Significant Predictors of Learning from Student Interactions with Online Learning Objects

L. Dee Miller, Leen-Kiat Soh
Department of Computer Science and Engineering
University of Nebraska, Lincoln
{lmille, lksoh}@cse.unl.edu

***Abstract*—Learning objects (LOs) are self-contained, reusable units of learning. Previous research has shown that using LOs to supplement traditional lecture increases achievement and promotes success for college students in the disciplines of engineering and computer science. The computer-based nature for LOs allows for sophisticated tracking that can collect metadata about the individual learners. This tends to result in a tremendous amount of metadata collected on LOs. The challenge becomes identifying the predictors of learning. Previous research tends to be focused on a single area of metadata such as the learning strategies or demographic variables. Here we report on a comprehensive regression analysis conducted on variables in four widely different areas including LO interaction data, MSLQ survey responses (that measure learning strategies), demographic information, and LO evaluation survey data. Our analysis found that a subset of the variables in each area were actually significant predictors of learning. We also found that several static variables that appeared to be significant predictors in their own right were simply reflecting the results from student motivation. These results provide valuable insights into which variables are significant predictors. Further, they also help improve LO tracking systems allowing for the design of better online learning technologies.**

Keywords—Learning Objects, Predictors of Learning, Regression Analysis

I. INTRODUCTION

The last 20 years has seen the rapid proliferation of online instructional materials to support distance education and to supplement the traditional classroom environment. One such type of instructional material is learning objects (LOs). LOs are most often described are independent and self-standing units of learning content that are predisposed to reuse in multiple instructional contexts [1]. An example of a LO is a self-contained lesson on a computer science topic (e.g., recursion) with a tutorial, interactive exercises, and assessment questions. LOs have been used to supplement traditional lectures in college-level engineering and computer science courses [2][3]. LOs have two features that can increase achievement and promote success in students:

First, the reusable nature of LOs and the availability of searchable LO repositories mean that the quantity of available LOs is growing, not only in terms of their numbers and content, but also in their usage [4]. This results in a multiplicity of LOs on the same topic but with different content, design strategies, etc. This diversity offers a tremendous potential for providing instructional support for college students with differing attitudes, diverse backgrounds, and prior knowledge on the subject matter.

Second, the computer-based nature of LOs facilitates sophisticated data tracking, which can collect and combine information about the learners, their interactions with the LOs, and their learning efficiency into LO metadata. Tracking is typically based on learner attributes and the content and pedagogical characteristics that are expected to be associated with or predicting learning, but the capability exists for literally every interaction with the LO (i.e. mouse click) to be tracked and time stamped.

However, there is a problem with the data tracking feature for LOs. LO designers have no way of knowing a priori what variables derived from the metadata collected are going to be significant predictors of student learning. Since there is no practical way of redeploying the LOs to collect additional metadata from the same set of students, LO designers tend to err on the side of collecting too much metadata [5]. What results is a tremendous amount of metadata on individual usage of LOs as well as data aggregated across courses, learning strategies such as motivation and demographic characteristics such as gender. This resulting "mountain of metadata" makes it extremely difficult for anyone to benefit from the LO results; for example, instructors will have no idea what metadata to look for when selecting appropriate LOs for their students.

The most common solution to the data tracking problem is an a posteriori analysis of the metadata to determine which of the variables collected are actually significant predictors of student learning. After this analysis is done, the subset of significant predictors is made available and interested groups can take appropriate actions. LO content developers, for example, could improve the tracking system for current and similar LOs to focus on the significant predictors. Previous approaches on analysis of significant predictors tend to focus only on the metadata in a similar area. For instance, Chyung et al. [6] focus on MSLQ survey while Kay & Knaack [7] focused on evaluation survey. These approaches tend to assume that the significant predictors are found in a single area when, in fact, the most significant predictors may be dispersed across multiple areas. Further, they do not consider interactions between variables such as a variable dependent on another variable in a different area and, as such, not truly a significant predictor.

978-1-4673-5261-1/13/$31.00 ©2013 IEEE

We report on a regression analysis that involves comprehensive metadata from multiple areas including LO interaction data, MSLQ survey results, demographic information, and evaluation survey results. The metadata analyzed was based on 1335 distinct sessions of student interactions with LOs. These sessions involved 16 different LOs and 134 different students from three different undergraduate computer science courses at two different universities. We used a least-squares regression analysis to identify the variables that are significant predictors of learning based on the LO assessment score from that session. We provide an extensive discussion of the variables found in each area that were significant predictors and also discuss possible interactions between variables.

The rest of this paper is organized as follows. Section 2 describes the learning objects and metadata in more detail. Section 3 discusses the regression analysis. Section 4 discusses the analysis results on the LO metadata including significant predictors of learning while Section 5 provides the conclusions and future work.

II. Learning Objects

Here we provide background on the LOs used in our regression analysis. We start by describing the overall design of the LO. We then discuss, in more detail, the metadata collected during the LO deployment.

A. *Learning Object Design*

The LOs used in the regression analysis are part of the intelligent learning object guide (iLOG) framework [8]. The iLOG LOs follow the Sharable Content Object Reference Model (SCORM) standard for web-based e-learning. These LOs are designed as self-contained lessons on introductory CS concepts (e.g., searching and sorting). Each of these LOs consists of the same three basic components. First, the LO contains a tutorial with the learning objectives and a set of pages explaining the concept using text and figures (analogous to a traditional textbook). Second, it contains a set of interactive exercises that further explain the concept and provide immediate feedback for students (analogous to homework problems). Third, it contains a set of assessment questions designed to measure whether students have learned the CS concept (analogous to a quiz or test). Students are allowed to freely navigate back and forth between the tutorial and exercises, to allow for review, but once they start the assessment they are locked out of the rest of the LO.

The regression analysis below uses the student score on the LO assessment questions as the measure of whether a student has learned the CS concept. Previous work has shown that assessment scores have a strong connection with actual student learning outcomes [9][10].

B. *Metadata Collected*

The iLOG framework provides automatic, comprehensive metadata collection for the LOs [8]. Each LO contains software that automatically tracks user interactions with the LO. Examples of the user interactions tracked in LO components are given in Fig. 1. The LOs are also bundled together with surveys designed to collect additional data from the students. There are three different types of surveys: student demographic, motivated strategies for learning questionnaire (MSLQ), and LO evaluation. Students fill out these surveys as part of taking the LOs. After deployment, the iLOG framework automatically "crunches the numbers" and provides statistical metadata based on the data collected; for example, the average time spent on each page. The metadata is provided in a flat file, suitable for further analysis, with a single row for each session between a student and LO.

There are four diverse sets of metadata used in the analysis below: (1) student interaction metadata such as average time spent on a page, (2) demographic metadata such as gender and GPA, (3) MSLQ metadata such as student motivation and self-efficacy, and (4) evaluation metadata such as LO ease of use. Now, we "cast our net" and examine this wide range of diverse metadata for two main reasons. First, we expect variables that are significant predictors of learning to be dispersed across multiple areas. In this event, focusing on a single area of metadata may actually provide an incomplete picture on how the LOs impact student learning. Actions taken based on this incomplete picture are likely to result in unexpected or even unwanted results; for example, actions taken to improve LO content could negatively impact undiscovered predictors resulting in reduced student learning. Second, we expect there to be dependencies between the variable in different areas. In this event, dependent variables appear as significant predictors when only a single area is considered. Actions taken based on dependent variables are, again, likely to have unexpected results.

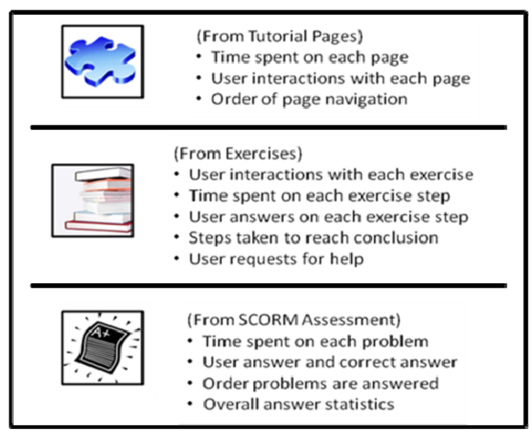

Fig. 1. Examples of user interactions tracked in iLOG LOs.

III. Regression Analysis

Here we provide details on the regression analysis used in our results and discuss previous work using regression to find significant predictors of learning.

A. *Methodology*

We use a least-squares regression analysis to determine whether a given variable is a significant predictor of learning. This process consists of two separate steps. First, we compute the slope for the least-squares estimate that best "fits" the data:

$$slope = \frac{\sum xy - (\sum x)(\sum y)/n}{\sum x^2 - (\sum x)^2/n} \quad (1)$$

where x contains the session values for the variable in question and y contains the assessment scores on same sessions and n is the number of sessions. The y-intercept is trivial to compute with the slope. Next, we plug the results of the least-squares estimate into an F-test and compute the p-value based on the distribution. If the p-value is less than or equal to 0.05, the variable is deemed a significant predictor of learning.

B. *Previous Work*

Regression analysis has previously used to find significant predictors of learning for related areas of metadata. First, regression analysis was used in a study by Kruck & Lending [11] to find the significant predictors of learning in an introductory college-level information systems course. This study considered student demographics including gender, major, SAT scores, etc. and student motivation. The authors reported that, based on a regression analysis, GPA and motivation were significance predictors of learning whereas SAT was not. One important detail is that student motivation was measured using homework grades as opposed to the more widely accepted MSLQ used in this study.

Second, regression analysis was used in a study by Chyung et al. [6] to find significant predictors of learning for nine LOs on material science and engineering concepts (e.g., mechanical properties of metals). This study gave students a MSLQ survey to measure motivation, goal orientation, and self-efficacy as well as a pre- and post-test survey to measure student learning of the concepts. Based on regression analysis, student intrinsic goal orientation and e-learning practice were significant predictors of learning but self-efficacy was not. One important detail is that authors were unable to analyze student interactions directly (as we do) since the LOs lacked tracking capability. Instead, the authors relied exclusively on self-reported student knowledge from the pre- and post-test surveys.

Third, regression analysis was used in a study by Kay & Knaack [7], to predict the validity of a wide range of LOs. This study gave students an evaluation survey asking about how much the LOs helped student learning and the quality of the LOs. The authors reported that, based on a regression factor analysis, student responses to the learning questions were significant predictors of learning. One important detail is that the factor analysis was based on a wide range of LOs from different developers. These LOs covered widely different concepts as opposed to the iLOG framework LOs on similar, CS concepts considered in this study.

To summarize, regression analysis has been previously used to find significant predictors of learning for demographic, MSLQ, and evaluation areas. However, to the best of our knowledge, no study exists that *provides a compressive analysis on metadata in all these areas*. Further, analysis to find significant predictors of learning based on actual student interactions with LOs in much more limited.

IV. RESULTS

Here we discuss the results for the comprehensive regression analysis broken down by metadata area.

A. *Student Interaction Metadata*

Table I shows that nine student interaction variables are significant predictors of learning.

First, assessmentTotalClicks is a significant predictor of learning based on the coefficient and p-value in Table I. In the assessment, students were not allowed to go back to the tutorial but they could go back and change answers to previous assessment questions. We observed that numerous times for each LO, the number of assessmentTotalClicks was greater than the minimum number of clicks required to complete the assessment test. This means that students reviewing and revisiting the test pages had higher assessment scores.

Second, we look at metadata related to student interactions with the LO tutorials. The results for tutorialTotalSeconds and tutorialAverageSecondsOnAPage indicate that a higher assessment score is associated with more time invested by students on each page of the tutorial and the overall tutorial section of the LO. Students need to spend substantial amounts of time to fully understand and comprehend the underlying notion of the LOs provided by these tutorial pages. The more time they spend studying the content on the tutorial pages the better prepared they are for the assessment test, resulting in higher assessment scores. This makes sense given that the assessment questions are designed to evaluate student understanding on the tutorial content. The results for tutorialMinClicksOnAPage show that students who interacted more each page of the LO tutorial received higher assessment scores. The more minimum number of clicks students had on tutorial pages, the higher their test scores. This makes sense because each tutorial page contains a significant amount of information requiring students to click the scroll button a few times to look at all the information provided on the page.

Third, we look at metadata related to student interactions with the interactive LO exercises. The positive results for two variables, namely exerciseAverageSecondsOnAPage and exerciseMinSecondsOnAPage show that students who had visited and spent some time on the exercises did well in the assessment test. Additionally, the results for exerciseAverageEntries, show that students who tried the exercise pages received higher test scores. On the surface, these results make sense given that each exercise is fairly complicated requiring time for students to go through completely. The benefit of the exercises is that they provide active feedback on the LOs which has been shown to improve student learning [10]. However, the other results cast doubt on the effectiveness of the LO exercises. We found that exerciseTotalInterval is a negatively associated significant predictor of learning. The exerciseTotalInterval is the time students spent reading the provided feedback after answering the questions contained in the exercise about the LO topic on an exercise page and clicking the submit button. Further, the exerciseTotalEntries, which is the total number of entries students made in the exercise pages, is also a negative predictor of student learning. Taken together, these two negative predictors seem to indicate that students who struggle with the exercises, as measured by the time spent reading the feedback and making additional entries, also struggle to do well on the assessment questions. In a sense, the exercise questions failed to prepare these students properly for the assessments questions. These results could also result in student frustration from struggling with difficult LOs—we will return to this in Evaluation section. Based on these results, one action to improve the LOs would be to revise the exercises to provide more effective feedback.

TABLE I. SIGNIFICANT PREDICTORS OF LEARNING IN STUDENT INTERACTION METADATA WITH REGRESSION COEFFICIENTS AND P-VALUES.

Student Interaction Var.	Coefficient	p-value
assessmentTotalClicks	**0.1255**	**0.0031**
tutorialTotalSeconds	**0.0028**	**0.0009**
tutorialAverageSecondsOnAPage	**0.0434**	**0.0001**
tutorialMinSecondsOnAPage	**0.2654**	**0.0002**
tutorialMinClicksOnAPage	**7.5953**	**0.0000**
exerciseAverageSecondsOnAPage	**0.0130**	**0.0081**
exerciseMinSecondsOnAPage	**0.0172**	**0.0010**
exerciseAverageEntries	**0.1410**	**0.0336**
exerciseTotalEntries	-0.615	0.446
exerciseTotalInterval	**-0.0001**	**0.0001**

B. MSLQ Metadata

Table II shows that eight MSLQ variables are significant predictors of learning.

First, we observe that Extrinsic Goal Orientation (EGO) is a negatively associated significant predictor of learning based on the negative coefficient and p-value in Table 2. Note that EGO refers to student motivation to do well in class for better grades, rewards, evaluation by others, and competition. To further investigate the negative association, we divide the data by courses. We observe that EGO is positively associated predictor with all but one course. On this course, the students were able to take the learning object again and again, with only their last assessment score being used for course grade calculation. As a result, it is possible that those students randomly answered all the questions the first time they took the test, only performing well the second time after they knew the test content. To confirm our suspicious, we found that the time spent was *significantly lower* the last time students took the assessment test compared to the first time they took the test while assessment scores are *significantly higher*.

We think that, after their introduction to the assessment test, these students realized that they could score better when they repeated the test if they learned the test content the first time. That these students were "gaming the system" to achieve higher assessment scores also helps explain the lack of significant results for Intrinsic Goal Orientation (IGO) which has previously been reported to be a significant predictor of learning [6]. Again, if we leave out students in that course, the IGO becomes a significant predictor of learning. Based on these results, one action to improve student learning with the LOs is not allowing students to retake the LO assessment multiple times. Alternatively, if instructors still want this capability, the LOs should be equipped with a larger battery of questions that are randomly presented to students making "gaming the system" less practical for students.

TABLE II. SIGNIFICANT PREDICTORS OF LEARNING IN MSLQ METADATA WITH REGRESSION COEFFICIENTS AND P-VALUES.

MSLQ Variables	Coefficient	p-value
Intrinsic Goal Orientation	0.2717	0.0912
Extrinsic Goal Orientation	**-1.0693**	**0.0000**
Control of Learning Beliefs	**1.2354**	**0.0000**
Self-Efficacy for Learning	**0.9340**	**0.0000**
Task Value	**0.8795**	**0.0000**
Elaboration	0.1245	0.5164
Organization	0.0965	0.4748
Self-Regulation - planning	**1.3040**	**0.0000**
Self-Regulation - monitoring	**0.6859**	**0.0000**
Self-Regulation - checking and correcting	**1.4880**	**0.0000**
Effort Regulation	0.1180	0.6004
Help seeking	-0.0754	0.5382
Problem-Solving	**0.6230**	**0.0006**

Second, we found that Self-Regulation - checking and correcting is a positively associated significant predictor of learning. Self-Regulation - checking and correcting refers to student checking and correcting behavior as they progress through a task. Those students who rated high in this category are more likely to change the way they study to fit course needs and the teaching styles of professors. They are also more likely to attend to feedback and make changes to improve their understanding of the course material.

Lastly, the results for the remaining four variables are expected and consistent with previous work that MSLQ variable are, in general good predictors of learning [12]. Control of Learning Beliefs means that the students control their own learning and their learning outcomes are proportional to their hard work [13]. They believe that results are not affected by the other factors such as teacher or luck. Students who put a high rate in this category strongly believe that the good assessment score they received is because of the hard work they put in. These students are likely to study the material more diligently leading to higher assessment scores. Self-Efficacy for learning is how the students judge themselves on their capability of learning the material [14]. Students who put a high rate in this category are more confident that they can learn the material. These students are likely to work harder at learning the material leading to higher assessment scores [15]. Task Value refers to the student determination about the value of the assessment [16]. The students work on the assessment if they think that the assessment is important to their score in the course. The higher the students rate on Task Value, the more attention they will pay to the assessment and thus they receive better assessment scores. Problem-Solving is a mental process that involves determining, examining, and solving problems [17]. With a high rating in Problem-Solving, it is to say that the students are able to understand the question and are capable of solving the question in the assessment.

C. Demographic Metadata

Table III shows the results for the eight demographic metadata variables. Four of these variables have categorical rather than numeric values so we use the F-test, rather than regression, to determine the p-values. Overall, seven demographic variables were found to be significant predictors of learning.

First, student's ACT score, the student's grade point average, and the student's highest math course taken (i.e., math background) are all positively associated significant predictors of learning based on the regression coefficients and p-values in Table III. These results are expected and consistent with previous work which found that ACT and GPA [18][19][20] and math background [21] are all significant predictors of student learning.

Second, student gender is not a significant predictor of learning. These results are *consistent* with previous research which has shown that gender is not a significant predictor of learning [22][23][24]. To further convince readers, since the influence of gender is still contested [25], we summarize previous work on the iLOG LOs. In the past, female students had more difficulty with these LO assessment questions than did male students [26]. This difficulty led to a systematic revision of the LOs using assessment validation tools from educational research [27]. After this revision, gender was no longer a significant predictor of learning [3].

Third, student's grade level, major, and required course all appear to be significant predictors of learning. These results are undesirable since the LOs are intended to allow students with a wide variety of backgrounds to all learn the CS topics [27].

- To explain grade level, we compared the average assessment scores for each grade level of students against all the other grade levels (e.g., freshmen vs. non-freshmen). We found that the freshmen students achieved significantly higher average assessment scores than all the other years (82.02% vs. 69.34%) based on a Welch's t-test ($p <= 0.05$) contributing to the significance of grade level as a predictor of learning. Looking further, we realized the majority of freshmen were in the CS honors course and, as such, were highly motivated to get good grades (i.e., high EGO motivation). When these honors students were removed, the significant difference between freshmen and non-freshmen students disappeared (71.35% vs. 69.34%). These results seem to indicate a dependency between grade level and EGO motivation in some students.
- To explain major and required course, we looked at the assessment scores broken down by CS students required to take the course and students taking the course as an elective. Surprisingly, non-majors taking the course as an elective achieved significantly higher average assessment scores than majors (84.58% vs. 76.57) based on a Welch's t-test. These results can be explained in terms of IGO motivation. Non-majors taking it as an elective have significantly higher average IGO than non-majors required to take the course (4.81 vs. 4.54 on the 7-point Likert scale). Further, majors with high IGO motivation (>4) achieved significantly higher average assessment scores than majors with low motivation (81.02% vs. 61.85%). These results seem to indicate a dependency between these variables and IGO motivation.

Overall, based on these results, the differences in assessment scores for the students have more to do with motivation than the grade level, major, or required course. In a sense, these dependent variables are simply reflecting the differences between students with high and low motivation. Thus, these variables should not be considered significant predictors of learning. As a result, actions taken to engage lower motivation students are likely to be more effective at improving the LOs than customizing LOs for specific grade levels or majors.

TABLE III. SIGNIFICANT PREDICTORS OF LEARNING IN DEMOGRAPHIC METADATA WITH REGRESSION COEFFICIENTS AND P-VALUES. THE (*) INDICATES CATEGORICAL VARIABLES WHERE F-TESTS WERE USED.

Demo. Variables	Coefficient	p-value
Gender*	N/A	0.0669
Grade Level*	N/A	**0.0000**
College Major*	N/A	**0.0006**
Required Course*	N/A	**0.0000**
ACT	**2.6645**	**0.0000**
GPA	**11.0135**	**0.0000**
Highest Math Course	**4.7810**	**0.0000**
Programming Courses	**-4.2619**	**0.0000**

D. Evaluation Metadata

Table IV summarizes gives the results for the ten evaluation metadata variables. Six evaluation survey variables are significant predictors of learning. The negative regression coefficients for Q8 and Q9 can be explained by the negative wording. On these two questions, students who agreed with these questions (higher survey values) struggled with learning the content more than other students providing the negative coefficients. In general, these results are consistent with previous work which found these questions to be significant predictors of student learning [28]. Based on these results, actions to improve the LOs should focus on improving LO user friendliness. One prime section for such improvements is in the exercises that, as discussed in Section IV.A, are troublesome for the students.

We looked further at the remaining four survey variables. We repeated the regression analysis broken down by course and found that the results were still not significant. To dig deeper, we broke these variables into two natural groups. Group 1 contains Q2 and Q7 that have similar wording and compare the LOs and the professor using traditional lecture. Group 2 contains Q4 and Q6 which ask students about future preferences. For Group 1, we found a significant negative correlation (dependency) between these variables and the MSLQ Self-Regulation – checking and correcting (-0.09 for Q2 and -0.08 for Q7 with DF = 400). Students with high checking and correcting seemed to prefer lectures where they could ask for immediate feedback. For Group 2, we found that non-majors were less likely to use the LOs in the future than majors (3.23 vs. 3.38, p-value 0.12). Based on these results, we have increased confidence that these questions are, indeed, not significant predictors of learning. Therefore, no actions need to be taken to address these questions.

TABLE IV. SIGNIFICANT PREDICTORS OF LEARNING IN EVALUATION METADATA WITH REGRESSION COEFFICIENTS AND P-VALUES.

Evaluation Survey Variables	Coeff.	p-value
Q1- LO was easy to use	**6.2157**	**0.0000**
Q2-LO maintained interest more than professor	1.7167	0.0606
Q3- LO was valuable addition to course	**2.7967**	**0.0014**
Q4-More course material via the web	-0.3622	0.6546
Q5- LO helped me understand this topic	**3.7838**	**0.0000**
Q6-I will use the same LO again in the future	1.1039	0.1580
Q7-Learned more from LO than professor	0.1279	0.8908
Q8-LO material was difficult to understand	**-6.5382**	**0.0000**
Q9-LO needed to go into greater detail	**-4.9595**	**0.0000**
Q10-Overall how would you rate this LO	**4.1535**	**0.0000**

V. CONCLUSIONS AND FUTURE WORK

Learning objects (LOs) are reusable instruction material that support distance education and supplement the traditional classroom environment. The computer-based nature of LOs allows for sophisticated data tracking such as recording user interactions with the LOs. However, such data tracking results in a "mountain of metadata" only some of which are actually significant predictors of learning. This makes it difficult for anyone to utilize the LO results. Previous studies have focused on a posteriori analysis to identify the metadata variables that are actually significant predictors of learning. Unfortunately, these studies tend to focus on metadata in a specific area (e.g., MSLQ or demographic) and, thus, fail to identity variables dispersed across multiple areas. Further, these studies do not take into account variables that are dependent on those in different areas and, as such, not truly significant predictors.

In this work, we found a number of significant predictors of learning based on student interactions with the LOs. Students who spent more time on the LOs and reviewed their answers on the assessment questions received significantly higher scores. Further, students with higher motivation, self-efficacy, and self-regulation also tended to receive higher scores—consistent with previous work. Lastly, student gender was not among the demographic variables found to be significant predictors of learning.

This work has two further general contributions. First, we provide a regression analysis that involves metadata from multiple areas: interactions, MSLQ, demographic, and evaluation. This regression analysis shows that, indeed, significant predictors of learning are dispersed across multiple areas. Second, we further investigate the dependencies between variables that appear (initially) to be significant predictors of learning. This investigation shows that several of variables in the demographic and evaluation areas are, in fact, dependent on variables in other areas such as student motivation.

The above contributions establish the need for more comprehensive analysis on metadata collected for LOs. We see three main avenues for continuing this work in the future.

- First, we intend to make use of the findings (predictors) to improve the LOs. Based on the findings, engaging the students with the LOs may be more effective than customizing the LO content for, say, a given major. One possible way to *further engage* students is to tap into creative thinking. The LO content could be redesigned to engage multiple senses and require imaginative thought as opposed to the current LO content that may be too dry and pedantic—like reading a text book.
- Second, we intend to continue our investigation on the previously collected metadata. This metadata shows clear signs of dependencies between variables in different areas. However, we have likely not found all the interdependencies between these variables. In particular, we still question whether several of the variables in the evaluation and demographic areas are truly significant predictors or simply reflecting other variables.
- Third, we are extremely interested as to whether the significant predictors for these LOs, which are taken individually by the students, are still applicable for students in collaborative learning environments. In particular, we intend to investigate whether motivation and other MSLQ variables remain significant predictors for teams of students working together to write online wikis or essays.

ACKNOWLEDGMENT

This material is based upon work supported by the National Science Foundation under grant no. 1122956 and IIS-0632642. The authors would also like to thank Sing Hwa Lee for his help on the initial regression analysis.

REFERENCES

[1] P. Balatsoukas, A. Morris and A. O'Brien, "Learning Objects Update: Review and Critical Approach to Content Aggregations," in Educational Technology & Society, vol. 11, 2008, pp. 119-130.

[2] G. Francia, "A Tale of Two Learning Objects," in Journal of Educational Technology Systems, vol. 3, 2003, pp. 177-190.

[3] L. D. Miller, L-K. Soh et al., "Evaluating the Use of Learning Objects in CS1," in SIGCSE, 2011, pp. 57-62.

[4] X. Ochoa and E. Duval, "Quantitative Analysis of Learning Object Repositories," in IEEE Trans. On Learning Technologies, vol. 2, 2009, pp. 226-238.

[5] C. Romero and S. Ventura, "Educational Data Mining: A Review of the State of the Art," in IEEE Transactions on Systems, Man, and Cybernetics, 2010, pp. 601-618.

[6] S. Chyung, A. Moll, and S. Berg, "The roll of intrinsic goal orientation, self-efficacy, and e-learning practice in engineering education," in Journal of Effective Teaching, vol. 10, 2010, pp. 22-37.

[7] R. Kay, L. Knaack, "Faculty of Education, Assessing learning, quality and engagement in learning objects: the Learning Object Evaluation Scale for Students (LOES-S)," in Journal of Education Technology Research Development, vol. 57, 2008, pp. 147-168

[8] L. D. Miller, L-K. Soh, A. Samal, and G. Nugent, "iLOG: a Framework for Automatic Annotation of Learning Objects with Emperical Usage Metadata," in IJAIED, vol. 21, 2011, pp. 215-236.

[9] G. Nugent, L-K. Soh et al., "Design, Development, and Validation of a Learning Object for CS1," in Journal of Educational Technology Systems, vol. 34, 2006, pp. 271-281.

[10] G. Nugent, K. Kupzyk et al., “Emperical Usage Metadata in Learning Objects,” in FIE, 2009, pp. 1285-1290.

[11] S. E. Kruck and D. Lending, “Predicting Academic Performance in an Introductory College-Level IS Course,” in Information Technology, Learning, and Performance Journal, vol. 21, 2003, pp. 9-15.

[12] A. Kitsantas, A. Winsler, and F. Huie, “Self-regulation and ability predictors of academic success during college: A predictive validity study,” in Journal of Advanced Academics, vol. 20, 2008, pp. 42-68.

[13] N. Somtsewu, The applicability of the motivated strategies for learning questionnaire (MSLQ) for South Africa. Nelson Mandela Metropolitan University, 2008.

[14] A. Bandura, “Self-Efficacy,” in *Encyclopedia of Human Behavior*, V.S. Ramachandran, Ed. Elsevier Inc., 1994, pp. 71-81.

[15] D. Shunk and F. Pajares, “Self-efficacy theory,” in *Handbook of Motivation at School*, K. R. Wentzel and A. Wigfield, Eds. Rutledge, 2009, pp. 35-54.

[16] A. Marcou and G. Philippou, “Motivational beliefs, self-regulated learning and mathematical problem solving,” Proceedings of the 29th Conference of the International Group for the Psychology of Mathematics Education, vol. 3, 2005, pp. 297-304.

[17] S. K. Reed, “Problem solving,” in *Encyclopedia of psychology*, A. E. Kazdin, Ed. Oxford University Press, 2000, pp. 71-75.

[18] H. Stumpf and J. C. Stanley, “Group data on high school grade point averages and scores on academic aptitude tests as predictors of institutional graduation rates,” in Educational and Psychological Measurement, vol. 62, 2002, pp. 1042-1052.

[19] R. Zwick and J. C. Sklar, “Predicting college grades and degree completion using high school grades and SAT scores: The role of student ethnicity and first language,” in American Educational Research Journal, vol. 42, 2005, pp. 439-464.

[20] R. Enbody, W. Punch, and M. McCullen, “Python CS1 as preparation for C++ CS2,” in SIGCSE Bull., vol. 41, 2009, pp. 116-120.

[21] B. Wilson and S. Shrock, “Contributing to Success in an Introductory Computer Science Course: A study of Twelve Factors,” in SIGCSE Bull., vol. 33, 2001, pp. 184-188.

[22] S. Beyer, K. Rynes, J. Perrault, K. Hay, and S. Haller, “Gender Differences in Computer Science Students,” in SIGCSE Bull., vol. 35, 2003, pp. 49-53.

[23] S. Bergin and R. Reilly, “Programming: Factors that Influence Success,” in SIGCSE Bull., vol. 37, 2005, 411-415.

[24] P. R. Ventura, “Identifying Predictors of Success for an Objects-First CS1,” in Computer Science Education, vol. 15, 2005, pp. 223-243.

[25] R. Kay, “Exploring the Influence of Context on Attitudes toward Web-Based Learning Tools (WBLTs) and Learning Performance,” in Interdisciplinary Journal of E-Learning and Learning Objects, vol. 7, 2011, pp. 125-142.

[26] S. Riley, L. D. Miller, L-K. Soh, A. Samal, and G. Nugent, “Intelligent Learning Object Guide (iLOG): A Framework for Automatic Emperically-Based Metadata Generation,” in AIED, 2009, pp. 515-522.

[27] L. D. Miller, L-K. Soh et al., “Revising Computer Science Learning Objects from Learner Interaction Data,” in SIGCSE, 2011, pp. 45-50.

[28] G. Nugent, K. Kupzyk, et al., “A Learning Analytic Approach to Identify Attributes of Learners and Multimedia Instruction That Influence Learning,” in ED-MEDIA, 2011, pp. 2021-2028.