Small Data in the LMS: Are We Using it Correctly?

A scholarship of teaching and learning (SOTL) project

Brief Abstract

This scholarship of teaching and learning project (SOTL) examined the use of student usage data (also referred to as small data in this paper) in the learning management system (LMS) to understand its usefulness in teaching in online courses. Findings indicated that students felt that the data was accurate but did not fully represent their learning activity and effort in the course. As a teaching practice, it is recommended that small data be used minimally or for information purposes, but not to form judgements on student effort or participation.

Abstract

This scholarship of teaching and learning project (SOTL) examined the accuracy and usefulness of student data in an online graduate course in an LMS. Student usage data are not always visible to students and students may be unaware of what data is collected related to their activity. This study examined student data and its relationship to performance, as well as how students felt about this data after reviewing it. Findings suggest that time spent in the LMS is not necessarily a measurement of student learning, but that students did feel the data of time spent in the LMS was correct. Students also felt that the data was limiting: they would like to see more of their own data, and their work outside of the LMS was not evident by usage data. As a teaching practice, using small data can be helpful but should not be used to form judgment on student work and effort.

Introduction

Can small data help us teach better? This project began after realizing that I was using small data, or student usage data, within the LMS to learn more about how much time students were spending in different parts of the online course. However, I began to question my own use of this data, wondering if it fairly represented student time spent learning, and if this type of data was even helpful for understanding students. As part of a community of practice on the scholarship of

teaching and learning (SOTL), I used SOTL to understand if using this small data was helpful to my teaching.

Definitions

Learning management system (LMS): An LMS is an application for teaching and learning which provides a digital platform for sharing materials, collaborating, discussing, assessing and interacting outside of the physical classroom. Blackboard and Canvas are examples of LMS.

Usage data: Usage data describes the numerical data available within the course reports of the LMS. This data includes time spent per active student, type of activity, items accessed and number of times accessed. In this study, we are referring to usage data as *small data*.

Literature Review

Prior research indicated that spending time in the LMS does relate to improved achievement, pending the resources within the LMS were useful to student learning. Dutt and Ismail (2018) did not find a relationship between visiting online resources and higher scores on assessments. This is similar to Castano-Munoz et al. (2013) who also did not find a relationship between the number of visits to a resource and student achievement. However, Castano-Munoz et al. (2013) also found that time spent online itself is not a cause for improvement and success in online, but instead time devoted to studying and interactive learning is the cause of success in online learning. Similarly, Firat (2016) also found a statistically significant relationship between the time spent in the LMS and a student's GPA. In addition, students felt like the LMS positively affected their academic achievement when it had certain features which encourage engagement, motivation and interaction.

Other authors have described the issues of surveillance or privacy related to increased use in learning analytics (big and small data) at higher education institutions (Dede et al, 2016; Desouza et al., 2016; Jones, 2019; York, 2020;). York (2020) suggested that this type of data can be harmful as misconceptions can be made about students, including imagining false scenarios. Jones (2019) explained that tracking and analyzing student profiles is a privacy concern, as students have few controls over what data is shown or not shown to instructors in educational platforms. However, Roberts (2019) found that students generally had positive attitudes towards learning analytics, thinking this data may help institutions with supporting students who need interventions.

Research Questions

1. What is the relationship between time spent in the LMS and academic performance in this course?
2. How does awareness of student data change student perceptions in the course?
3. How do students feel about the accuracy of their student data to measure their performance in the course?

Methods

The scholarship of teaching and learning (SOTL) is a systematic, 5-step process of evaluating one's teaching which includes identifying the research question, designing the study, collecting data, analyzing the data and drawing conclusions, and reporting the findings (Bishop-Clark and Dietz-Uhler, 2012). Within this process, researchers apply a research approach, and use reflective critique to evaluate the data to identify the impact and effectiveness of teaching. SOTL was a good choice for this study because I was reflecting on my own use of student data for teaching, to understand more about its accuracy and value.

Research Approach

This SOTL project used a mixed methods explanatory sequential design. SOTL often includes a mixed methods approach to understand the phenomena fully and from various perspectives (Divan et al., 2017). In this study, quantitative data was analyzed first, followed by analysis of qualitative data to understand how the qualitative data might explain the quantitative results (Creswell and Creswell, 2018).

Participants and Setting

This project took place within an online, graduate level research course at a regional comprehensive university in the Midwest. In this course, students learned participatory research methods. Students were given instructions on accessing their own data, and were also given a

report of their data at midterm which showed activity and time spent in the LMS. All students in the course (n=12) were invited to participate, and all students received the same treatment.

Data Collection Procedures

This two-phase design studied the accuracy of student data in Blackboard, as well as how students perceived the data. Students experienced two treatments throughout the semester. In treatment 1, students were taught how to view their own data. In treatment 2, students were provided a report from data which is viewed only by instructors. After experiencing both treatments, data collection began and continued throughout the semester. The study was approved by the human subjects board.

Data Collection Instrumentation

Existing usage data. Summary reports of time spent in the LMS were run and collected at various points in the course and archived for later use. At the end of the course time spent in Blackboard, total points earned, and activity were extracted for analysis.

Survey. An original survey was developed for this study which asked students to reply on the accuracy of the data, changes they made after viewing the data, and responses to the data. After survey development, the survey was reviewed by a SOTL community of practice.

Reflections. A discussion forum was created within the online course for reflecting on user data throughout the semester. All posts were anonymous. The prompt for this discussion included the following questions: How did you feel about seeing your user activity in Blackboard? Was it accurate? Would you use learning analytics with your own students? What concerns might there be? What value might there be? How could learning analytics in learning management systems be improved?

Analysis

Data was collected both qualitatively and quantitatively. For the analysis of usage data, descriptive statistics of frequency were used to measure changes in time spent using the LMS. Since time descriptive statistics cannot reveal a relationship between variables, a simple linear regression analysis was applied. Descriptive statistics were also used to measure the attitudes of students about the usage data. Qualitative data, gathered in both the open-ended questions of the survey and an anonymous discussion board, was analyzed through thematic analysis. Codes were

applied to extracted data and themes identified to answer the research questions. Data was not viewed or analyzed until after the semester concluded to protect student privacy.

Findings

Student Activity in the LMS

A report of student hours spent in Blackboard was run at three different intervals: at the beginning of the course (baseline), after treatment 1 and after treatment 2. The results indicated that the longest amount of time spent in Blackboard was during treatment 1, though both treatments indicated increased usage from the baseline (see Figure 1).

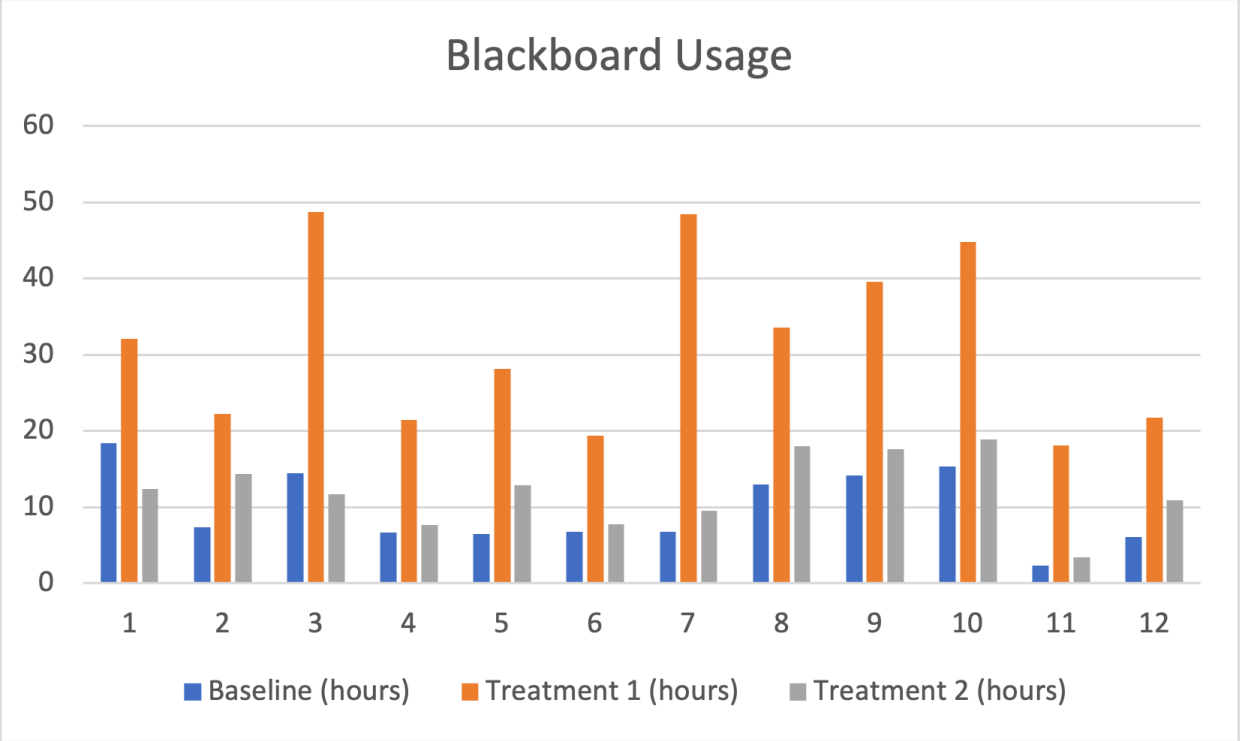


Figure 1. Hours spent in Blackboard at different intervals of the research.

Simple linear regression was used to test if hours spent in the LMS significantly predicted total success in the course.The regression analysis showed that the total time spent by the students in the LMS did not predict the total success in the course ($F_{(1,10)} = 2.908$, $p=.1190 > .05$, Adjusted $R2 = .225$). The simple scatterplot of these two variables can be seen in Figure 2.

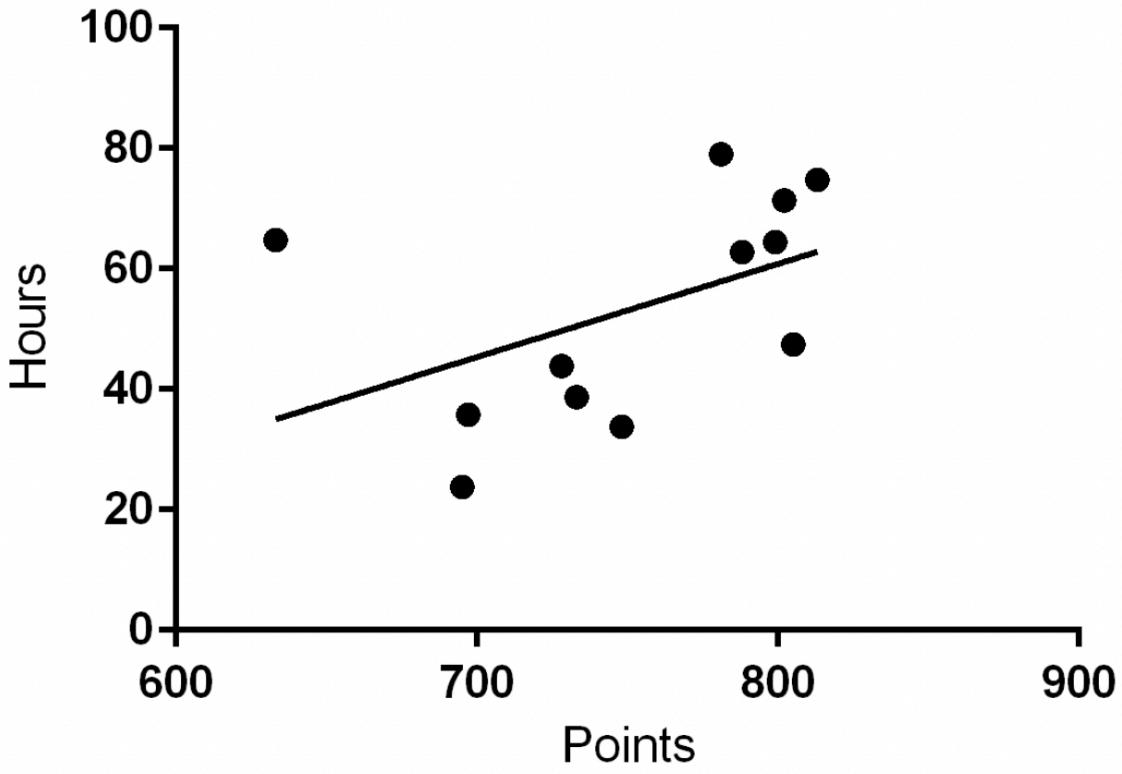
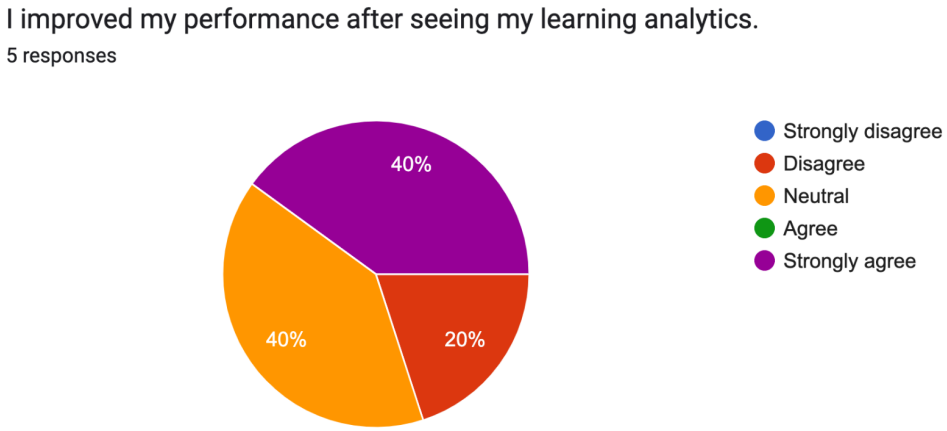


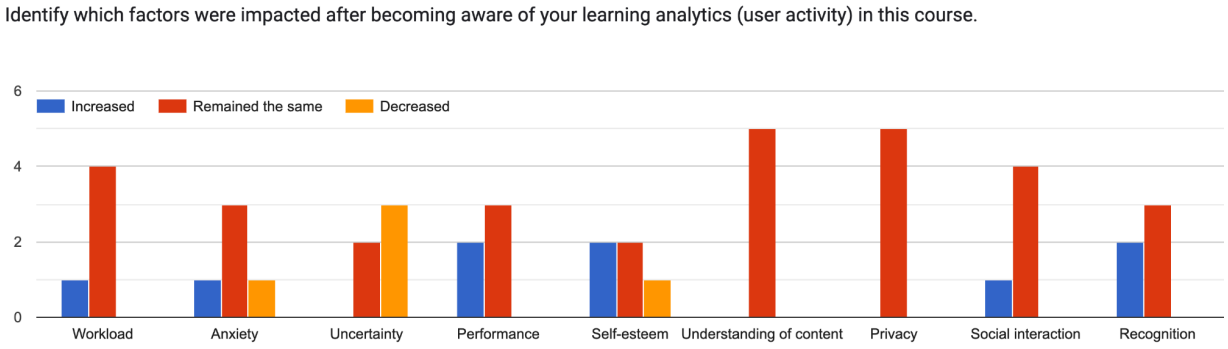
Figure 2: Scatterplot of total time spent in LMS and total points earned in the course

Changes in Activity After Awareness

Students were asked to respond to a survey which asked them several questions related to how they made changes after seeing their own usage data. Of the students in the class (n=12), 5 replied to the survey (n=5, 42%). When asked if they improved their performance, 40% (n=2) said they did, 40% (n=2) had no opinion, and 20% (n=1) said they did not change their performance (see figure 3).



Students were also asked to identify how viewing the user data impacted other psycho-social factors of the course. Of the five students who responded, 20% (n=1) indicated that it increased their workload, while 80% (n=4) remained the same. Viewing usage data reduced (n=1) and increased (n=1) anxiety, while 60% (n=3) remained the same. Uncertainty was the factor that decreased the most, with 60% (n=3) of students indicating viewing data reduced their uncertainty, and 40% (n=2) remained the same. Performance remained the same for 60% (n=3) of students, and increased for 40% (n=2). One student indicated that self-esteem decreased when they viewed their data, but self-esteem increased or stayed the same for the other four students. Viewing student data did not change how any student (n=5, 100%) felt about their understanding of content or their privacy. Social interaction remained the same for 80% (n=4) of students, and increased for one student. Recognition increased for 40% (n=2) students, but remained the same for 60% (n=3).



Students Feelings about Accuracy of Data

Students were invited to share their reflection on the accuracy and usefulness of their usage data in open-ended questions of the survey (n=5) and in an anonymous forum of the discussion board (n=3). This data was compiled and coded. Codes were organized into two themes: 1.) concerns about data being misleading, 2.) Potential for using data in professional work..

Theme 1. Concerns about data being misleading. While students indicated that they felt the data did measure their time spent in the LMS, qualitative data indicated that they were concerned about their time spent working on the course that was not represented by the data in the LMS (see Table 1). One student said, "Much of the work I do for the course is not done within Blackboard.". Students were also concerned about what the instructor might think. One student wrote that the concern about data being misleading led her to login more, stating, "I felt like I needed to be logging in to BlackBoard more often or for longer periods of time to "prove" to my professor that I was putting forth effort in the class, and I was taking it seriously."

Table 1. Student activity not included in LMS Usage Data

Working on learning activities	"The amount of time spent on Google Drive working on the material." "Activity spent within Docs that have been opened as a separate window or tab, or printed and referred back to that way" "Writing (I write most content outside of Blackboard then copy and paste it in). Thinking." "Some concerns with analytics are that they may not show all ways students have to access work. For example, the analytics that were sent to me showed my time on Blackboard, but it had no way of measuring or analyzing the amount of time I spent working with documents on Google Docs." "The Blackboard analytics have no way of showing all of the time I spent working in my separate Google Doc, so it may mislead analysts into thinking I only spent five or ten minutes on my work, when really I spent a couple of hours designing my responses for the discussion board."
Consuming learning materials	"Downloading, printing and reading articles. Reading the text. Previewing lecture slides. Attending CITI training."

	"The time I spent on preparing the lesson without connecting to BB." "Some of the sections on time spent (ex. Syllabus and schedule) showed that I spent 0.00 hours on the activity, but I added a shortcut to my own Google Drive of the document, so I accessed it there, rather than through BlackBoard."

Theme 2. Potential for using data in professional work. Students described how it was interesting to see their data, and also data from the instructor perspective. However, instead of describing concerns about instructors using data (except for one student who said, "I would also want to trust that my teacher wasn't using this data for participation points because it showed that I wasn't accessing the module page when I am, but I am looking at the embedded version rather than clicking the link to see the full page."), students described how they would use data with their own students or in libraries.

Table 2. Potential uses in professional work

- "One of the data points I can see is how long a student spends reading an article, taking a comprehension quiz, and so on. I can tell by this if students are whipping through something very quickly or if they are giving it the appropriate amount of attention."
- "When I've used learning analytics with my own students, it was great at holding my students accountable. There was also no opportunity to make up excuses (or lies), because the data was there showing how long they had accessed the document, and Google's revision history could show the minute by minute changes they may (or may not have) made."
- "I've used similar applications with my own students with assignments assigned through Google Classroom that use either the Google Docs or Slides format. The Chrome extension I used that allowed me to view the analytics of my student's work habits was Classwork Zoom."
- "In an educational setting I think making this kind of analytic information available can be very useful, both to teacher and student, as long as both parties understand how to use it, how the other is using it, and that it is very likely an imperfect source of information that should not be the only data source in making an evaluation. I use some of the activity data from a program called Achieve3000, which is a reading program my school uses for struggling readers in middle school."

Discussion & Reflection

This study examined small data within an LMS to see how closely it aligned to student activity. By asking students to review their own data (including data they can access and data they cannot access), students were best able to determine the ability of the data to provide an accurate measurement of the time spent in the LMS. The findings indicate that the data does provide a measurement for time spent in the LMS, however time spent in the LMS is just one small part of learning in an online course. Students indicated that most of their coursework and learning did not occur within the LMS itself, but through the readings, videos, discussions and collaborations which happened outside of the LMS.

There is useful small data that instructors can use within the LMS for information purposes. For example, data such as last course access is helpful to instructors to know that the student has checked in for new content. Discussion board activity can provide instructors with an indication of the number of times the student has engaged, and an item can tell an instructor if a student has accessed an important item in the course. This is useful to know if a student missed a reading or other important course related item, so that instructors can guide students to review that information if needed. In this study, students were not concerned with use of student data for this purpose, however, small data does not indicate learning, and accessing content does not necessarily lead to increased academic performance. Students described well-planned methods for accessing content on their devices or in a way that suited their study habits, such as printing, adding a bookmark to their computer, or downloading to read offline. Data about these activities is not included in LMS statistics, however this practice of information management could indeed lead to increased academic performance.

In this study, students were eager to reflect on their own data, but did so at the same time as reflecting on how they might use student data with their own students and patrons. Even though students described that the data they were shown did not fully represent their learning activity, they still described wanting to use it with their own students or patrons. They described how it could give them insight into student engagement in the platform, or how much effort students were putting in. This was interesting to me because as the instructor for their course, this is also how I had been using their data. After this study, I no longer feel that the LMS usage data is helpful to understanding student effort or engagement in the course.

Implications for Practice

There are several implications for practice that are relevant to other instructors who teach online courses using an LMS:

Instructors using student data in the LMS should remember that usage data may not show the critical learning and processing of materials which students in this study indicated occur outside of the LMS. For example, students working in another document, who copy and paste their work to the discussion board will show less activity than those who write in the discussion board. Students who download files will show less time than those who read within the interface.

Teaching students about the small data in the LMS, and how it can be useful to them, can increase students' digital literacy. As students move into the professional world, they will use small and big data and can benefit from instruction on value, privacy and transparency.

LMS designers should consider opportunities for students to view their own data. In this study, students indicated that they would like to see their own data, including how it compares to others.

References

Bishop-Clark, C. and Dietz-Uhler, B. (2012). *Engaging in the scholarship of teaching and learning*. Stylus Publishing.

Castaño-Muñoz, J., Duart, J. M. & Sancho‐Vinuesa, T. (2014). The Internet in face‐to‐face higher education: Can interactive learning improve academic achievement? *British Journal of Educational Technology, 45*(1), 149-159.

Creswell, J. & Creswell, D. (2018). *Research Design*. Sage.

Dede, C., Ho, A., and Mitros, P. (2016). Big data analysis in higher education: promises and pitfalls. *EDUCAUSE Rev.* 51, 22–34.

Desouza, K. C., and Smith, K. L. (2016). Predictive analytics: nudging, shoving, and smacking behaviors in higher education. *EDUCAUSE Rev.* 51, 10–20.

Divan, A., Ludwig, L. O., Matthews, K. E., Motley, P. M., & Tomljenovic-Berube, A. M. (2017). Survey of research approaches utilized in the scholarship of teaching and learning publications. *Teaching and Learning Inquiry*, *5*(2), 16-29.

Dutt, A., & Ismail, M. A. (2019, June). Can we predict student learning performance from LMS data? A classification approach. In the 3rd *International Conference on Current Issues in Education (ICCIE 2018)* (pp. 24-29). Atlantis Press.

Firat, M. (2016). Determining the effects of LMS learning behaviors on academic achievement in a learning analytic perspective. *Journal of Information Technology Education-Research*, *15*.

Jones, K. M. L. (2019). Learning analytics and higher education: a proposed model for establishing informed consent mechanisms to promote student privacy and autonomy. *International Journal of Educational Technology in Higher Education*, *16*(1), 2-22.

Roberts, L., Howell, J., Seaman, K. & Gibson, D. (2016). Student attitudes toward learning analytics in higher education: "the fitbit version of the learning world." *Frontiers in Psychology*, *7*.

York, E. J. (2021). Digital surveillance in online writing instruction: Panopticism and simulation in learning management systems. *Computers and Composition*, *62*, 102680.

Contact

Jenna Kammer
Associate Professor of Library Science
University of Central Missouri
jkammer@ucmo.edu

About Me

Jenna Kammer, Phd, is an Associate Professor of Library Science at the University of Central Missouri (jkammer@ucmo.edu). Prior to working as a professor, Jenna was an instructional designer and academic librarian and researches in the areas of libraries and learning, and information access in online learning.