [50]). See Table 1 for subscales and example items. Questionnaires measured a range of constructs related to student motivation, self-regulation, learning strategies, and studying behaviors, with some modest overlap between the different inventories. While questionnaires have been criticized as a way to assess the application of SRL and other learning skills or processes (see [3,27]), questionnaires have nonetheless shown suitable reliability and contribute to valid judgments in predicting SRL behaviors and outcomes in higher educational systems which facilitate abstract reflection and conceptualization of learning [43].

The breadth of these self-report tools was motivated by the complexity of SRL processes and the data corpus (see section 2.3). As a cyclical process, SRL is inherently dynamic and can easily lead to repetition and development of these phases over time within a particular social and technical context [34, 55, 56, 58]. This complexity led to our 'wide net' approach to understand learners perceptions of their SRL processes across the cycles and they believe they engaged in within a single semester and the common technical system used within their academic contexts.

2.2 Procedure

The study was posted in an online sign-up system used for fulfillment of research participation requirements in our institution's Department of Psychological and Brain Sciences during Summer and Fall 2019. The study was completed entirely online. When a student first accessed the study, they saw an information sheet about the study, and then they agreed to release their student information for the purposes of the study, by electronic signature.

Participants were then shown instructions for taking the surveys. Instructions included that it would take approximately 30 minutes to respond to all the questions, and suggested for participants to take their time, read the questions carefully, and answer the questions thoughtfully. Additionally, participants were told that if they got tired, they could take a short break between survey screens.

Each questionnaire was shown on a single page, with items appearing in order as described in the questionnaire's cited article. The order of questionnaires, however, was randomized for different participants. Participants could not leave items blank; progress to the next screen required that all items were filled-in. Each questionnaire was displayed with the response scale shown at the top of the screen. Many of the MSLQ's items refer to a specific class, so we preceded the MSLQ with the instructions, "For the questions below, please try to consider a single, typical class that you are enrolled in." The polarity of the response scales was spatially-consistent across all questionnaires; agreement with or endorsement of an item was always on the left-side of the response options. Upon completing the last survey, participants were shown a short explanation of the current study, and they clicked once more to have the credit applied to their account in the sign-up system.

2.3 Student Data

For each participant, we accessed their enrollment records from the Spring 2019 academic term, the most recently-completed standard academic term prior to participants filling-out the surveys. Participants who did not have enrollments in Spring 2019 (e.g., those who were first-semester freshmen when filling-out the surveys) were excluded from the study (see *Participants*, above).

For participants who met the criteria for inclusion, we extracted their cumulative scores and constructed features from our institution's learning management system (LMS) data store (Canvas; Instructure, Salt Lake City, UT). Thirteen features were extracted that summarized behavioral logs measured throughout the entire Spring 2019 academic term. Many of these features centered on participants' activity pertaining to course assignments, as these reflect the primary and most generalizable mechanism within the LMS for students to regulate their interactions with course assessments. Specifically, the assignment tools indicate the primary prompts and information on assignment submission (e.g., due date, assignment instructions, submission method, criteria for grading, etc.). We also examined participants' use of the calendar features within Canvas in order to identify whether participants kept track of upcoming deadlines or calendar events through that system. Tables 2 describes and summarizes the features extracted. Finally, courses were categorized based on the subject category used within institution. Enrollments that had no submission data were excluded from final analysis. This left 748 enrollment records and 188 participants in the sample for the analysis stage.

2.4 Data Analysis

Queries for constructing behavioral features of students' LMS activity, and scripts for carrying out all analyses of data, are available at https://osf.io/8yhdp/.

Approximately half of the LASSI items, and 4 of the MSLQ items were reverse scored, so responses were reordered for these items prior to subsequent analysis. Items in the SBI that were not associated with the four factors (labeled as "No loadings" in [5] were excluded from analysis (they were still included in the questionnaire, however, to preserve authenticity to the original instrument). All SESRL item responses were treated as measuring one construct, as there are no reported subscales for the SESRL questionnaire [50].

All questionnaires used ordinal response scales (e.g., *Strongly agree* to *Strongly disagree*; *Very much like me* to *Not at all like me*). Rather than treating ordinal responses as metric values, we instead used Bayesian methods to estimate the mean value underlying each participant's ordered responses, separately for each subscale (see Table 1), using the hierarchical ordered-probit model described in Liddell & Kruschke [29]. We used Markov chain Monte Carlo (MCMC) to sample the posterior distributions (75,000 steps thinned to 15,000 saved samples across 4 chains, following 500 adaptation steps and 1,000 burn-in steps), using Gibbs sampling (JAGS; [36]), and the runjags package ([10]) for R. For each subscale, we measured a participant's score as the mode of the posterior distribution of estimated means for continuous normal distributions mapped onto participants' ordered ratings within that subscale. This score represents the most credible summary of a participant's underlying response tendencies to ordinal survey items within a subscale. All subsequent analyses of survey results

Feature	Description
timeOnAssignments	The average amount of time (secs.) spent on assignment pages within Canvas
timeBetweenFirstAccessandDeadline	The average time (hours) between a participants first access of an assignment and its submission.
sessionDuration	The average length of time per web session within Canvas where an assignment is accessed.
numberOfRequests	The average number of HTTP requests (e.g., a proxy for page views) in sessions in which an assignment is accessed within Canvas.
timeBetweenFirstAccessandSubmission	The average time (minutes) between first access of an assignment and its submission.
submissions	The number of assignment submissions created.
numberOfAssignmentAccessPreDeadline	The total count of views of assignments occurring prior to the deadline of the assignment.
numberOfAccessPostDeadline	The total count of views of assignments occurring after the deadline of an assignment.
numberOfSessions	The total count of web sessions in which an assignment was accessed.
numberOfCalendarAccess	The total number of calendar page views
largestPeriodOfInactivity	The longest period (in hours) between sessions within Canvas.
submissionTimeFromDeadline	The average time (minutes) between the submission and deadline for assignments.

Table 2: Features extracted from behavioral logs used in the current study. Each feature was z-scored within-course.

are performed on these estimated subscore values for each participant.

All features constructed from Canvas data were rescaled through z-scoring to adequately reflect a participants' interactions in a given enrollment in comparison to the rest of the students in their course. An analytic set with random identifiers was constructed by combining the z-scored Canvas features with the estimated mean value of survey responses. Since this analysis was focused on the relation between perceived regulation of participants' interactions with their actual behaviors in digital learning tools, cases where enrollments were missing two or more features were removed. This left 653 total enrollment observations. While deletion of missing data is commonly decried, this procedure was conducted in order to conservatively preserve the relationship of actual use of Canvas tools used within the course. The participants were further collapsed into an aggregated data set, averaging their z-scored behavioral interactions with the learning management system to the individual student level (rather than the enrollment level), and joined with their self-reported SRL strategies. This resulted in 181 total participant observations (see *Participants*, above).

Explorations of the covariance structure of the questionnaire subscale scores and the features of LMS behavioral logs were conducted through a factor analysis using the psych package [38] in R. Factors were extracted using the *oblimin* rotation method and the maximum likelihood factoring method. Factor retention was determined through parallel analysis [22], which evidence suggests may be among the most accurate approaches to determining the appropriate number of factors for an optimal solution [52, 59]. For factor analysis of survey subscale scores, a five factor solution was determined to be the most appropriate fit. For factor analysis of features measured from behavioral logs, a 3 factor solution was selected.

Extracted factors were then compared, in *top-down* and *bottom-up* approaches, to their respective comparison measurements through a Pearson correlation. Pearson correlations between participants' factor scores was conducted to identify a concise alignment between the measurement processes.

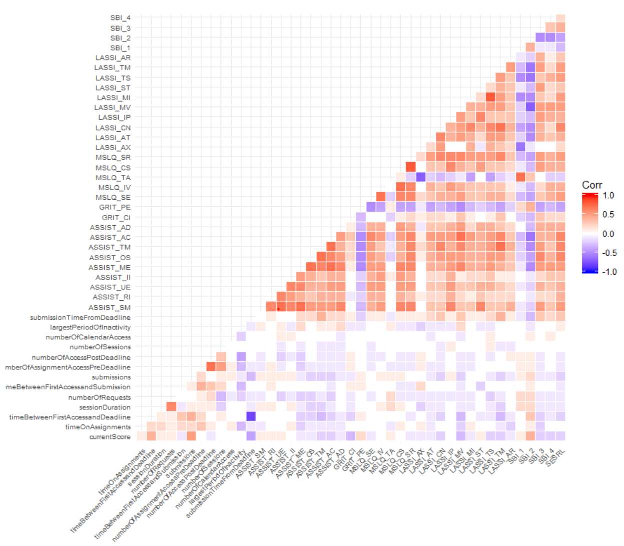

Figure 1: Correlation between and among behavioral features of LMS activity and survey subscale scores

3 Results

The median amount of time it took participants to complete the full set of surveys was 29.5 minutes. On average, participants were enrolled in 3 courses during the Spring 2019 academic term that had active LMS sites. These enrollments were well-distributed across our institution's academic units, with 13% of enrollments in humanities, 36% natural or applied science, 13% social sciences, and 8% within business or management courses. The remaining 30% were composed of a variety of subjects ranging from art, education, technology or engineering, and music.

Descriptive correlations of the relations between participants z-scored Canvas behavior and questionnaire subscale scores indicated little to no relations generally. Figure 1 describes the overarching correlational structure observed within these data. Correlations between related subscale scores appeared to be strong, as would be expected. For instance, test anxiety subscales (LASSI_TA, MSLQ_TA), and time management subscales (LASSI_TM, ASSIST_TM) covaried as anticipated. Considering that participants' responses to these items appeared on different questionnaires, on different pages, in different orders, these predicted relationships within the survey data reveal that participants were not answering randomly or haphazardly.

3.1. Top-down approach

Bartlett's test for sphericity indicated the correlation matrix of survey scores was sufficiently distinguished from the identity matrix ($\chi^2(903) = 4724.928$, $p < .001$) and thus factor analysis was an appropriate process to understand the underlying structure of survey responses. With an RMSR of .04, a RMSEA of .09, and a Tucker Lewis Index of Factor Reliability of .84, the overall fit of the model was relatively good. Figures 2 describes the relations between the questionnaire subscale scores and the 5 factor solution. High loadings (>0.7) on the ASSIST use of evidence (UE), relating ideas (RI), seeking meaning (SM), and interest in ideas (II) were suggestive of participants' perceptions on their own *Depth of Processing* information within their courses. Similarly, high loadings

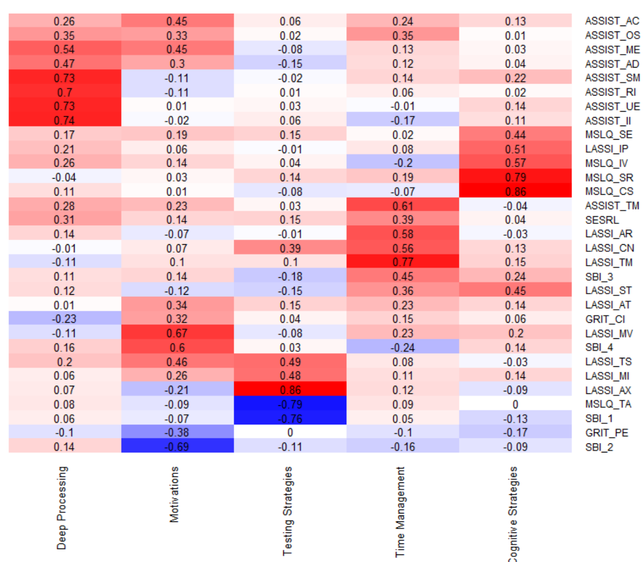

Figure 2: Factor loadings for Questionnaire Subscale Scores and 5 Factor Solution

(0.45-0.67) on the ASSIST achieving (AC), LASSI motivation (MV) and moderate loadings (0.32) on the GRIT consistency (CI) scale were glossed to be participants' perceived *Motivation*. The third factor, which we labeled *Testing Strategies*, indicated moderate to high loadings (0.49-0.86) on the LASSI test strategies (TS), selecting main ideas (MI), and anxiety (AX) subscales as well as discriminant relations between participants perceived anxiety as measured by the MSLQ anxiety (AX) subscale. The fourth factor, labeled *Time Management*, showed high loadings (0.45-0.77) on the LASSI and ASSIST time management (TM) subscales, and the LASSI concentration (CN) and using academic resources (AR) subscales. Finally, the 5th factor, labeled as *Cognitive Strategies*, indicated high loadings (0.44-0.86) on MSLQ self-efficacy (SE), self-regulation (SR), and cognitive strategy (CS), and LASSI information processing (IP).

Correlations between participants' z-scored Canvas behaviors and their factor scores are shown in Figure 3. In general, as was the case in Figure 1, little to no relation was apparent between participants' behaviors within Canvas and their factor scores. Some notable exceptions, however, included the submission time from deadline, submission counts, and the time between a student's first access of an assessment and the deadline of that work. Participants with an, on average, larger duration of time between their submission and the deadline of the assignment showed moderate

Figure 3: Correlation between factors derived from survey responses and behavioral features of LMS activity.

relations to self-reported time management (r=.39) and motivation factors (r=.37). Similarly, participants with a higher number of submissions on average and larger durations between the first access and deadline, were found to be somewhat negatively correlated with the time management (r=-.31 and r=.-.27, respectively) and motivation factors (r=-.24 and r=-.29, respectively).

This finding suggests the potential that students who scored higher on SRL time management and motivational scales tended to access assignments further from the deadline and had fewer submissions. Conversely, students with lower time management and motivation factors tended to access assignment instructions early, but (re)submit assignments more often and closer to the deadline. In other words, it is suggestive that there were relations between their time management and motivational factors exhibited behaviorally insofar as students used these tools and resources within Canvas. However, it is also notable that other common behavioral metrics, such as the time spent looking at assignments, the number of web sessions, and number of views of assignment instructions after the deadline (which might've been assumed to be representative of self-reflection, a central dimension in many SRL frameworks), were not associated with SRL subscales in any substantive way.

Of course, these findings must also be interpreted within the context of the collapse of participants' 3 enrollments into a single, averaged observation. This process may have reduced the impacts of specific course designs and processes that could mediate student behavior within Canvas. However, considering that SRL survey instruments commonly measure learning strategies at the student-level (and not at the class-level, with the exception of the MSLQ), aligning these two requires some abstraction of the behavioral indices to aggregate values for individual students.

3.2. Bottom-up approach.

Using the Canvas features, a separate factor analysis was constructed in order to detect the underlying relationships within our behavioral data. A Bartlett test for sphericity revealed that a factor analytic approach was appropriate ($\chi^2(55)=449$, p < .001). Parallel analysis indicated a 3 factor solution was appropriate, which had an RMSR of .04, a RMSEA of .06, and a Tucker Lewis Index of Factor Reliability of .92. Figure 4 summarizes the factor solution for the behavioral features. High loadings on access of assignments across various time windows and number of sessions indicated a general relationship between visits and were therefore labeled *Visit Frequency*. The number of requests and average session duration were also highly correlated with a factor and appeared to describe the relations between participants' time or depth within a visit to their course materials, which we labelled as *Visit Depth*. Finally, moderate loadings of the number of submissions, and a negative relation with the length of time between a submission and deadline suggested a pattern of sparse effort that was highly concentrated effort near to the assignment deadlines, which we labelled as *Concentration of Effort*.

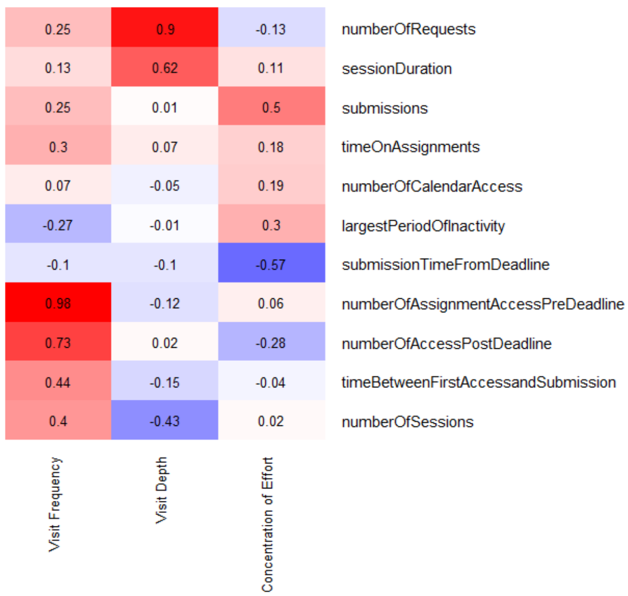

Figure 4: Factor loadings for behavioral features of LMS activity and 3 Factor Solution.

The relations between these extracted factors and the subscale scores are indicated in Figure 5. As might be expected given the scarcity of substantive relationships between the behavioral and self-reported variables (Figure 1), there was little connection between these variables and the factors generated from behavioral data. More than 2/3rds of the SRL subscales had correlations with factors of student's activity data beneath *r*=0.20. But as an exception, again, survey measures of time management (e.g., LASSI TM), and [lack of] motivation (e.g., SBI_1, SBI_2) seemed to vary systematically with how students behaved in the LMS.

3.3. Convergent comparison.

As a final comparison, we draw correlations between factor scores extracted from survey responses, and factor scores extracted from behavioral features of LMS activity. After distilling prominent sources of variance within these different measures to a small number of factor scores, the comparison between these factor scores enables examination of alignment from a more abstract and theoretically-informative frame. Figure 6 illustrates this comparison, at the convergence of our top down and bottom up triangulation function. As might be expected from analytical results thus far, there were only a couple modest relationships between participants' factor scores from the two approaches, one relating *Time Management* to *Concentration of Effort* (r=-0.29, p<0.001), and another relating *Motivations* to *Concentration of Effort* (r=-0.23, p<0.01). The remaining dimensions of self-regulated learning as measured by survey scores (*Cognitive Strategies, Testing Strategies, Deep Processing*) were not significantly related to the measured behavioral features of LMS activity. Similarly, the other factors derived from these behavioral metrics (*Visit Frequency, Visit Depth*) were not significantly correlated with self-reported survey measures of SRL. Finally, to the extent that the 3 factor solution provides an appropriate rendering of the variance structure of our behavioral features of LMS activity, the factor *Concentration of Effort* (which seems to reflect whether students "mass" their effort close to a deadline), comingles two different (and otherwise separable) dimensions of SRL: *Motivation* and *Time Management*. These inverse relationships suggest that students who concentrate their activity near to deadlines have *lower* self-reported measures in survey responses related to motivation and time management.

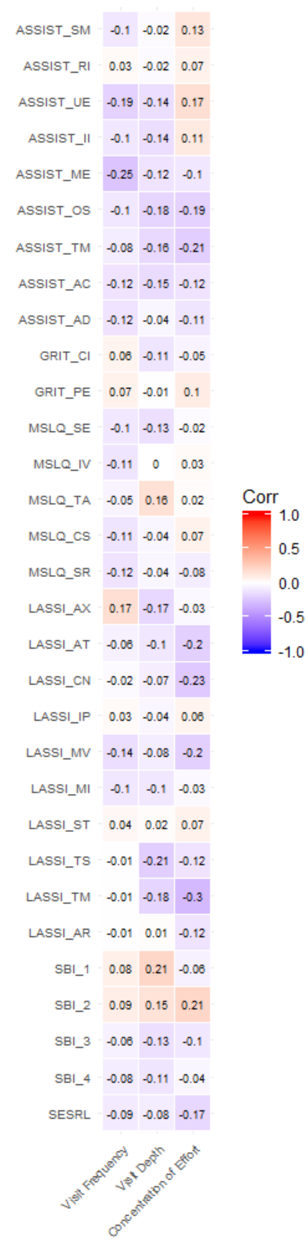

Figure 5: Correlation between factors scores derived from behavioral features of LMS activity and survey response scores.

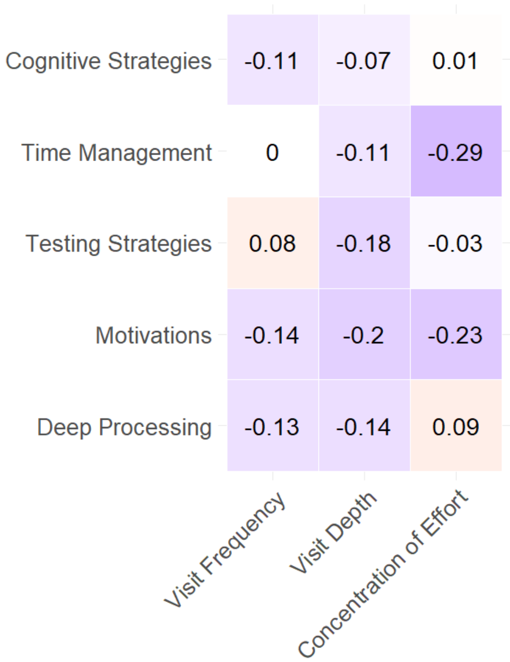

Figure 6: Correlation between factor scores. Rows (*Cognitive Strategies, Time Management*, etc.) are factors derived from survey responses. Columns (*Visit Frequency, Visit Depth, Concentration of Effort*) are factors derived from behavioral features of student activity in the LMS. Values are correlation coefficients, and cell shading matches the scale as shown in Figure 5.

4 Discussion

Our results suggest there is possibility of alignment in some respects with aspects of constructed behavioral measures and SRL self-reported methods, but that standard behavioral measures fall short of capturing the range of SRL dimensions commonly identified through survey research. The fact that there was some observed relation between the more concretely grounded behavioral components afforded by common LMS tools (e.g., clearly defined workflows for assignment schedules and due dates) and students' perceived time management and motivations is telling of what measurements from some aspects of digital systems interactions reveal. Recent work [51] corroborates these results, suggesting that conventional behavioral derivatives of LMS trace data demonstrate modest construct validity as measuring time management strategies and motivations.

In contrast, the behavioral derivatives are perhaps not as revelatory on learners' information processing and strategic practices as, perhaps, their self-assessments in relation to said assessments might be. Similarly, the analysis of discursive features and functions in online mediated tools (e.g., discussion forums) presents an intriguing path to gain insight into broader learning and regulatory practices [4] and have received substantive methodological focus see [33, 41]. Future work, then, might examine alignments between discourse metrics and SRL frameworks in addition to other theoretical perspectives.

Another consideration for the utility and validity of analytic behavioral measures is the overarching function of such measures in terms of what it is being compared with. Depending on the context of comparison, a feature such as "submission counts" could be indicative of more engagement if the determination is understood in terms of passing or failing a course (e.g., [32]). In contrast, when measured against perceived SRL processes, higher-than-average submission counts may reflect hasty corrections, extra attempts, "do-overs," or other behaviors representative of poor regulatory skills (see Figure 3). For predicting these two outcomes, passing or appropriately regulating study, the same variable, submission count, may have opposite qualitative interpretations.

These results are likely mediated, however, by a variety of functional design choices implemented within the array of courses included in this analysis. Course design structures influence the way in which these common measures can be more fully understood and used in analytic processes [16]. Examination of the relationship between these metrics therefore represents a viable path to further understand these measures in context. However, middle ground approaches also exists in the alignment of instructor perceptions of student engagement and online behaviors (see [32]). In either case, these avenues represent an array of methodological choices with which to align self-regulated learning processes. Furthermore, such alignment ought to develop across fundamentally different measurement processes, in order to scaffold more robust convergence within analytics as a theoretically informative and practically impactful tool.

Obstacles to this alignment are on full display in the current results, as some intuitive associations between students' behaviors and survey responses were simply not observed. For example, deep processing (a factor derived from students' survey responses) was not significantly associated with the depth of user interactions with course materials on the LMS (see Figure 6). Additionally, there was no correlation ($r = 0.0$) between the frequency of activity in the LMS (a factor derived from students' behaviors in Canvas), and students' time management skills (a factor derived from students' survey responses). This null finding merits some attention. Past research commonly observes that the frequency and duration of activity in an LMS reflects positive evidence of student engagement (e.g., [27, 32]), so it may be the case that these positive associations exists, but do not generalize across courses (see [9]), thus diluting the aggregate measures in the current study.

But it is also possible that disconnects observed in our results reflect discrepancies between self-reported measures, which have been used to construct theory on self-regulated learning for the past 20 years, and behavioral measures, which have more recently become the focus of learning analytics. Students' responses to survey items inquiring about studying habits are fundamentally indirect measures, as these rely on introspective reports of one's own behavior, rather than direct measurement of this behavior. On the one hand, as with any self-report measure, it is possible that aspects of the questionnaire's administration or the questions themselves could have systematically biased some respondents [37], perhaps in the direction of more socially-desirable responses [12]. On the other hand, students simply might not have accurate introspective access to their own studying effort [18, 20], or are unaware of how their individual effort compares with that of their classmates.

Rovers, Clarebout, Savelberg, de Bruin, and van Merriënboer's [42] recent review on the relation between self-report and behavioral measures of SRL provides more expansive insights into these issues. Their narrative review indicates that students may have more insight into their global SRL processes, which can be measured through self-report tools; in contrast, behavioral measures can provide good indicators of specific strategies and functions of learners within a particular context. Granularity of SRL processes is therefore a fundamental factor in considering the alignment of these methodological approaches. This distinction is a relevant one and suggests the need for further consideration on the functional purpose of these measurements in terms of the consequential and social functions of these approaches for developing productive and responsible interventions. Whether measuring SRL with self-report questionnaires or with LMS activity logs, validity is not a generic property of the measurement approach; rather, validity is a property of one's interpretation of the measurement. Beyond suggesting that researchers must consider the purpose of a measurement, this distinction also suggests that more novel measurement processes for interpreting the complexities of SRL may be necessary.

One such novel approach, and a likely future direction of our own approach, is process modeling. Process modeling has been used to detect sequential patterns of behaviors as particular events of SRL within learning analytics relatively recently [see 44, 47,51]. Such approaches recognize the temporality and cyclicality of SRL processes but also generate challenges in effective interpretation and construction of SRL events. One such constraint is the issue of scale across pedagogical design and social contexts. Given that instructional designs mediate the behavioral processes within those designs [16], it is likely that analyses across contexts, such as ours, would be hampered by variability in SRL behaviors, and processes that are muddled due to an array of social and contextual factors. Future work, then, should consider how one might extend the scale and functionality of process metrics, for examining their alignment with SRL constructs.

In the current analyses, we aimed to quantify this alignment by first reducing our measurements (using factor analysis) to a set of abstract and meaningful dimensions, and then examining correlation between factor scores along these dimensions. Due to the large quantity of pairwise comparisons in the current study, we have intentionally avoided calculating p-values for these comparisons (except in the final convergent comparison, Section 3.3), as these would have dramatically inflated the familywise error rate. Moreover, our overarching analytical goal was not to merely discover significant coefficients or to reject null hypotheses, but rather to examine the relative agreement between two different, but increasingly overlapping, measurement traditions. Indeed, among the current study's key findings is not only that a couple significant associations exist, but that some intuitive and expected associations were simply not observed.

One criticism of our current results might suggest that our set of measured behavioral features provided incomplete, if not outright *poor* coverage of the range of student activities quantifiable from LMS activity logs. This suggestion is undeniably true. However, the behavioral features we included in this study are representative of those commonly discussed in learning analytics research and conventionally viewed as meaningful and useful (e.g., [9, 27, 32]). Alternatively, data mining approaches would succeed in identifying patterns of students' behavioral activities that are *maximally* associated with scores on SRL survey subscales, but the generalizability, the meaningfulness, and the utility of those patterns would be questionable. Similarly, one might view SRL theories and survey instruments as being incomplete accountings of the dynamic range of students' agentic learning processes with digital coursework. Perhaps, if one had complete and unrestricted access to students' mental states, a data mining algorithm could similarly mine students' minds for hitherto-unidentified SRL patterns that correlate with students' activity in the LMS (perhaps motor readiness for one's finger to click on one's computer mouse), but that might be practically useless for the goals of improving and optimizing student learning. Our goal is not to suggest that we've analyzed the full range of possible alignments between behavior and introspection during SRL processes, but rather to suggest that the conventional measures that frequently appear in learning analytics discourse show underwhelming alignment in this sample.

The development of coherent, robust measures for learning analytics fundamentally involves the alignment between theoretical frameworks of learning, knowing, and assessment [26]. Mapping out these methodological frameworks in coherent ways should therefore work within the intersections of these alignments, specifically in areas where there is divergence on what learning, knowing, and the measurement thereof entails. Such discussions have been ongoing concerns within educational research more generally [1, 2, 7, 19]. The translation of this discussion into the context of learning analytic inquiries therefore represents an emergent challenge that should be addressed at the intersections of research and practice. Considerations on the functional purposes of the measurements and the processes and practices with which they are compared to, who benefits from the measurement, and why these measurements are necessary [25] should also be developed into frameworks of learning analytic research and practice designs. The continued development of learning analytics as an inquiry method, then, is contingent on the convergence and contradictions of the implications of these theoretical frameworks towards metrics generated in digital tools and systems.

REFERENCES

[1] John Anderson, James Greeno, Lynne Reder, and Herbert Simon (2000). Perspectives on learning, thinking, and activity. *Educational Researcher*, 29, 4, 11-13.

[2] John Anderson, Lynne Reder, and Herbert Simon. 1997. Situative versus cognitive perspectives: Form versus substance, *Educational Researcher,* 26, 1, 18-21.

[3] Cordula Artelt. 2000. How predictive are self-reported strategies for their actual use? *Zeitschrift fur padagoische psychologies*, 14, 2-3, 72-84.

[4] Roger Azevado. 2015. Defining and measuring engagement and learning in science: Conceptual, theoretical, methodological, and analytical issues. *Educational Psychologists,* 1, 84-94.

[5] Leonard Bliss and Richard Mueller. 1993. An instrument for the assessment of study behaviors of college students. *Literacy Research and Instruction,* 32, 4, 46-52.

[6] Jim Broadbent and Walter Poon. 2015. Self-regulated learning strategies & academic achievement in online higher education learning environments: A systemic review. *The Internet and Higher Education, 27*, 1-13.

[7] Robbie Case. 1996. Changing views of knowledge and their impact on educational research and practice. In *The Handbook of Education and Human Development*, Blackwell Publishers, Cambridge, MA. 75-99.

[8] Louis Cohen, Lawrence Manion, and Keith Morrison. 2007. *Research Methods in Education* (6th ed.). Routledge, New York, NY.

[9] Rianne Conjin, Chris Snijders, Aad Kleingeld, and Uwe. Matzat. 2016. Predicting student performance from LMS data: A comparison of 17 blended courses using Moodle LMS. *IEEE Transactions on Learning Technologies*, 10, 1, 17-29.

[10] Matthew Denwood. 2016. Runjags: An R package providing interface utilities, model templates, parallel computing methods and additional distributions for MCMC models in JAGS. *Journal of Statistical Software*, 71, 9, 1-25.

[11] Angela Duckworth and Patrick Quinn. 2009. Development and validation of the Short Grit Scale (Grit-S). *Journal of Personality Assessment*, 91, 2, 166-174.

[12] Allen Edwards. 1953. The relationship between the judged desirability of a trait and the probability that the trait will be endorsed. *Journal of Applied Psychology*, 37, 2, 90-93.

[13] Rebecca Ferguson. 2012. Learning analytics: drivers, developments, and challenges. *International Journal of Technology Enhanced Learning*, 4, 5/6, 304-317.

[14] Nigel Fielding. 2012. Triangulation and mixed methods designs: Data integration with new research technologies. *Journal of Mixed Method Research*, 6, 2, 124-136.

[15] Dragan Ga ević, Jelena Jovanovic, Aberlado Pardo, and Shane Dawson. 2017. Detecting learning strategies with analytics: Links with self-reported measures and academic performance. *Journal of Learning Analytics*, 4, 2, 113-128.

[16] Dragan Ga ević, Shane Dawson, Tim Rogers, and Danijela Ga ević. 2016. 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.

[17] Dragan Ga ević, Shane Dawson, and George Siemens. 2015. Let's not forget: Learning analytics are about learning. *TechTrends,* 59, 1, 64-71.

[18] Alison Gopnik. 1993. How we know our minds:The illusion of first-person knowledge of intentionality. *Behavioral and Brain sciences*, 16, 1, 1-14.

[19] James Greeno. 1997. On claims that answer the wrong questions. *Educational Researcher*, 26, 1, 5-17.

[20] David Hammer and Andrew Elby. 2002. On the form of a personal epistemology. Personal epistemology: *The psychology of beliefs about knowledge and knowing*, 169-190.

[21] Nina Hood, Allison Littlejohn, Colin Milligan. 2015. Context counts: How learners' contexts influence learning in a MOOC. *Comptuers & Educations,* 91, 83-91.

[22] John Horn. 1965. A rationale and test for the number of factors in factor analysis. *Psychometrika*, 30, 2, 179-185.

[23] Kristy Kitto, Simon Buckingham-Shum, Andrew Gibson. 2018. Embracing imperfection in learning analytics. 2018. In *Proceedings of the 8th International Conference of Learning Analytics and Knowledge*, 451-460.

[24] René Kizilcec, Mar Pérez-Sanagustín, and Jorge Maldonado. 2017. Self-regulated learning strategies predict learner behavior and goal attainment in Massive Open Online Courses. *Computers & Education*, 104, 18-33.

[25] Simon Knight and Simon Buckingham-Shum. 2017. Theory and Learning Analytics. In *The Handbook of Learning Analytics*, SOLAR, 17-22.

[26] Simon Knight, Simon Buckingham-Shum, and Karen Littleton. 2015. Epistemology, assessment, pedagogy: where learning meets analytics in the middle space. . *Journal of Learning Analytics,* 1, 23-47.

[27] Vitomir Kovanović, Dragan Ga ević, Shane Dawson, Srecko Joksimović, Ryan Baker, and Marek Hatala. 2015. Penetrating the black box of time-on-task estimation. In *Proceedings of the Fifth International Conference on Learning Analytics & Knowledge*, 184-193.

[28] Kathryn Ley and Dawn Young. 1998. Self-regualtion behaviors in underprepared (developmental) and regular admission college students. *Educational Psychology*, 23, 1, 42-64.

[29] Torrin Liddell and John Krusche. 2018. Analyzing ordinal data with metric models: What could possibly go wrong*? Journal of Experimental Social Psychology*, 79, 328-348.

[30] Allison Littlejohn, Nina Hood, Colin Miligan, and Paige Mustain. 2016. Learning in MOOCs: Motivations and self-regulated learning in MOOCs. *The Intenet and Higher Education, 29,* 40-48.

[31] Jorge Maldonado-Mahauad, Mar Pérez-Sanagustín, René Kizilcec, Nicolás Morales, and Jorge Munoz-Gama. 2018. Mining theory-based patterns from big data: Identifying self-regulated learning strategies in Massive Open Online Courses. *Computers in Human Behavior*, 80, 179-196.

[32] Benjamin Motz, Joshua Quick, Noah Schroeder, Jordon Zook, and Matthew Gunkel. 2019. The validity and utility of activity logs as a measure of student engagement. In *Proceedings of the 9th International Conference on Learning Analytics & Knowledge*, 300-309.

[33] Trena Paulus and Alyssa Wise. 2019. *Looking for Insight, Transformation, and Learning in Online Talk.* Routledge, New York, NY.

[34] Paul Pintrich. 2000. Multiple goals, multiple pathways: The role of goal orientation in learning and achievement. *Journal of educational psychology,* 92, 3, 544-555.

[35] Paul Pintrich and Elisabeth de Groot. 1990. Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology,* 82, 1, 33-40.

[36] Martyn Plummer. 2013. JAGS: Just another Gibbs Sampler version 3.4.0. *http://mcmc-jags.sourceforge.net*.

[37] Phillip Podsakoff, Scott MacKenzie, Jeong-Yeon Lee, and Nathan Podsakoff. 2003. Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, 88, 5, 879-903.

[38] William Revelle. 2011. An overview of the psych package. Department of Psychology Northwestern University Accessed on March, 3, 2012, 1-25.

[39] Tim Rogers, Dragan Ga ević, and Shane Dawson. 2016. Learning analytics and the imperative for theory driven research. In *The SAGE Handbook of E-learning Research*, Sage Publications, Los Angeles, CA, 232-250.

[40] Ido Roll and Phillip Winne. 2015. Understanding, evaluating, and supporting self-regulated learning using learning analytics. *Journal of Learning Analytics*, 2, 1, 7-12.

[41] Carolyn Rosé. 2017. Discourse analytics. In Charles Lang, Alyssa Wise, Geogre Siemens, and Dragan Ga ević (Ed.). *The Handbook of Learning Analytics,* SOLAR, 105-114.

[42] Sane Rovers, Geraldine Clarebout, Hans Savelberg, Angieqe de Bruin, and Jerroen van Merriënboer. 2019. Granularity Matters: comparing different ways of measuring self-regulated learning. *Metacognition and Learning,* 14, 1. 1-19.

[43] Anne Roth, Sabine Ogrin, and Bernhard Schmitz. 2016. Assessing self-regulated learning in higher education: a systemic literature review of self-report instruments. *Educational Assessment, Evaluation, and Accountability*, 28, 2016, 225-250.

[44] John Saint, Dragan Ga ević, and Aberlardo Pardo. 2018. Detecting learner strategies through process modeling. In *European Conference of Technology Enhanced Learning.* Springer, 385-398.

[45] James Segedy, John Kinnebrew, and Gautam Biswas. 2015. Using coherence analysis to characterize self-regulated learning behaviors in open-ended learning environments. *Journal of Learning Analytics,* 2, 1, 13-48.

[46] George Siemens. 2013. Learning analytics: The emergence of a discipline. *American Behavioral Scientist*, 57, 10, 1380-1400.

[47] Christoph Sonnenberg and Maria Bannert. 2015. Using Process Mining to examine the sustainability of instructional support: How stable are the effects of meta-cognitive prompting on self-regulatory behavior. *Computers in Human Behavior*, *96*, 259-272.

[48] Hilary Tait, Noel Entwistle, and Velda McCure. 1997. ASSIST: a reconceptualization of the approaches to studying inventory. In *Improving Student Learning: Improving students as learners*, Oxford, 262-271.

[49] Stephanie Teasley. 2019. Learning analytics where information science and the learning sciences meet. *Information and Learning Sciences*, 120 ½, 59-73.

[50] Ellen Usher and Frank Pajares. 2008. Self-efficacy for self-regulated learning: A validation study. *Educational and Psychological Measurement*, 68, 3, 443-463.

[51] Nora'ayu Ahmad Uzir, Dragan Gašević, Wannisa Matcha, Jelena Jovanović, Abelardo Pardo. 2019. Analytics of time management strategies in a flipped classroom. *Journal of Computer Assisted Learning*, 1-19.

[52] Wayne Velicer, Cheryl Eation, and Joseph Fava. 2000. Construct explication through factor or component analysis: A review and evaluation of alternative procedures for determining the number of factors or components. In *Problems and solutions in human assessment*, Springer, Boston, MA, 41-71.

[53] Claire Weinstein, Stephen Zimmermann, and David Palmer. 1988. Assessing learning strategies: The design and development of the LASSI. In *Learning and Study Strategies*, Academic Press, San Diego, CA, 25-40.

[54] Phillip Winne. 2017. Learning analytics for self-regulated learning. In Charles Lang, Alyssa Wise, George Siemens, and Dragan Ga ević (Ed.). *The Handbook of Learning Analytics*, 241-249. SOLAR.

[55] Phillip Winne and Allyson Hadwin. 2008. The weave of motivation and self-regulated learning. In *Motivation and Self-Regulated Learning: Theory, Research, and Applications 2,* Routledge, New York, NY, 297-314.

[56] Phillip Winne. 2001. Self-regulated learning viewed from models of information processing. In *Self-regulated learning and academic achievement: Theoretical Perspectives,* (2nd ed.)., Lawrence Eribaum Associates, Mahwah, NJ, 153-189.

[57] Alyssa Wise and David Shaffer. 2015. Why theory matters more than ever in the age of big data. *Journal of Learning Analytics*, 2, 2, 5-13.

[58] Barry Zimmerman and Dale Schunk (eds.). 2001. *Self regulated learning and academic achievement: Theoretical perspectives.* (2nd ed.). Lawrence Eribaum Associates, Mahwah, NJ

[59] William Zwick and Wayne Velicer. 1982. Factors influencing four rule for determining the number of components to retain. *Multivariate Behavioral Research,* 17, 2, 253-269.