Analyzing Learning Management System Data to Guide Teaching – Experiences from a Data-informed Approach to Configuring Online Assignments

Ryan Orchard

Analyzing Learning Management System Data to Guide Teaching – Experiences from a Data-informed Approach to Configuring Online Assignments

Author (sole author):

Ryan Orchard, MacEwan University, Edmonton, AB, Canada

Corresponding Author:

Ryan Orchard

Department of Decision Sciences, Macewan University

10700 - 104 Avenue, Edmonton, Alberta, Canada

T5J 4S2

Email: orchardr@macewan.ca

Abstract

Learning Management Systems (LMS) allow for a variety of ways in which online multiple-choice assessments ("tests") can be configured, including the ability to allow for multiple attempts and options for which of and how the attempts will count. These options are usually chosen according to the instinct of the instructor; however, LMS also provide an opportunity to make data-informed decisions based on data captured by the LMS itself. This paper describes an experience with extracting and analyzing LMS data for determining online test option settings that encourage behaviors that promote learning (and/or discourage behaviors that don't). The data extraction and cleaning process was relatively straight-forward, but not without some challenges, and required beyond-novice spreadsheet skills. It seems to have been worth the effort, though, as the insights gained led directly to a change in online test administration, and current option settings seem more appropriate and are supported by data.

Key Words

multiple attempts, online assessment, online test, learning analytics, learning management systems, data-informed teaching

Introduction

Data-Informed Teaching and Multiple Assessment Attempts

A list of the merits of evidence-based teaching that is informed by scholarship of teaching and learning (SoTL) research is unlikely to surprise many instructors in higher education, yet, even if untrained as a teacher, many "tend to choose their teaching techniques and tools based on intuition and previous experience rather than on scientific evidence of effectiveness" (American Psychological Association, 2019). In addition to the value of SoTL to guide pedagogical choices, there is another likely under-utilized resource – data from the Learning Management System (LMS) – that could be used to make data-informed decisions for teaching. Fact-based decision making is a key principle of quality management and continuous improvement frameworks, including (to name but a few) ISO 9000 and the Baldrige Award criteria for performance excellence (Summers, 2009), and the availability of data to teachers has never been greater than in recent years with the proliferation of LMS in post-secondary education and the continuing trend toward educational data mining (Cantabella, Martínez-España, Ayuso, Yáñez, & Muñoz, 2019). Thus, it may be incumbent upon an instructor to use LMS data, when possible, to guide decisions related to teaching; Clow (2013) contends that "educators have a professional responsibility to use tools and methods [such as learning analytics systems] that can improve students' learning" (p. 692). Consider this proposition as it relates to the delivery of online (LMS) assessments that consist of multiple-choice questions and allow for multiple attempts, and from the perspective of the "lay-instructor" who may not be versed in advanced data analytics.

When administering online multiple-choice assessments using an LMS, there are a multitude of question types and assessment options, with the latter including (for example) whether students may attempt an assessment multiple times (usually with different question versions appearing in

subsequent attempts) and if so, what information they are shown between the attempts and how a final grade is calculated. During the early stages of my adoption of these assessments, I allowed only a single attempt of each assessment, but found that some students were caught off-guard by the the questions that required higher-order thinking and/or Excel modeling, and many students performed worse than they had expected. I then re-configured the assessment option settings based on what *seemed* to make sense – I allowed up to two attempts but with certain limitations to what the students saw between attempts and with the condition that the final grade for the assessment would be based entirely on the last attempt submitted. (Some colleagues teaching different courses used different test option settings, but like me, they based their decision on intuition only). After a few years of hearing opinions from students and other Faculty teaching the course (all based on instinct rather than fact), I decided to analyze some of the LMS data, and I was surprised to find that my original instinct for how to configure the assessments was probably doing more harm than good, in terms of student learning (see Author, 2016, for more; further discussion will be provided later in this paper). In response, I modified the assessment option settings so that the *average* of the grades for multiple attempts is now used. Although this wasn't my initial hunch-based policy, or the most common suggestion by colleagues and students, follow-up analysis suggests that it has resulted in a change in student behaviour that is for the better (although this is debatable, as will be discussed later).

The point of the current piece is not to provide definitive instructions for online assessment option settings, although hopefully it will provoke thought and provide some useful insights on this matter, but rather it is to describe an experience with data-informed decision-making for teaching. The paper will thus first provide context, and then will move into a description of the process that was used for extracting LMS data, including a discussion about the opportunities

and challenges with the data that was available. The paper will then briefly recount the data analysis (including skills and tools required) and offer some of the findings, and will close with a discussion about making the connection between the results of the analysis and student learning and the teaching-related decision that was made. But first, a brief a brief aside is offered to consider the connection between this work and the formal topic of *Learning Analytics*, which has been described by Zhong (2016) as an "emerging trend in education especially in higher education" (p. 39) and seems worthy of attention.

Is this 'Learning Analytics'?

For the purpose of improving teaching as per the context of the current paper, the question of whether or not this analysis can be classified as "Learning Analytics" is moot; however, it was a useful exercise to investigate this, and it may bear some relevance to the reader given the growing use of the term and related terms of similar meaning.

My personal preferred definition of the increasingly prevalent (and perhaps fashionable) discipline *analytics* is that of The Institute for Operations Research and the Management Sciences [INFORMS]: "analytics [is] the scientific process of transforming data into insights for the purpose of making better decisions" (INFORMS, 2018). *Learning Analytics* (LA) is a somewhat new but growing field concerning the application of data analysis to education-related realms from administration to pedagogy. Larusson and White (2014) define LA as "…the collection, analysis, and application of data…to assess the behviour of educational communities" (p. 1) and list a number of goals, including refining pedagogical strategies, but also streamlining institutional costs and identifying students who may be at risk of failure. They state that LA "ideally attempts to leverage data to provide insight into the classroom" (p. 2). It might seem, then, that something as simple as the calculating and using the class average on an assessment to

inform teaching is an example of LA. Some (Clow, 2013; Wilson, Watson, Thompson, Drew, & Doyle, 2017) imply that LA refers to applications that involve 'Big Data' and automated analysis techniques and are initiated at the IT and Institution level (rather than in ad hoc manners by individual instructors). Examples that they provide include predictive modeling that is based on frequency and type of student interactions with the LMS and the subsequent interventions with at-risk students. Cantabella et al. (2019), who use the term *E-learning analytics*, discuss the need for a framework based on big data methods to analyze the massive amounts of LMS data, and they derive conclusions (by way of a case study) about student behaviour in terms of use of LMS tools. Zhang (2016) differentiates between *academic analytics* (for the purpose of aiding in operational and financial decision-making at the institutional level) and *learning analytics* (done for the purpose of understanding and improving learning). They use the term *analytics techniques* in a non-definitive way, but seem to imply the use of advanced methodologies on large amounts of data, based on their description of use in business (e.g. Amazon) and medical applications.

My interpretation is that the term *Analytics* is generally referring to the use of data analysis, with some sophistication beyond the simple and routine, in order to understand and improve. *Learning Analytics, then,* generally refers to the investigation of questions regarding teaching and learning that perhaps previously couldn't be answered before modern technologies, due to a lack of data and/or analysis methodologies. On the whole, the analysis project described in the current paper seems consistent with the definition of Learning Analytics offered in McKenney & Mor (2015): "teachers engage in learning analytics when they analyse learner data for the purpose of informing, refining and/or customising teaching and learning" (p. 265).

Context

LMS, and assessment option settings

Blackboard Learn (see Blackboard, 2018; referred to as *Blackboard* for the remainder of this paper) includes the ability for an instructor to develop multiple-choice assessments (referred to as *tests* in Blackboard and for the remainder of this paper). Tests in Blackboard allow for a number of different question types (17 different types in version 9.1, including anything from essay to multiple-choice to jumbled-sentence questions), and with a number of different options for any individual question type (for multiple-choice questions, options include the ability to give partial credit for some answer choices, and to randomize the order in which questions appear, for example). And then there are the test-level options; examples here (to name but a few) include the ability to randomly select question versions from a pool (which reduces opportunities for academic dishonesty), the ability to require the test to be completed in one sitting, and the ability to allow for multiple attempts (an *attempt* means *submitting* the test, but it is important to note that in any individual attempt, students may or may not be allowed to open the test, save a few question answers, and close (without submitting) as many times as they want, up to the point that they do submit or the due date passes). When allowing multiple attempts, the maximum number of attempts can be set (or it can be unlimited), and options for how multiple attempt scores are used include taking the highest, lowest, first, last, or average of the attempt scores. In addition, there are options for when and what results are shown to students after they submit an attempt and/or after the due date; for example, students can only be shown their test score after an individual attempt, or they can be shown everything, including what the correct answers were and even a pre-set explanation of the correct answer (called *feedback* in Blackboard); there are a number of options in between these two extremes as well.

Our *Introduction to Operations Management* course includes five online tests throughout the term, with an average of twelve questions per test, where each individual question is drawn from a pool of four-to-six versions of the question (all question versions being based on the same concept). Thus, in a second attempt it is likely that a student will see different versions of the individual questions, but each individual question will be *based on* the same concept as their previous attempt. We do not require the test to be completed in one sitting (they can come back to it as many times as they please up until they submit or until the due date), and we allow up to two attempts. After submitting an attempt (before the due date), students are only shown whether each answer was correct or not, but not what the correct answers are; they are shown all correct answers as well as feedback after the due date. (All test answers are hidden at the end of the school term for academic integrity purposes; overall average grades have remained consistent for years and we have not had problems with academic dishonesty – if they exist, they are minor and, considering the weight of the assignments in the overall course grade, not worth sacrificing the benefits of online tests for.) Of note is that the course and the tests strive to incorporate higher-order thinking, and some test questions include an Excel file that requires that the student build a model to determine an answer, and thus the tests are not conducive to being completed in a single sitting, which might influence a students' willingness to attempt an individual test more than once.

My choice of capping the number of allowed attempts at two was based on Yourstone, Kraye, and Albaum (2010), who found in a similar setting that two attempts lead to better learning than four attempts. However, I used instinct (i.e., my gut) to navigate the remaining options for administering the test. This included my decision for the 2015-16 school year to count only the last attempt, meaning that if a student attempted twice, only the second (most recent) attempt

would count toward their course grade. My reasoning was that the purpose of the tests was to get the students to roll up their sleeves, consult their resources (textbook, class notes, peers), and do some thinking and learning; thus, two attempts should generally lead to strong results, at least on the second attempt.

Previous analysis

At the end of the 2015-16 school year, I extracted detailed attempt data from the LMS and was quite surprised with what I saw and reported in Author (2016). The results showed that, when comparing students in the same general grade category for the final (written) exam in order to control for differences in overall academic ability, the average test grades for single attempts were higher (with statistical significance) than average grades of attempts two of two. In other words, when comparing students with similar overall abilities (based on final exam scores) online test attempts one of one were on average higher than attempts two of two. My hypothesis was that some of the students that attempted twice may have planned to do so as a way to avoid any effortful (and perhaps painful) thinking. It seemed that those that attempted once and only once, on the other hand, were more deliberate and careful. Those avoiding deep effort were very possibly being enabled by the fact that I was only counting the second of two attempts, which in essence gave them a "throwaway" first attempt to use for clues for a second attempt, sometimes perhaps as a way to avoid deep and painful thought on *either* attempt. We thus began using the *average* of the attempts in the 2016-17 school year, hoping that the absence of a throwaway attempt would encourage students to be careful and deliberate with each and every attempt, which hopefully would lead to better learning (more on this later).

I performed some follow-up data collection/analysis after the 2016-17 and 2017-18 school years and will use the latter as the context for walking through my ad hoc learning analytics project, beginning with a description of the data.

Data

The intention was to look in the LMS for data that could help to determine the best way to configure the online tests in order to promote student behaviours that enhance learning (and/or to prevent negative behaviours that interfere with learning). In general, I set out to determine *how many and which students were attempting tests twice,* and to hopefully gain some insight (from the LMS data) into *why and how students attempted tests multiple times.* In short, the data that was routinely available to me in Blackboard was helpful, but not as helpful as I had hoped, and the specific nature of the data did cause a few headaches.

Extracting Data

The literature describes LMS data in manners such as "log files from student use of the LMS" (Martin & Whitmer, 2015, p. 60) and "massive volumes of data representing click-stream or click-flow data" (Cantabella et al., 2019, p. 263). However, not only is this data not visible or accessible to the lay-instructor, but based on discussions with a Blackboard administrator at my Institution, access would require a subscription to advanced "data and analytics products" from the software vendor, which does not seem to be aimed at my type of needs, based on the website description (Blackboard Analytics, 2018). Wilson et al. (2017) argue that "it is questionable whether any…off-the-shelf…generic approach could suit [my specific] context" (p. 992), and I certainly agree when I consider what my particular intentions were. Thus, the current paper will discuss a select few of the numerous *standard* reports that are available to instructors from

within the Blackboard interface and are potentially relevant for analyzing student use of multiple test attempts.

Data accessible through standard Blackboard *Course Reports* (there were eight of these reports available in total) provide information about user activity in different areas of the LMS course site, including how much time individual students spent on each online test. However, for analyzing test attempt behavior, I found that this data wasn't separated by test attempt number, and that it must be accessed for each student separately (I could not find a way to download this data for all users at once), and thus it was not very practical. In the end, it looked *possible* that Course Report data could be used to determine how much time each student spent on each attempt and when they did so (in relation to the due date), but not without an awful lot of manual work and/or having a way to access the data on the "back-end", which wasn't possible for me, and it wasn't certain that the data would serve my purposes even if I could get it.

Data is also accessible through the Blackboard *Grade Center*, which includes *test statistics* (average, median, range, grade distribution, etc.), and detailed *item analysis* for each test question, for example. (The latter is a straight-forward tool that is standard in Blackboard and can help a teacher "recognize questions that might be poor discriminators of student performance" (Blackboard Help, n.d.) and would be an excellent starting point for faculty wishing to get started with data-informed instruction. After a test due date, I often show the item analysis results to the students for transparency, and to inspire a data-informed approach to continuous improvement.) There is also (via the Grade Center) high-level data on *test attempts*, which is what I used for the current analysis. This data is a listing of every student attempt for a given test, including grade, the date/time that the test was submitted, and the "duration" (more on this in the following section). Table One shows an excerpt. For the 2017-2018 school year (two

school terms with 145 total students and five tests per term) the data amounted to 927 records. Note that the data is provided in Blackboard as a separate formatted report for each test (ten tests in total – five in each school term), which I had to copy from a web browser and paste into Microsoft Excel for cleaning and analysis (described in the next sections).

[Insert Table 1]

Data Cleaning and Challenges

None of the data challenges were enough to prevent what I would consider a generally successful analysis, but it wasn't a breeze, either. Some notable data challenges:

- Attempts that are opened but not submitted will show as "in progress" (presumably the student chose not to complete the attempt and thus these should not be included in the analysis). One nuance of the specific LMS (Blackboard Learn 9.1 update Q4 2018 CU4) is that the data will show the 'Completed' (submitted) attempt as Attempt No. 2, and the 'In Progress' (unsubmitted) attempt as Attempt No. 1 (see 'Student10' in the data sample provided, for example). Thus, part of the pre-processing before doing any analysis was to identify these records, delete the 'In Progress' attempts from the data, and change the Attempt Number of the corresponding completed attempt from '2' to '1'. This was generally a manual process, but a savvy Excel user could employ a number of tricks to make it relatively quick to identify and correct these, once one was aware of the data issues (which took some time).
- The "duration" for each attempt is not in fact the amount of time that the student spent actively working on the test, but it is the amount of time elapsed from when the test is *first opened* until it is submitted, and thus this information was essentially without meaning for

my purposes (since the student may have worked on it many separate times since first opening, and/or had it open without working on it).

- Numerous times during the copy/paste exercise (copying from the formatted Blackboard report being displayed in a web browser and pasting into Excel) there were some completed attempts in which the report showed the second completed attempt as having been submitted *before* the first completed attempt. Oddly enough, when the report was run a subsequent time in Blackboard the data appeared differently (this time with the attempts in the correct order of completion). This observation was validated by a member of the Blackboard Admin team at my Institution (I needed to check that I wasn't going crazy!) and was quite troubling (from a data integrity perspective). In the end, I created a number of formulas in Excel to validate the data, and by the completion of my data pre-processing I was quite confident that I had caught all of the data issues, but not without some frustration and time spent.

Analysis

Skills and Software

I performed my 2017-18 analysis entirely in Excel, using Pivot Tables and a variety of other functions, including VLOOKUP and IF functions, to name the few that are beyond simple Excel functions. In my previous analysis (Author, 2016) I used Excel for data cleaning and then imported the data into Microsoft Access (database application) for analysis. Neither of these software applications were lacking in any way for my purposes; in particular Excel provides efficiencies and conveniences that are difficult to surpass for the purpose at hand, given that the types of statistics that I was calculating were not considered advanced and the data set was not large, relatively speaking, and also given that I am an experienced user and teach Excel functionality in my Operations Management course.

Analysis Results

Note – this was not a controlled experiment with what should be considered generalizable results (the point of the essay is to discuss my data-informed decision-making experience moreso than the results), but the results were certainly used to inform my own teaching, and may provide useful insights for readers.

Recall that my purpose for analyzing the test attempt data was to see whether the change in how attempt grades are counted, from only using the last completed attempt to using the average of multiple attempts, affected student behavior, since my original analysis found that attempting tests multiple times did not provide better results than attempting only once. Thus, the first order of investigation was to see how many students were attempting tests multiple times under the new settings. Considering that students of certain general academic ability in this type of a course may use multiple attempts in different ways, I again used final exam grades to partition the students into groupings for the analysis (e.g. test results for students who had a final exam grade between 70.00 and 79.99% were grouped together, 80.00 – 89.99% were grouped together, etc.) Adding this dimension to the analysis does take the complexity of the spreadsheet work required to a higher level (and possibly beyond the abilities of a novice Excel user), but is probably necessary in order to make informed decisions. The results, as per Table Two below, clearly show that for all categories, fewer students are attempting the tests multiple times since the change was made (statistical significance will be discussed later).

[Insert Table 2.]

The second area of investigation was to see if similar relative grade results persisted after the setting change, in terms of average one-attempt grades being higher than average second of two

attempt grades. Within each grouping, I compared average test results for those that attempted once (1 of 1) with those that attempted twice (1 of 2 and 2 of 2). Once again (see Figure One) I found that, for the most part, attempting once and only once is producing the best test results, even though those attempting twice got so see which questions they had wrong the first time (statistical significance discussed later). This analysis required some skill with Excel Pivot Tables, in order to be able to perform the calculations on groupings of data; these skills are intermediate/advanced Excel skills but generally not beyond the reach of the lay-instructor in Higher Education.

[Insert Figure 1.]

(A few other analyses were performed, such as comparing the number of student-tests below 40%, before and after the setting change, as well as trying to determine the number of "chronic" multiple-attempters, but the results were not conclusive and are somewhat tangential to the purpose of the current paper so are not included.)

Statistical Significance

It is possible that my analysis results happened by chance. In layperson's terms, the *statistical significance* is a measure of the probability that the same results could have occurred even if there actually was no true difference or effect; statistical significance depends on the sample size and the degree of variability within the sample results. While it may not be completely necessary to run full statistical tests on my analysis if it is simply for the sake of improving what I do in class, it does add a level of confidence in my results for myself and for my colleagues teaching the same course if I am suggesting that they use the same option settings.

For my analysis, a p-test comparing two sample proportions was used for the proportion of students-tests that are attempted twice under each of the two settings, and a t-test was used to

compare the average grade of single attempts vs. second of two attempts. Detailed statistical results are not included in this paper due to research ethics and privacy rules, and because the purpose of the paper is not to detail a formal research experiment and prove generalizable results.

Interpretation, and Implementation

I have heard many intelligent opinions and read studies about how students treat multiple attempts for online tests, and it seems that it would be very difficult to isolate conditions and construct a perfect experiment with generalizable results, for many reasons, including that different courses have different evaluation needs, sometimes considerably (which is further motivation for doing one's own analysis of LMS data to gain situation-specific insight and make informed decisions). In my Operations Management course, the tests are designed to examine the *what* and *why* (i.e. to see if a student understands what the concepts mean) as much as the *how* (whether a student can perform a routine calculation), and thus the online tests are designed to require some deliberation and higher-order thinking, which may be why one careful and deliberate attempt leads to better results than two perhaps not as careful attempts. Archer & Olson (2018) extoll the virtues of "practice makes perfect" and show that, for their setting, multiple attempts of web-based homework management systems improve learning. However, they describe mastering "quantitative disciplines" by way of "massive practice," (p. 1) which seems to focus on the *how* (to solve a specific type of problem) moreso than the *what/why* (relationships between variables and what-if-something-changes type of question), and doesn't seem to fit what I am doing with my online tests. Faulconer, Griffith, and Frank (2018) also point to the virtues of multiple attempts (although their study used LMS tests that were configured quite differently than mine). They found that second attempts score higher than first attempts (for a given individual student), which I observed with my data as well and isn't

surprising (what I found that was surprising was that those that attempt twice don't achieve better results than those that attempt once). Recently, Tila and Levy (2019) reported a similar observation for their own course ("single submitted attempts were not worse off than multiple submitted attempts", p. 1), but they also investigated student perceptions related to the opportunity to submit multiple attempts and they reported that their students were very satisfied, and the authors strongly encourage this practice. My assertion is not that multiple attempts don't have some benefit to some students in some applications, but rather it is that certain test option settings may inadvertently (and without the awareness of the instructor) be encouraging a the deliberate use of a first attempt that is intended only as a way to become familiar with the test content and gain clues for subsequent attempts, and may not promote learning overall. The studies of Baker et al. (2008) sought to understand what motivates students to *game the system*, which they define as "exploit[ing] properties of the system…rather than attempting to learn the system" (p. 185). I don't necessarily have evidence that my students are attempting to game the system on my tests, but it is difficult to explain why second of two attempt scores are on average lower than single attempt scores, and thus I want to avoid any test option settings that might enable (or encourage) gaming. Baker et al. point to "dislike of subject matter", "lack of educational self-drive" and "frustration" as the main contributing factors to gaming, which probably brings up an entirely different matter (course and test design) than which attempt counts for the grade, and gives me something else to think about.

The bottom line of my own analysis of my own course is that for no group of students (where a group is based on final exam grade) did the average grade for the second of two attempts exceed with statistical significance that for single attempts. However, the advice of Tila and Levy (2019) - that allowing multiple attempts may "cultivate and nurture a growth mindset among

community college students" (p. 11) - would suggest that there may be benefits to multiple-attempt availability that go beyond what LMS data analysis can show. My overall perception is thus that *planning to use multiple attempts* does not seem to be a good strategy for a student, but having multiple attempts as an option if needed may not be a bad thing (e.g. if one has a bad first attempt, or if one is just not good with these tests). I therefore intend to keep option settings that allow multiple attempts but *encourage* a strategy of a single, careful attempt; avoiding the best-of-multiple-attempts or most-recent-attempt approaches and using the average-of-multiple-attempts seems to help with this, for my course, based on my interpretation of the data.

Conclusions

It is probably safe to assume that the variety of ways that students can be assessed using online tools (such as those in an LMS), as well as the availability of data to an instructor, will continue to increase with time. Although extracting data from an LMS and performing analysis requires a certain amount of experience and (at the least) spreadsheet skills, it does seem incumbent upon an instructor in higher education to at least start to think about how they are making their decisions about how they use the LMS. When I analyzed data from my LMS tests, I discovered that what I thought were the right test option settings probably weren't, and it was worth the trouble to find that out. Now when I am asked the question *why not use the best attempt as the test grade*, I can give a data-informed response that usually results in buy-in from students and other faculty. My actions were primarily done of my own accord, though, and hopefully my advice can also be heard by educational leaders; as Jimerson, Choate, & Dietz (2015) point out that "equipping teachers to use data in the service of teaching and learning is a critical piece of the school improvement puzzle" (p. 204), but they also acknowledge that it is easier said than done and recommend the use of mentors, while Watson (2016) points out that not only does

moving toward fact-based decision-making “involve getting existing employees to change their mindset [and] learn new skills but it also requires cultural changes” (p.5). Indeed the organizational and management aspects of data-informed teaching are vast, but the ultimate success of any data-informed teaching initiative will rely on the motivation of the teacher to reflect it in the teaching, and the intent of this piece was to inspire just that.

References

Archer, K. K., & Olson, M. (2018). Practice. Practice. Practice. Do Homework Management Systems Work? *International Journal for the Scholarship of Teaching and Learning*, *12*, 1-5. https://doi.org/10.20429/ijsotl.2018.120212

American Psychological Association (2019). Overview of the book *Evidence-Based Teaching for Higher Education* by B.M. Schwartz & A.R. Gurung. *Retrieved from* https://www.apa.org/pubs/books/4317288

Baker, R., Walonoski, J., Heffernan, N., Roll, I., Corbett, A. & Koedinger, K. (2008). Why Students Engage in "Gaming the System" Behavior in Interactive Learning Environments. *Journal of Interactive Learning Research*, *19*, 185-224. Retrieved from https://www.learntechlib.org/primary/p/24328/

Blackboard (2018). Retrieved from https://www.blackboard.com/index.html

Blackboard Analytics (2018). Retrieved from https://www.blackboard.com/education-analytics/index.html

Blackboard Help (n.d.). Item Analysis. Retrieved from https://help.blackboard.com/Learn/Instructor/Tests_Pools_Surveys/Item_Analysis

Cantabella, M., Martínez-España, R., Ayuso, B., Yáñez, J.A., Muñoz, A. (2019). Analysis of student behavior in learning management systems through a Big Data framework. *Future Generation Computer Systems, 90*, 262-272. Retrieved from https://doi.org/10.1016/j.future.2018.08.003

Clow, D. (2013). An overview of learning analytics. *Teaching In Higher Education*, *18*, 683-695. Retrieved from https://doi.org/10.1080/13562517.2013.827653

Faulconer, E., Griffith, J., & Frank, H. (2018). If At First You Do Not Succeed: The Student Benefits of Multiple Trials on Summative Assessments. *Scholarly Commons, Embry-Riddle Aeronautical University.* Retrieved from https://commons.erau.edu/publication/964

INFORMS (2018). *Operations Research & Analytics*. Retrieved from https://www.informs.org/Explore/Operations-Research-Analytics

Jimerson, J.B., Choate, M.R., & Dietz, L.K. Supporting Data-Informed Practice Among Early Career Teachers: The Role of Mentors. *Leadership and Policy in Schools*, 14, 204-232.

Larusson, J.A., & White, B. (2014). *Learning Analytics: From Research to Practice*. New York, NY: Springer-Verlag.

Martin, F., & Whitmer, J.C. (2015). Applying Learning Analytics to Investigate Timed Release in Online Learning. *Tech Know Learn*, *21*, 59-74. Retrieved from https://doi.org/10.1007/s10758-015-9261-9

McKenney, S., & Mor, Y. (2015). Supporting teachers in data-informed educational design. *British Journal of Educational Technology,* 46, 265–279.

Summers, D.C.S. (2009). *Quality Management: creating and sustatining organizational effectiveness* (2nd ed.). Upper Saddle River, NJ: Pearson Education.

Tila, D., & Levy, D. (2019, January). Revising Online Assignments and the Impact on Student Performance at a Community College. *Community College Journal of Research and Practice*. Retrieved from https://doi.org/10.1080/10668926.2018.1564089

Watson, H.J. (2016). Creating a Fact-Based Decision-Making Culture. *Business Intelligence Journal*, 21, 5-9.

Wilson, A., Watson, C., Thompson, T. L., Drew, V., & Doyle, S. (2017). Learning analytics: challenges and limitations. *Teaching In Higher Education*, *22*, 991-1007. Retrieved from https://doi.org/10.1080/13562517.2017.1332026

Yourstone, S.A. Kraye, H.S., Albaum, G. (2010). Online quantitative-based tests - are more attempts better for learning? *Decision Sciences Journal of Innovative Education*, *8*, 347-351.

Zhong, L. (2016). A Systematic Overview of Learning Analytics in Higher Education. *Journal Of Educational Technology Development & Exchange*, *8*, 39-54. Retrieved from https://aquila.usm.edu/jetde/vol8/iss2/3

Last Name	Grade (/12)	Attempt No.	Date	Status	Duration*
Student1	12	1	2/5/2018 21:37	Completed	81:15:12
Student2	8	1	2/5/2018 13:44	Completed	89:06:59
Student3	11	1	2/2/2018 13:23	Completed	1:27:36
Student4	12	1	2/5/2018 22:05	Completed	13:00:26
Student5	8.5	1	2/5/2018 15:21	Completed	26:11:21
Student5	9	2	2/5/2018 15:46	Completed	0:22:37
Student6	9	1	2/4/2018 22:03	Completed	35:30:01
Student7	11	1	2/3/2018 17:19	Completed	120:53:37
Student8	5.5	1	2/5/2018 21:01	Completed	105:04:12
Student8	8.5	2	2/5/2018 22:30	Completed	0:46:51
Student9	7	1	2/5/2018 19:16	Completed	8:07:43
Student10	0	1	2/5/2018 22:06	In Progress	
Student10	8	2	2/5/2018 20:06	Completed	80:36:02
Student11	12	1	2/5/2018 15:50	Completed	5:46:23
Student12	10	1	1/31/2018 19:10	Completed	125:39:28
Student12	12	2	2/2/2018 9:50	Completed	38:37:18

**Duration refers to time elapsed from when a student first opens a test and when they submit; this is not necessarily the time spent actually working on the test.*

Final Exam Grade	2015-16 (only the last attempt is used)	2017-18 (average of attempts is used)
<50%	32%	23%
50-59.99%	39%	29%
60-69.99%	41%	29%
70-79.99%	38%	25%
80-89.99%	40%	25%
90-100%	28%	18%
All	38%	26%

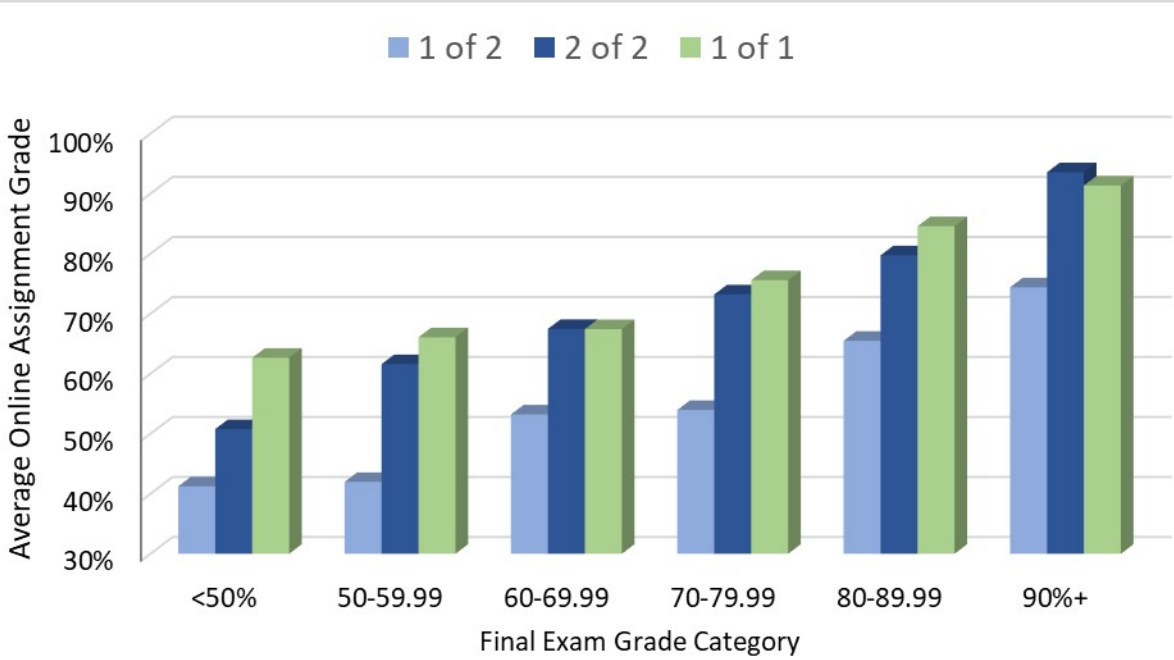
1 of 2
2 of 2
1 of 1
Average Online Assignment Grade
100%
90%
80%
70%
60%
50%
40%
30%
<50%
50-59.99
60-69.99
70-79.99
80-89.99
90%+
Final Exam Grade Category

Table 1. Excerpt of Data from 'View All Attempts' Report for a Test

Table 2. Percent of Student-Tests that were Attempted Twice, by Student Group (based on final exam grade)

Figure 1. Average Online Test Grade by Attempt Number, Categorized by Final Exam Grade