Associations Between Self-Regulated Learning Strategies and Science Assignment Score in a Digital Literacy Platform

Effat Farhana, North Carolina State University, efarhan@ncsu.edu
Teomara Rutherford, University of Delaware, teomara@udel.edu
Collin F. Lynch, North Carolina State University, cflynch@ncsu.edu

Abstract: Self-regulated learning (SRL) strategies are connected to student achievement and learning outcomes. However, students may get few opportunities to learn, practice, and apply these strategies in traditional classrooms. We examine the associations between SRL strategies and performance in a digital literacy platform, Actively Learn, for K-12 education. This platform gives students opportunities to engage in SRL and collects trace data on their use of the system. We examined four SRL-related features from 14,575 students' activity data. We applied hierarchical linear models using assignment scores as the outcome variable. We found that all SRL strategies had positive associations with science scores, and that these associations were stronger for short-answer scores than for multiple choice question scores. Our results support the construct validity of log measures of SRL and offer insights into how the benefits of SRL may vest differently depending on assessment type.

Introduction

Self-regulated learners—those who plan, monitor, and evaluate progress toward goals—tend to be metacognitively, motivationally, and behaviorally active participants in the learning process (Azevedo & Cromley, 2004; Pintrich & De Groot, 1990; Zimmerman, 1990). Successfully engaging in SRL processes has been linked to greater learning and achievement (Pintrich & De Groot, 1990). Prior research has shown that through instruction, students can internalize and apply skills related to SRL, such as planning and monitoring (Brown, 1989), as well as annotating and using double entry journals (Herman et al., 2010). However, students in classroom settings often get insufficient opportunities to learn and practice these and other SRL strategies (Paris & Paris, 2001). Further, the typical teacher/student ratio and priority for topic completion make it difficult for teachers to monitor students' SRL processes (Herman et al., 2010).

Computer-based learning environments (CBLEs) can provide opportunities to learn and practice SRL strategies in classroom settings. The learning activities logged by CBLEs provide granular records of student activities as learners interact with the system. This rich array of data can be analyzed to derive students' cognitive and metacognitive states and learning behaviors. In this paper, we use student-system interaction data gathered while students engage with science content on the digital literacy platform, Actively Learn (AL), to present an exploratory analysis assessing the relation between students' SRL learning behaviors and science performance.

Theoretical background and literature review

SRL, as described by Zimmerman (1990), includes goal setting, organizing, self-monitoring, and self-evaluating. A learner uses self-monitoring strategies to assess their progress towards a specific goal (Ertmer & Newby, 1996). When engaging with texts, expert learners use monitoring strategies, such as reading and re-reading, to check their comprehension (Krathwohl & Anderson, 2001), and highlighting to identify important passages (Winne et al., 2017).

Existing research on SRL behaviors within educational technology has examined student learning strategies within hypermedia (Azevedo & Cromley, 2004), in an open-ended environment (Munshi et al., 2018), and in game-based environments (Rutherford, 2017; Taub et al., 2018). We expand this work to focus on a suite of SRL behaviors within a popular CBLE and link these behaviors to student performance.

The current study

The present study investigates the relationship between students' observed SRL activities and their performance on reading comprehension assignments in a scientific domain. We focus on reading comprehension as an integral part of science learning (Norris & Phillips, 2003). Despite the importance of reading comprehension, 70% of eighth-grade US students were below proficiency in assessments of academic text reading (National Center for Education Statistics, 2005). The AL platform allows instructors to engage their students in science reading by assigning reading assignments with embedded questions. These assignments range in length from one to a few pages, depending on student levels, and the questions include multiple choice, fill in the blank, and free text. Instructors may use predefined assignments and questions or add their own. The questions appear embedded in the reading and students answer the questions at will as they work through the readings. Some questions are automatically graded by the system and provide immediate feedback, whereas others, notably the free-text questions, are not. Students using the system may highlight text in the system and take notes within the platform. The AL platform can be used in any domain. For this study, we focus on physical science assignments at the middle-school level. The predefined science assignments in AL are designed following the Next Generation Science Standards (NGSS) guidelines and question standards. The platform's developers claim the platform promotes deep learning by close reading: annotating, highlighting, and engaging with text (Reading_AL, 2019). In this study, we assess students' SRL reading strategies and the connection of these strategies to science performance. To achieve this goal, we analyze four SRL features related to reading strategies: re-reading, highlighting, note taking, and vocabulary lookups from logs of middle school physical science assignment data.

We answer the following research questions

RQ1. What is the association between SRL activities and students' overall assignment score?

RQ2. Does this association differ depending on question format?

Methodology

Dataset preparation

AL developers provided us a dataset of middle-school physical science assignments completed on the platform during the 2017-2018 school year. On average, assignments contained 2.52 multiple choice questions and 4.19 short answer questions. Examples of questions of an AL assignment are presented below.

1. (MCQ) Which of the following is NOT a characteristic of protons?
    - Protons have a charge of +1.
    - Protons are located in the nucleus of an atom.
    - A proton has about the same size mass as an electron.
    - A proton is a tiny, dense region at the center of the atom.
2. (SA) In what ways are quarks and gluons related to protons?

The dataset covers 69 unique readings, used in 1,033 classes by 17,886 user accounts. After examining the data we opted to exclude classes containing more than 60 or fewer than 10 participants from our analysis. This left 83.45% of the students; we also excluded any participant that was enrolled in multiple classes (an additional 1.95% students), as we believed that these accounts may have been used by teachers or for testing purposes. This resulted in a final sample size of 14,575 unique users in 701 classes. We extracted SRL features from the logged data. The system logs each time the student views the assignment page. We classified the students' first access to a page as the first read and treated all subsequent accesses as rereading. Table 1 presents descriptive statistics for all our variables.

Table 1: Descriptive statistics of variables. MCQ = multiple choice question, SA = short answer

Variable	Range	Mean	SD
Students Per Class	10 - 59	26.57	9.29
Total Question Per Assignment	2-70	7.092	7.712
MCQ per Assignment	0 - 8	2.52	2.53
SA Per Assignment	0 - 68	4.19	5.96
Score Overall	0 - 143	10.001	10.370
MCQ Score	0 - 32	5.831	8.474
SA Score	0 - 135	4.170	6.583
Re-reading	0 - 23	0.354	0.806
Note Taking	0 - 20	0.218	0.978
Highlighting	0 - 31	0.044	0.596
Vocabulary Lookup	0 - 20	0.155	0.821

Modelling

Given the complex nature of our data, we used hierarchical linear models (HLMs) to model the relationship between observed behaviors and performance, with assignment at level one, nested within students (level two), nested within classes (level three). We built three models for three different response variables: overall assignment score, MCQ score, and SA score, which combined fill-in-the-blank and free-text questions. The fixed-effect variables were the SRL features and number of questions in assignment; these variables were at Level 1. Assignment, student, and class were all modeled as random intercepts.

Results and discussion

Table 2 displays the results of three models. We report standardized effect size using the formula ß = (B*SDx)/SDy (see e.g., Rutherford et al., 2017). All SRL-related variables had positive and statistically significant association with overall science score. Re-reading had the highest predictive power (B = 1.667, ß = 0.128, $p < 0.001$), followed by note taking (B = 0.582, ß = 0.055, $p < 0.001$), highlighting (B = 0.492, ß = 0.028, $p < 0.001$), and vocabulary lookups (B = 0.275, ß = 0.021, $p < 0.001$).

In order to address RQ2 and how predictive associations varied based on question format, we fit two models, one with the MCQ score as the dependent variable, and one with the SA score. All of the SRL strategies continued to be positive predictors of SA score with statistically significant associations; all but the vocabulary lookups continued to be statistically significant positive predictors of MCQ score. Standardized betas were also higher in SRL/SA associations. One possible explanation for this could be that when answering MCQ, students could leverage some partial knowledge to rule out one or more distractors. Thus, deeper processing, as is associated with SRL (Bliss, 1980), may be less necessary for answering MCQs.

Table 2: Results from HLM measuring association between SRL and science score

L1 (Assignment) Level	ß	B	SE	*p*
RQ1: Overall Score				
Intercept		6.533	0.402	< 0.001
No. Times Reread	0.128	1.667	0.069	< 0.001
No. Notes	0.055	0.582	0.062	< 0.001
No. Highlights	0.028	0.492	0.072	<0.001
No. Vocabulary Lookups	0.021	0.275	0.055	< 0.001
No. Questions	0.303	0.410	0.035	< 0.001
RQ2:MCQ Score				
Intercept		5.510	0.369	< 0.001
No. Times Reread	0.032	0.345	0.041	< 0.001
No. Notes	0.024	0.206	0.038	< 0.001
No. Highlights	0.016	0.228	0.045	< 0.001
No. Vocabulary Lookups	-0.003	-0.036	0.031	0.259
No. Questions	0.171	0.188	0.030	< 0.001
RQ2:SA Score				
Intercept		1.699	0.232	< 0.001
No. Times Reread	0.150	1.223	0.043	< 0.001
No. Notes	0.040	0.271	0.038	< 0.001
No. Highlights	0.019	0.210	0.043	< 0.001
No. Vocabulary Lookups	0.036	0.289	0.035	< 0.001
No. Questions	0.210	0.180	0.021	< 0.001

Our analyses demonstrate positive associations between SRL activities and assignment scores, supporting associations between SRL and achievement found in prior studies, such as the relationships between annotations and science reading in Herman et al. (2010). Although these results do not establish a causal relationship between science achievement and SRL variables, they add to the growing body of literature on how SRL behaviors can be extracted from CBLE trace data and how such behaviors relate to performance. Extensions of this work can build toward leveraging these traces toward actionable SRL enhancements.

Conclusion and future work

We investigated the association between SRL activities and science assignment score within middle school physical science articles in a CBLE. Three-level hierarchical linear models were fit to analyze how four SRL activities associated with overall scores and with scores on two question formats. SRL variables were positively correlated with science score; relationships were stronger for short-answer questions. Future work could unpack these associations by including article content and question type (e.g., Bloom level; NGSS mapping).

References

Azevedo, R., & Cromley, J. G. (2004). Does training on self-regulated learning facilitate students' learning with hypermedia? *Journal of Educational Psychology*, *96*(3), 523 -535.

Bliss, L. B. (1980). A test of Lord's assumption regarding examinee guessing behavior on multiple-choice tests using elementary school students. *Journal of Educational Measurement, 17*(2), 147-152.

Brown, A. L. (1989). Guided, cooperative learning, and individual knowledge acquisition. *Knowing, Learning, and Instruction: Essays in Honor of Hillsdale*.

Ertmer, P. A., & Newby, T. J. (1996). The Expert Learner: Strategic, Self-Regulated, and Reflective. *Instructional Science*, *24*(1), 1-24.

Herman, P., Perkins, K., Hansen, M., Gomez, L. M., & Gomez, K. (2010, June). The effectiveness of reading comprehension strategies in high school science classrooms. In *Proceedings of the 9th International Conference of the Learning Sciences-Volume 1* (pp. 857-864).

Hawkins, J., & Pea, R. D. (1987). Tools for bridging the cultures of everyday and scientific thinking. *Journal for Research in Science Teaching*, *24*, 291-307.

Krathwohl, D. R., & Anderson, L. W. (2001). A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives. *Theory Into Practice*, *41*(4), 212.

Munshi, A., Rajendran, R., Ocumpaugh, J., Moore, A., & Biswas, G. (2018). Studying the Interactions between Components of Self-Regulated Learning in Open Ended Learning Environments. *International Society of the Learning Sciences, Inc.*

National Center for Education Statistics. (2005). The Condition of Education 2005 (NCES 2005–094).Washington, DC: U.S. Government Printing Office.

Norris, S. P., & Phillips, L. M. (2003). How literacy in its fundamental sense is central to scientific literacy. *Science Education*, *87*(2), 224-240.

Paris, S. G., & Paris, A. H. (2001). Classroom applications of research on self-regulated learning. *Educational Psychologist*, *36*(2), 89-101.

Pintrich, P. R., & De Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, *82*(1), 33

Reading_AL(2019).https://www.activelylearn.com/post/infographic-close-reading-strategies-with-actively-learn [Website]

Rutherford, T. (2017). Within and between person associations of calibration and achievement. *Contemporary Educational Psychology*, *49*, 226-237.

Rutherford, T., Long, J. J., & Farkas, G. (2017). Teacher value for professional development, self-efficacy, and student outcomes within a digital mathematics intervention. *Contemporary Educational Psychology*, *51*, 22-36.

Taub, M., Azevedo, R., Bradbury, A. E., Millar, G. C., & Lester, J. (2018). Using sequence mining to reveal the efficiency in scientific reasoning during STEM learning with a game-based learning environment. *Learning and Instruction*, *54*, 93-103.

Winne, P. H., Nesbit, J. C., Ram, I., Marzouk, Z., Vytasek, J., Samadi, D., & Stewart, J. (2017). Tracing Metacognition by Highlighting and Tagging to Predict Recall and Transfer. *AERA Online Paper Repository*.

Zimmerman, B. J. (1990). Self-regulated learning and academic achievement: An overview. *Educational Psychologist*, *25*(1), 3-17.

Acknowledgments

We would like to thank Actively Learn for the provision of the data and our research assistants for initial assignment content coding.