Mining the frequent use-contexts of learning management system tools and assessing their impact on learning outcomes

Taha Hassan

ARTICLE INFO

Keywords:
evaluation methodologies
data science applications in education
distance education and online learning

ABSTRACT

Measuring the utilization of learning management system (LMS) resources, while computationally expensive and hard to scale, is critical to evaluating and improving the design, management and delivery of academic course content. To that end, we propose depth-of-use (DOU): a novel, resource-specific view of the utilization of a university-wide LMS. We hypothesis-test the relationship between DOU and meta-variables like course participation (enrollment, viewership), modality (course level, mode-of-delivery, third-party app use), logistics (teaching support and digital skills training) and outcomes. In a large-scale study of metadata from over forty thousand university courses offered at [INSTITUTION NAME OMITTED] over the last three years, we find that our framing of DOU can help identify resource-level needs and preferences of micro-cohorts of courses. We discover that LMS use is near-consistently linked to better learning outcomes, and a pervasive need for scale, interoperability and ubiquitous access drives high LMS utilization. We also identify three key applications of our analyses. One, we demonstrate that DOU can help faculty members identify the opportunity-cost of transition from legacy apps to LMS services. Two, it can help instructional designers evaluate and improve their design interventions. Three, it can also aid LMS administrators in detecting three unique types of actionable low-adoption course cohorts (junk-drive, gradebook-only, and access-portal).

1. Introduction

Learning management systems (LMS) have been the primary infrastructure for hosting and disseminating information between key stakeholders in the higher education domain, near-exclusively (Coates et al., 2005; edutechnica, 2018). A contemporary LMS is a full-blown ecosystem of communication, productivity, assessment and class-management applications. Understanding the adoption and impact of these LMS apps and services is central, to university administrators, academics, and instructional designers, in improving the design, management and delivery of course content. For these reasons, there has been ample work on qualitative driving factors of LMS adoption (Berggren et al., 2005; West et al., 2007; Adeyinka & Mutula, 2010; Mtebe, 2015). However, this research is largely limited to self-reported LMS use, and there is no real consensus on how to model frequent LMS use-contexts and their relationship with learning outcomes, especially at scale. Note that a *use-context* is any meaningful set of course attributes with a potential impact on learning outcomes. This includes myriad aspects of course content, mode-of-delivery, participation, and logistics. This diversity of contextual variables, coupled with a variety of LMS data sources (app metadata, course site content, team drives, social media), and the large volume of raw LMS page requests, makes this a challenging problem.

This, in turn, informs the fundamental two-fold research question behind our study: a) what aspects of the course design, content and delivery drive the faculty and students to utilize or ignore individual LMS tools? and b) do LMS tools have a material, consistent impact on learning outcomes? To answer this research question, we assign an ordinal 'depth-of-use' (DOU) score of LMS use (low, medium or high) to each of forty-thousand college courses offered between 2017 and 2019 at [INSTITUTION NAME OMITTED]. DOU uses a taxonomy of LMS use by resource (table 1), developed in collaboration with instructional designers at [INSTITUTION NAME OMITTED], to rate the overall LMS use of an academic course. Figure 1 describes our overall approach, and section 3 details the DOU estimation method. We then hypothesis-test these scores against course attributes like modality, participation, logistics and outcomes. This lets determine the frequent contexts where faculty and staff might deem a subset of LMS services effective. For instance, we discover a consistent impact of overall LMS use on learning outcomes, and increasing reliance by faculty on tools that favor scale, ubiquitous access and interoperability. Finally, we discuss three key applications of our analyses, to a) help faculty members assess the relative utility of LMS services and legacy apps, b) aid instructional designers in measuring and improving the scope of interventions and LMS evangelism, and c) help LMS administrators identify the technology needs of actionable low-adoption cohorts.

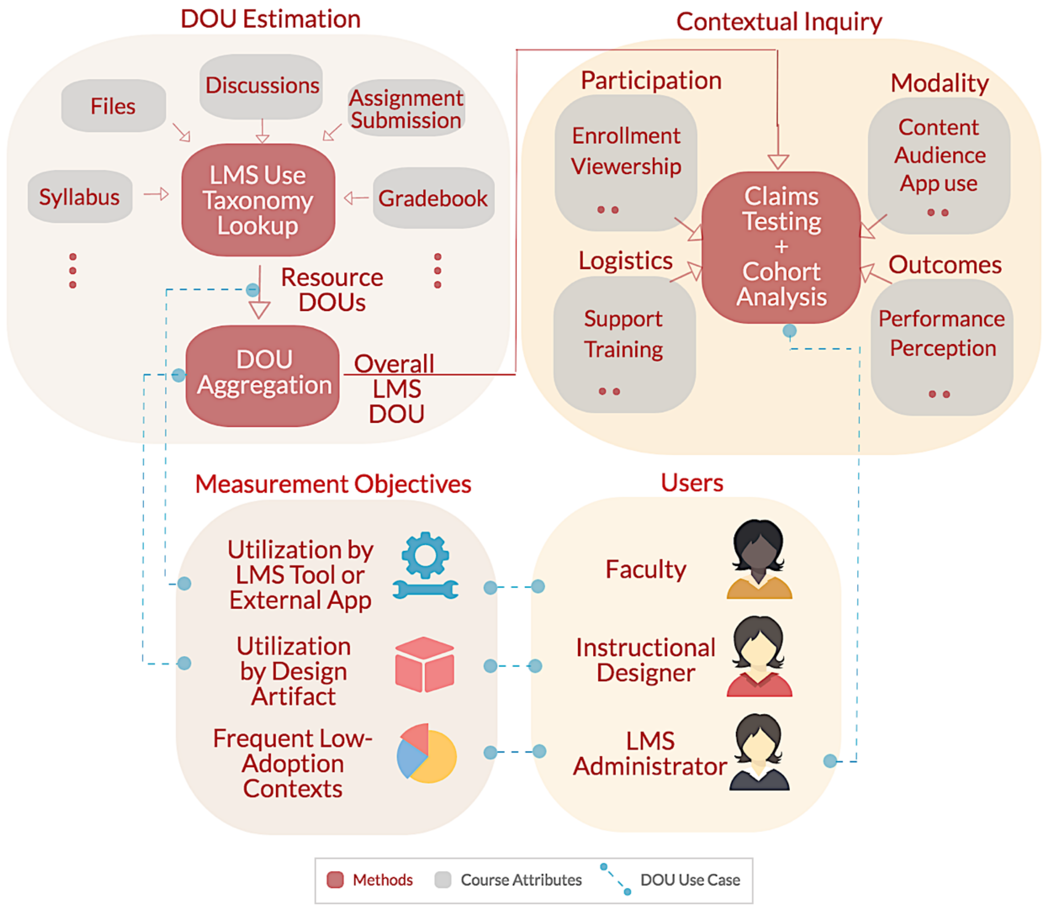

Figure 1: Depth of use (DOU) measurement overview (clockwise from top left): DOU Estimation (data sources and methods), Contextual Inquiry (hypothesis variables and methods), key user groups, and measurement objectives fulfilled by DOU

The rest of this paper is organized as follows. Section 2 extends our analysis of related work. Section 3 defines the problem of depth-of-use (DOU) estimation, and describes the research questions and corresponding hypotheses. Section 4 details the datasets, methods and results from hypothesis tests performed on DOU and its constituent dimensions. Section 5 and 6 conclude the study with a discussion of the broader faculty needs driving LMS use, and implications of our approach for data-driven design evaluation of university courses.

2. Related work

2.1. LMS adoption: human factors

There is considerable prior work on qualitative grounds for LMS adoption, like teaching and learning efficiency, generational student expectations, and institutional expansion and consolidation (Coates et al., 2005; West et al., 2007). For course instructors, the basic predictors of the pace of LMS adoption are departmental affiliation (STEM vs. non-STEM, say) and course modality (online vs. face-to-face, say). West et al. (2007) conducted semi-structured interviews with 30 college instructors over two semesters, about primary use cases, teaching efficacy and efficiency, and overall satisfaction with Blackboard LMS. The study identified so-called 'integration challenges': course instructors finding it difficult to integrate LMS services into their teaching practices. This notion of 'integration' was echoed by McGill & Klobas (2009) for the case of student adoption of WebCT, whereby students with a more favorable view of the 'task-technology fit' of LMS services were more likely to have higher LMS utilization. The authors also noted that instructor norms (instructor's view of LMS usability, support staff availability, and access to training resources) affected student utilization of LMS services favorably. Following an institution-wide transition to Canvas LMS, Wilcox et al. (2016) surveyed user perceptions on frequent modes of use and platform limitations for Canvas LMS. They identified a generation gap in expectations between students and course instructors, wherein the pervasive student use of the mobile LMS app rendered a subset of Canvas sites - designed by faculty members for the desktop - ineffective in navigation, flow and content organization.

2.2. LMS adoption: information systems

Likewise, an information systems (IS) perspective on LMS adoption has been thoroughly explored over the years (Adeyinka & Mutula, 2010; Adeyinka, 2011; Ozkan & Koseler, 2009). A bulk of these studies apply and evaluate a canonical model of IS success first discussed by DeLone & McLean (1992). The model factorizes the individual and organizational success of an IS into quality (system, information, and service), use (utilization, intention of use) and net benefits (impact on overall satisfaction, and intention of use) (Delone & McLean, 2003). Adeyinka & Mutula (2010) conducted a university-wide study of IS success factors underlying WebCT adoption and operationalized LMS utilization using nature of use (mandatory or optional), frequency of use, access and availability. They found use and intention of use both to be strong correlates of WebCT success. Fathema et al. (2015) evaluated TAM using survey data on faculty and student attitudes about Canvas LMS at two public universities. They discovered that system quality and user self-efficacy were strongly linked to system use and perceived usefulness. They also noted that system quality is a multi-faceted notion that incorporates issues like design aesthetics, flexibility of access, degree of customization, and multimedia support. Ngai et al. (2007) reported a stronger effect of the perceived usefulness and ease-of-use on system use relative to that of attitude (interest expressed towards adopting a new system). These studies largely employ user-reported system use in their analyses. Nonetheless, there are some early instances of LMS use modeling such as Ozkan & Koseler (2009), where study participants reported system use as the number of hours spent daily, on course-related activities with U-Link using a desktop or web application.

2.3. Learning analytics and educational data mining

A discussion of the key drivers of learning analytics research in Ferguson (2012) and Dawson (2010) notes how native LMS data analysis, visualization and recommendation capabilities are presently non-existent or quite limited, even with standard tracking software features. A lot of student activity is external to the LMS, the data volume is huge and ever-expanding, and there is little standardization of the data aggregation and reporting methods, viz-a-viz critical use-cases for all stakeholders involved (faculty, students, instructional designers, LMS administrators, department leadership). These problems persist even as in the past two decades, inroads in educational data mining (Romero et al., 2008; Romero & Ventura, 2010; Elbadrawy et al., 2017) have helped advance the state of the art in predictive modeling of student engagement, learning and achievement (Henrie et al., 2015; Black et al., 2008; Cocea & Weibelzahl, 2006, 2007). Simultaneously, LMS log data analyses have been used extensively to model student and faculty use-contexts (Casany Guerrero et al., 2012; Mazza & Milani, 2004), and to improve LMS features (Fenu et al., 2017), often for specific disciplines and pedagogies (Hassan et al., 2020). Improving existing pedagogies, assessing learning outcomes and risk-of-failure for students (He et al., 2015; Elbadrawy et al., 2015), and recommending interventions are all important use-cases that call for a convergence of data sources and a synthesis of approaches. One of the early instances of this approach is Course Signals at Purdue (Arnold & Pistilli, 2012). Course Signals uses students' course outcomes, frequency of interaction with the LMS (Blackboard Vista), prior academic history and demographic information to ascertain a failure-risk measurement. In Wolff et al. (2013), a short-term warning system for ailing students models the early-term drop in clickthrough rates for modules of an online course. Macfadyen & Dawson (2010) describe a similar early-warning system which identifies isolated students using an analysis of ego networks and micro-communities of high-ability students on an online course forum.

The breadth of qualitative correlates of LMS adoption reviewed in prior research highlights how complex (and potentially useful) it is to assign context to LMS data. A variety of stakeholders (figure 1) bring competing standards to evaluate the quality of the content delivered via LMS course sites. This suggests the need for a thorough, quantitative, and scalable means of evaluating LMS use by resource and context (table 1). To the best of the authors' knowledge, our work contributes a first formal, fine-grained, and vendor-agnostic method of measuring user engagement and discovering micro-cohorts of courses aboard an LMS.

2.4. Contributions

We make the following contributions in this study.

Table 1
A taxonomy of LMS use-contexts

LMS Resource	Use Context
Announcements (**An**)	**0**: None; **1**: Placeholder announcements; **2**: At least one per week or course instrument
Syllabus (**S**)	**0**: None; **1**: Syllabus under *Files*; **2**: File previewed/embedded under *Syllabus*
Discussions (**D**)	**0**: Discussions disabled; **1**: No discussion activity; **2**: Discussion groups with activity
Assignment Delivery ($\mathbf{A}_d$)	**0**: No assignments on LMS or placeholders; **1**: Link to DOC, ZIP or 3rd-party app; **2**: Assignments fully hosted on LMS
Quiz Delivery ($\mathbf{Q}_d$)	**0**: No assignments on LMS or placeholders; **1**: Link to DOC, ZIP or 3rd-party app; **2**: Quizzes fully hosted on LMS
Assignment Submission ($\mathbf{A}_s$)	**0**: No file upload, likely paper or 3rd-party app; **1**: LMS file upload; **2**: LMS text entry
Quiz Submission ($\mathbf{Q}_s$)	**0**: No online submission, likely paper or 3rd-party app; **1**: Submission within LMS
Gradebook (**G**)	**0**: No grading activity in LMS; **1**: Comprehensive grading for all assessments
Files (**F**)	**0**: No files; **1**: Course resources under *Files*

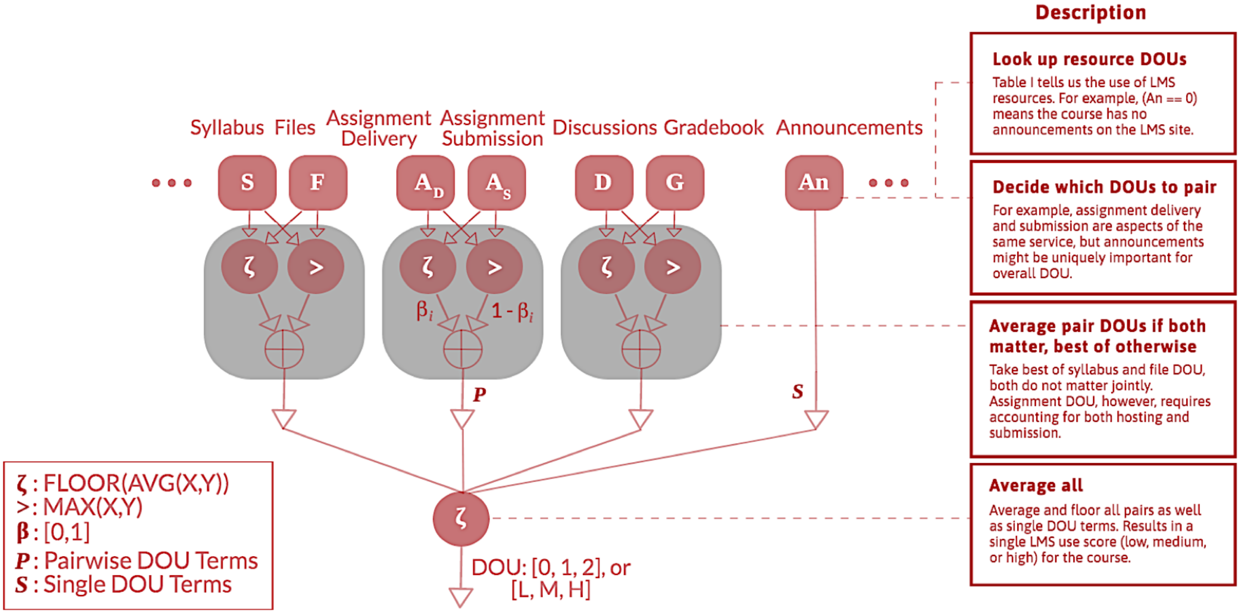

Figure 2: A schematic (left), and descriptions of all steps involved in the DOU calculation (right). S, F, D, etc. refer to LMS resource labels in table 1

- We present a first-principles, resource-specific view of course-level depth-of-use (DOU),
- We hypothesis-test the relationship between DOU and course attributes like modality, participation, logistics and outcomes,
- We identify three key use-cases of low LMS adoption (junk-drive, gradebook-only, access-portal) and survey feedback from instructional designers on ways to intervene and improve said use,

3. Depth-of-use estimation

In this section, we define a resource-level LMS depth-of-use (DOU) and describe how multiple resource DOUs can be aggregated into a single course-level DOU. We then describe four research questions (and ten corresponding hypotheses) which test how strongly DOU for a course is correlated with its modality, participation, logistics and outcomes.

3.1. Notation and definitions

Table 1 describes a taxonomy of LMS use, developed with aid from instructional designers at Virginia Tech. This taxonomy forms the basis of course-level DOU estimation in our study. We define depth-of-use for an LMS resource R_i as a simple logic rule DOU_i of the form ($R == k_i$) where k_i is a whole number. For instance, per table 1, (**An** == 1) for a given course implies *some* use of announcements (placeholders or class schedules, no instructor or TA activity). A total of N resource DOUs are accounted towards each course. As visualized in figure 2, the overall DOU for the course, DOU$_c$ is *aggregated* from the resource DOUs as follows.

$$\mathrm{DOU}_C \triangleq \zeta(P_1, P_2, ..P_{M'}, S_1, S_2, .., S_{N'}) \tag{1}$$

where

$$P_{ij} = \beta_i\Big(MAX\Big(DOU_i, DOU_j\Big)\Big) + (1-\beta_i)\Big(\zeta\Big(DOU_i, DOU_j\Big)\Big) \tag{2}$$

In equation 2, P-terms refer to pairs of LMS resource DOUs, and S-terms refer to single resource DOUs. Intuitively, we choose to pair up resource DOUs as needed, say (**An** == 1) and (**S** == 0). For each pair, we then choose between MAX() and $\zeta()$, by setting β_i to 0 or 1. $\zeta()$ is the logic equivalent of a real-valued floored-average $AVG(X, Y)$ function. $MAX()$ assigns the output to the larger of the two input contributions ($P_{\mathbf{An,S}} = 1$, for our example), while $\zeta()$ reverts to the lower of the two ($P_{\mathbf{An,S}} = 0$, that is). Picking $\beta_i = 1$ implies that the instructional staff intends to consider the MAX() or the *best* of announcement and syllabus DOUs towards the overall LMS DOU. On the other hand, $\beta_i = 0$ rewards contributions from both DOUs when necessary. This is useful say, with assignment delivery and submission considered together. It is critical to note that DOU allows flexibility in both *pairings* and *weights*, to encourage research on the usability and perceived efficacy of custom DOUs for a variety of learning environments. In addition, table 1 is vendor-agnostic, in that it can measure the use of multiple LMS ecosystems, and taxonomies for LMS services can be added or subtracted on a need-basis. Finally, in equation 1, we average all of the pairwise (P), and single (S) terms using $\zeta()$ to create a final score (low, medium or high) of overall LMS use.

3.2. Research questions and hypotheses

In this section, we describe four key research questions which address how significant the connection is between DOU and course attributes like STEM focus, mode-of-delivery, and viewership. It also explores the effect on DOU, of course instructor's work experience, and the use of training resources.

RQ1: Course type and modality What is the relationship between DOU and course type, mode of delivery, and use of third-party apps?

- **H1:** Undergraduate courses have significantly higher DOUs relative to graduate DOUs.
- **H2:** STEM courses have significantly higher DOUs relative to non-STEM courses.
- **H3:** Online-only courses have significantly higher DOUs relative to face-to-face courses.
- **H4:** Third-party app use significantly affects course DOU.

RQ2: Course participation What is the relationship between DOU and student participation in a course?

- **H5:** Course DOU is significantly linked to the number of students enrolled full-time in the course.
- **H6:** Course DOU is significantly linked to pageviews for the LMS course website.

RQ3: Course logistics What is the relationship between DOU, the size of instructional staff for the course, and participation in skills training and coursework?

- **H7:** Course DOU is significantly linked to the number of teaching staff members for the course.
- **H8:** Course DOU is significantly linked to the instructor's prior enrollment in on-demand coursework and training.

Table 2
Key counts and DOU breakdown (% **Lo**, **Med**, **Hi**) for course cohorts in the spring 2017 dataset

	Courses	Undergrad	STEM	Online	App use	Viewership (μ, σ)	Enrollment (μ, σ)	#TAs (μ, σ)	Skills
#	6117	4470	3730	981	2124	682, 5e4	49, 96	0.4, 2.8	2286
%	58, 29, 11	56, 30, 12	60, 29, 9	54, 31, 13	34, 47, 18	-	-	-	55, 33, 10

Table 3
High, medium and low DOU group composition (%) by course and instructor attributes

DOU	Undergrad	STEM	Online	App use	Skills
Low	69	63	16	20	35
Medium	75	60	19	55	42
High	83	52	23	57	33

RQ4: Course outcomes What is the relationship between course DOU, learning outcomes and student perceptions?

- **H9:** Course DOU is significantly linked to the average GPA for that course.
- **H10:** Course DOU is significantly linked to the DFW rate of that course.

4. Evaluation

4.1. Datasets

The primary dataset for this study is course page requests collected for 39580 courses during the fall and spring academic terms in 2017, 2018 and 2019, from Canvas, the enterprise LMS in operation at [INSTITUTION NAME OMITTED]. It is important to note that a university-wide transition to online instruction in the 2020 academic term in response to the COVID-19 epidemic makes the 2020 dataset unreliable for longitudinal analysis, hence excluded from our study. Tables 2 list key aspects of the courses analyzed during Spring '18, respectively. For instance, a majority (73%) of the courses offered in Spring '18 are intended for undergraduate audiences, 61% deal with STEM content, and 84% use traditional, face-to-face instructional format. These majorities are also retained in each of the three DOU groups as per table 3, with important differences. Section 6 discusses these patterns in detail.

We used a combination of manual and automated strategies (web scraping, entity resolution, and topic modeling) to create LMS utilization metadata for each course. Key textual sources include, and are not limited to, the [INSTITUTION NAME OMITTED] course catalog and historical timetable, Canvas page request logs, course descriptions on the [INSTITUTION NAME OMITTED] website (INSTITUTION NAME OMITTED, 2019), as well as syllabus files and assessment page content from Canvas course sites. STEM tagging of courses in the dataset is in accordance with the DHS classification of STEM fields US Department of Homeland Security (2016).

4.2. Methods

DOU is ordinal and not normally distributed, so we use non-parametric Kruskal-Wallis H-test (Kruskal & Wallis, 1952), in addition to an independent two-sample t-test, for hypotheses with discrete-valued meta-variables (Table 4). We evaluate group differences in viewership and enrollment for each of low, medium and high DOUs using one-way ANOVA (F-test, Table 4). To expand our analysis, we also test each of the hypotheses (ANOVA: table 5, OLS regression: figure 3) against all constituent dimensions of DOU.

4.3. Results

4.3.1. Modality (H1-H4)

As per Table 4, undergraduate courses have higher average DOUs relative to graduate courses (t-statistic is positive), consistent with their higher average enrollment (61 as opposed to 19 for graduate courses). As per table 5, undergraduate courses have higher relative DOUs for announcements ($F = 76.9, p = 2e^{-18}$), grading ($F = 119.8, p =$

Table 4
Spring 2017: Hypothesis-testing the relationship between LMS DOU and key course attributes

Hypothesis	t, p	F, p	H, p
H1: Undergraduate	8.3, 1e-16*	68.4, 1e-16*	63.9, 1e-15*
H2: STEM	-5.0, 5e-7*	25.1, 5e-7*	21.3, 3e-6*
H3: Online	3.6, 3e-4*	12.6, 3e-4*	12.1, 5e-4*
H4: App use	28.4, 2e-166*	8e2, 2e-166*	7e2, 1e-175*
H5: Enrollment	-	73.8, 3e-32*	614.4, 3e-134*
H6: Viewership	-	4.1, 1e-2*	0.9, 6e-1
H7: #TA	-	97.9, 1e-42*	1e3, 5e-289*
H8: Skills	1.5, 1e-1	2.3, 1e-1	6.3, 1e-2*
H9: GPA	-	7.4, 6e-4*	9.0, 1e-2*
H10: DFW	-	6.9, 1e-3*	2.5, 2e-1

*stat. signif., $\alpha = 0.05, p <= \alpha, F > F_{crit}, (F_{overall}, p) = (3.5^{*}, 3e^{-4})$

Table 5
Hypothesis-testing: $|t|$ and $|F|$ magnitudes for the relationship between resource DOU and course attributes

(a) announcements, syllabus, files and assignment delivery

Hypothesis	An	S	F	A_d
H1: Undergraduate	8.8*, 76.9*	5.5*, 30.2*	5.4*, 29*	-1.3, 1.8
H2: STEM	-1.5, 2.2	-5.8*, 33.4*	-1.1, 1.1	-5.7*, 32.8*
H3: Online	1.5, 2.3	3.5*, 12.5*	2.2*, 4.8*	3.9*, 14.9*
H4: App use	27.8*, 772*	29.3*, 855.7*	29.2*, 854*	0.4, 0.2
H5: Enrollment	- , 89.8*	- , 77.9*	- , 75.7*	- , 8.06*
H6: Viewership	- , 4.8	- , 3.2*	- , 7.4*	- , 4.4*
H7: #TAs	- , 89.1*	- , 79.4*	- , 117.1*	- , 0.74
H8: Skills	0.3, 0.1	3.8*, 14.6*	3.9*, 14.8*	-0.4, 0.2
H9: GPA	- , 5.5*	- , 11.2*	- , 4.6*	- , 5.9*
H10: DFW	- , 4.7*	- , 11.5*	- , 5.1*	- , 4.2*

*stat. significant, $\alpha = 0.05, p <= \alpha, F > F_{crit}$

(b) assignment submission, quiz delivery and submission, gradebook and discussions

Hypothesis	A_s	Q_d	Q_s	G	D
H1: Undergraduate	2.4*, 6*	1, 1	-0.2, 0.1	10.9*, 119.8*	-1.6, 2.4
H2: STEM	-4.9*, 24.1*	-0.7, 0.5	-1.8, 3.2	0, 0	-10.3*, 105.7*
H3: Online	3.8*, 14.2*	1.2, 1.4	0.4, 0.2	2.9*, 8.3*	8.5*, 72.1*
H4: App use	-0.3, 0.1	-0.8, 0.6	-0.8, 0.6	29.4*, 864.9*	13.5*, 181.9*
H5: Enrollment	- , 14.4*	- , 0.39	- , 4.82*	- , 160*	- , 8.75*
H6: Viewership	- , 1.2	- , 0.02	- , 2.6	- , 8.6*	- , 0.72
H7: #TAs	- , 2.6	- , 2.6	- , 0.5	- , 179*	- , 12.5*
H8: Skills	-2.7*, 7.5*	-0.6, 0.3	0.3, 0.1	-0.1, 0	2.6*, 6.7*
H9: GPA	- , 2.2	- , 0.78	- , 15.1*	- , 7.7*	- , 6.6*
H10: DFW	- , 2.6	- , 0.3	- , 15.7*	- , 7.7*	- , 4.05*

*stat. significant, $\alpha = 0.05, p <= \alpha, F > F_{crit}$

$1e^{-27}$) and online syllabi ($F = 30.2, p = 4e^{-8}$), among others. Non-STEM courses feature higher use of the LMS for assignment delivery ($F = 32.8, p = 1e^{-8}$) and submission ($F = 24.1, p = 9e^{-7}$), among others. Traditional in-class instruction loses out to online-only courses in overall DOUs. Online instruction is linked to in-depth use of online syllabi ($F = 12.5, p = 4e^{-4}$), as well as assignment delivery ($F = 14.9, p = 1e^{-4}$) and submission ($F = 14.2, p = 1e^{-4}$). Roughly 70% of each of low, medium and high DOU courses rely on third-party apps (table 3). Reliance on third-

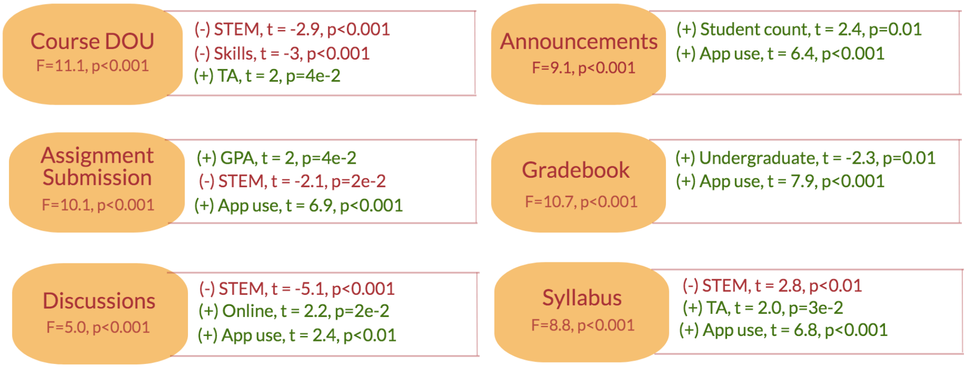

Figure 3: Positive (green) and negative (red) predictive correlates of course DOU (top-left) and a subset of resource DOUs. Each of the six panels reflect OLS regression output with the target variables and overall validity (panel-left) and their respective correlates (panel-right)

party apps coincides with the use of announcements ($F = 772, p = 4.1e^{-160}$), gradebook ($F = 864.9, p = 6e^{-178}$) and discussion forums ($F = 181.9, p = 7e^{-141}$), among others.

4.3.2. *Participation (H5, H6)*

Higher *DOU* courses feature larger overall enrollment ($F = 73.8, p = 3e^{-32}$) and viewership ($F = 4.1, p = 1e^{-2}$), as per table 4. Both of these are strong correlates of LMS utilization overall, and across a number of LMS resources considered individually (table 5). High enrollment is linked to high use of detailed online announcements ($F = 89.8, p = 6e^{-39}$), assignment delivery ($F = 8.06, p = 3e^{-4}$) and discussion forums ($F = 8.75, p = 1e^{-4}$), among others. High site viewership is similarly linked to the use of syllabi ($F = 3.2, p = 3e^{-2}$), assignment delivery ($F = 4.4, p = 1e^{-2}$) and gradebook ($F = 8.6, p = 3e^{-3}$), etc.

4.3.3. *Logistics (H7, H8)*

As per table 5, the number of teaching assistants is significantly linked to higher DOUs for announcements ($F = 89.1, p = 6e^{-39}$), discussion forums ($F = 12.5, p = 3e^{-6}$) and gradebook ($F = 179.8, p = 1e^{-40}$). Enrollment in a broad-charter professional development program does not appear to improve LMS use significantly. It is negatively linked to a number of resource DOUs.

4.3.4. *Outcomes (H9, H10)*

The average course GPA is significantly linked to overall DOU as per table 4 ($F = 7.4, p = 6e^{-4}$), and the use of announcements ($F = 5.5, p = 3e^{-3}$), syllabi ($F = 11.2, p = 1e^{-5}$) and discussion forums ($F = 6.6, p = 1e^{-3}$), among others. In comparison, DFW rate is a weaker correlate of DOU (compare the magnitudes of overall F- and H-statistics). Smaller DFW rates coincide with higher use of announcements ($F = 4.7, p = 8e^{-3}$) and gradebook ($F = 7.7, p = 5e^{-3}$), among others.

5. Applications and Discussion

Having surveyed the relationship between DOU and key course characteristics (modality, participation, logistics and outcomes), we discuss three applications of our analyses. We begin by describing how faculty members can use DOU to understand the utility of LMS services relative to legacy apps and the opportunity-cost of a future transition. We then describe how DOU can evaluate the efficacy of professional development programs and resource allocation at the department level. Finally, we describe how LMS administrators can use DOU to look for actionable low-adoption micro-cohorts of courses.

5.1. Helping faculty evaluate the opportunity-cost of LMS transition

Evident from literature surveyed in section 2 and 4, LMS adoption is a complex process, geared by the perceived quality of the overall system and the information it serves, as well as historical differences in pedagogies, and faculty-perceived opportunity-cost of transition (West et al., 2007). While determining the relative contribution of each of these

Table 6
Low-DOU course frequencies by context

Context	%	Context	%
Lo #TA ∧ No app use	79	Grad ∧ Online ∧ Lo #TA ∧ No skills	71
Lo enroll ∧ Lo #TA ∧ No app use	70	Grad ∧ Lo #TA ∧ No skills	67
Undergrad ∧ Lo #TA	67	Grad ∧ Online ∧ Lo #TA	65

factors is an open research problem, evidence in section 4 puts the need to scale as one of the most important potential correlates of LMS adoption, and provides insights into the relative utility of LMS resources for faculty expecting a transition. For instance, as per figure 3, and hypothesis **H1** in table 5a, larger class size coincides with higher or 'deeper' use of announcements, most likely because mailing lists become increasingly inefficient and harder to organize and search at scale. Larger audience sizes also coincide with more frequent LMS use for assignment submission and delivery. One key reason is that it allows for a larger range of content to be submitted and greater flexibility in scheduling and organizing take-home exams and offline evaluations. In comparison, according to hypotheses **H3** and **H5** in table 5, the use of third-party apps coincides with that of online discussion forums, but not for assignment delivery and submission. Services like Piazza are particularly favored by faculty because of their advanced forum management, content processing and tagging features, compared to the newer Discussions app aboard Canvas. This does not, however, take away from the utility of Canvas's file and assignment/quiz management apps, in part because of the ease of integration with grading apps that let course staff concurrently grade assessments without worrying about manual data imports, as well as data protection and privacy.

There are several important correlates of utilization that inform how relevant class size might be. In the previous section, we discover that while DOU is a strong positive-correlate of enrollment and viewership, graduate courses make a more exclusive use of LMS resources and have higher DOUs, with smaller class sizes on average. This points to the fact that undergraduate courses often rely on sophisticated legacy apps, especially for discussion forums (**H5**, table 5). Similarly, in table 3, which describes the fraction of courses with above-average enrollment and viewership for all DOU groups, the high DOU group has a slightly smaller fraction of these courses compared to the medium DOU group. Viewership, in contrast, is the aggregate of LMS and third-party app use, and both viewership and 3rd-party app use increase their relative share in the high DOU group.

5.2. Helping instructional designers identify opportunities for intervention

System administrators and instructional designers affiliated with the department can leverage this framework to begin to identify opportunities for meaningful LMS evangelism. DOU can point to faculty preferences about the use of legacy apps and resource allocation. For instance, in table 4, the hypothesis **H7** brings the relative utility of a comprehensive professional skills program into question (compared, for instance, to number of TAs in **H6**), as the cohort is at best indifferent to 'deeper' LMS use. Low DOU courses often frequent the cohorts with low #TAs, and faculty training alone does not appear effective in alleviating the cognitive burden of discovery required for rapid adoption. DOU can thus serve as a data-driven signal of the need for direct, personalized interventions or additional teaching support for facutly micro-cohorts. A similar picture emerges in table 6 where high enrollment courses with little to no teacher support staff results in a substantial fraction of low DOU courses (79.3%). The availability of digital skills training does not affect the wide majority (about 70%) of these courses.

5.3. Helping LMS administrators identify the needs of low-adoption cohorts

DOU can point LMS administrators to faculty preferences about the use of LMS tools and legacy apps, and their broader reasons like trade-off between teaching and research responsibilities, faculty self-efficacy, and cognitive burden of discovery (Coates et al., 2005; West et al., 2007). Table 6 describes some example low-adoption cohorts which highligh the connection between student viewership, course modality and logistics (low DOU courses are 58% of the dataset overall). In an expert review session with five instructional designers at [INSTITUTION NAME OMITTED], we identified three distinct types of low-adoption use-contexts, and their implications for design interventions.

5.3.1. Junk-drive

According to table 2, the overall frequency of low-DOU courses in the dataset is 58%. Compare these with the frequency of low-adoption courses for several micro-cohorts in table 6. These frequencies echo the connection between

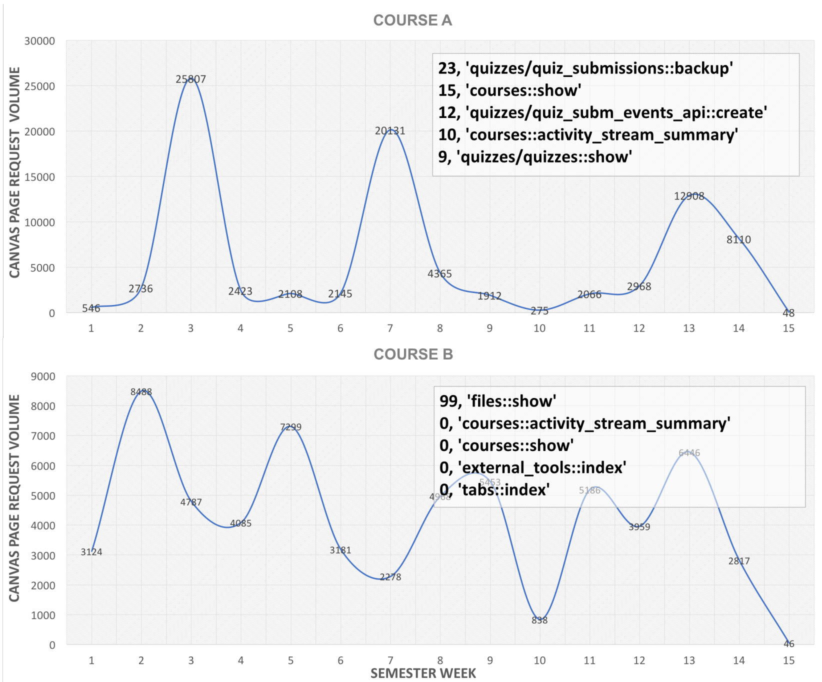

Figure 4: LMS page request volume and types for the 'junk-drive' use-case. Page requests for course A (top) occur around three key course milestone deadlines, and reflect a variety of LMS resources. Course B (bottom) uses the LMS site as a storage drive with no student activity.

Table 7
Junk-drive, gradebook-only and access-portal courses - group compositions

Course attribute	Junk-drive (%)	Gradebook-only (%)	Access portal (%)
Undergraduate	63	77	66
STEM	65	66	60
Online	12	12	15
3rd-party app use	18	58	100
Enrollment (Lo, Hi)	78, 21	57, 42	84, 16
Viewership (Lo, Hi)	36, 63	96, 3	98, 1
#TAs (Lo, Hi)	97, 2	80, 19	95, 4
Skills training	56	40	39

instructor and student engagement and how key aspects of course content and logistics might affect the system and information quality experienced by students while interacting with the LMS. An interesting scenario emerges in the connection between viewership and DOU. Figure 4 visualizes the weekly average pageviews for two STEM courses with medium weekly viewership and vastly different DOUs. The share of page requests by category (application controller::action) reveals the differences in LMS utilization: course B is primarily being used as a file drive despite having gone through design intervention. Course A, on the other hand, reveals heavy LMS use around two key deadlines for the course and a surge in page views early on in the semester (corresponding to add-drop period for the term). We identified 114 low-DOU high-viewership courses in the spring 2017 dataset. Table 7 details their attributes. They are slightly more likely to be undergraduate and STEM courses, and despite about half of them reporting digital skills training, only about 18% report the use of third-party apps. Nearly all of these courses do not have teaching assistance and the class sizes are mostly small, so there is evidence of instructor use or experimentation, however preliminary, with native LMS services. This micro-cohort is an important example of the potential for continual LMS evangelism and instructional support in order to drive up adoption rates.

5.3.2. Gradebook-only

According to table 7, exclusive use of the Canvas gradebook likely coincides with medium to high-enrollment, undergraduate (77%), STEM (66%) courses, with heavy reliance on third-party apps (58%), and an abundance of labs, recitations and group projects. Digital skills training is particularly ineffective for this cohort, which brings to attention its scarce teaching support staff (80%). It simultaneously points to the need for design interventions that help reduce the cognitive burden of faculty looking to make a fuller transition to LMS discussion forums, groups and assessments, especially at scale.

5.3.3. Access-portal

This micro-cohort refers to course sites that are collections of links to third-party apps. Per table 7, these courses are often undergraduate, STEM and unresponsive to digital skills training. Such an extreme reliance on these apps is often a function of both department-level precedents and faculty-perceived ease-of-use. This implies that a design intervention for this micro-cohort should make a particular note of faculty's technology self-efficacy and access to teacher support (note the high fraction of low #TA courses) in end-of-semester quality assessments.

6. Conclusions and Future Work

In depth-of-use, we devise a multi-factor, resource-specific view of LMS utilization. DOU helps us examine a variety of use-contexts in faculty and student adoption of LMS services. Our hypothesis-testing reveals that the needs for scale, ubiquitous access and interoperability drive a broad swath of courses across departments towards higher LMS use. We also discover that DOU helps us isolate low-adoption course clusters and reflect on faculty preferences, pedagogies, and patterns in administrative policies that might play a part in sustaining such clusters. Our research aims to combine expertise from course planning, policy design and quality assurance in order to test multi-level claims of efficacy and recommend interventions that leverage the totality of contextual evidence of historical LMS use.

Our dataset and analysis describes all Canvas course sites commissioned during the fall and spring academic terms of 2017, 2018 and 2019, at [INSTITUTION NAME OMITTED]. Its scope can be broadened in several important ways. We examine these as directions of future work as follows. To aid generalizability, we intend to reproduce our analyses for Scholar LMS - in use prior to Canvas - at [INSTITUTION NAME OMITTED]. We also plan to compare our results with courses hosted aboard Canvas at peer institutions. Beyond between-LMS and between-institution studies, we hope to hypothesis-test DOU as a function of course modality (flipped and blended classrooms (Dias & Diniz, 2014)), and content and system quality (example pervasiveness (Warnick, 2009), cognitive task models (Masapanta-Carrión & Velázquez-Iturbide, 2018; Prasad, 2018), early availability of course content, site aesthetics (Martin et al., 2008), mobile platform support (Casany Guerrero et al., 2012; Casany et al., 2012) and accessibility (Wilcox et al., 2016)), in order to analyze their impact on the usability of LMS services.

We also intend to expand the characterization of LMS use by resource (table 1) to include the use of content recommenders. The domain of educational recommendation has a large volume of literature on highly specialized interventions aimed at a multitude of use-contexts (Manouselis et al., 2011). Incorporating the use of topic, course and supplementary content recommenders in DOU can help evaluate if a specific low-adoption cohort is responsive, in perceived ease-of-use, novelty, trust and satisfaction (Hassan & McCrickard, 2019; Hassan, 2019), to instructor-aided curation of study materials. We also plan to account for user-activity within third-party apps hosted by the LMS. We plan to collaborate with several app vendors to better understand the relative satisfaction with interactional and content quality these apps might provide. Finally, the scope of our analysis is interpretive in that it examines the observed LMS usage as a function of high-level course meta-characteristics. In our future work, we plan to concurrently model instructor preferences, habits, and values that make up the said usage. We plan to incorporate instructor work experience and familiarity with instructional design practices in our analyses. We also plan to collect feedback from instructors and students, using semi-structured interviews and online academic forum analyses (Hassan et al., 2019) for key low-adoption micro-cohorts to better summarize and validate these reasons.

References

Adeyinka, T. (2011). Reliability and factor analysis of a blackboard course management system success: A scale development and validation in an educational context. *Journal of Information Technology Education: Research*, *10*, 55–80.

Adeyinka, T., & Mutula, S. (2010). A proposed model for evaluating the success of webct course content management system. *Computers in Human Behavior*, *26*, 1795–1805.

Arnold, K. E., & Pistilli, M. D. (2012). Course signals at purdue: Using learning analytics to increase student success. In *Proceedings of the 2nd international conference on learning analytics and knowledge* (pp. 267–270). ACM.

Berggren, A., Burgos, D., Fontana, J. M., Hinkelman, D., Hung, V., Hursh, A., & Tielemans, G. (2005). Practical and pedagogical issues for teacher adoption of ims learning design standards in moodle lms. *Journal of Interactive Media in Education*, *2*, 2.

Black, E. W., Dawson, K., & Priem, J. (2008). Data for free: Using lms activity logs to measure community in online courses. *The Internet and Higher Education*, *11*, 65–70.

Casany, M. J., Alier, M., Mayol, E., Piguillem, J., Galanis, N., García-Peñalvo, F. J., Conde, M. Á. et al. (2012). Moodbile: A framework to integrate m-learning applications with the lms. *Journal of Research and Practice in Information Technology*, *44*, 129.

Casany Guerrero, M. J., Alier Forment, M., Galanis, N., Mayol Sarroca, E., & Piguillem Poch, J. (2012). Analyzing moodle/lms logs to measure mobile access. In *UBICOMM 2012: The Sixth International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies* (pp. 35–40).

Coates, H., James, R., & Baldwin, G. (2005). A critical examination of the effects of learning management systems on university teaching and learning. *Tertiary education and management*, *11*, 19–36.

Cocea, M., & Weibelzahl, S. (2006). Can log files analysis estimate learners' level of motivation? In *LWA*. University of Hildesheim, Institute of Computer Science.

Cocea, M., & Weibelzahl, S. (2007). Cross-system validation of engagement prediction from log files. In *European Conference on Technology Enhanced Learning* (pp. 14–25). Springer.

Dawson, S. (2010). 'seeing'the learning community: An exploration of the development of a resource for monitoring online student networking. *British Journal of Educational Technology*, *41*, 736–752.

DeLone, W. H., & McLean, E. R. (1992). Information systems success: The quest for the dependent variable. *Information systems research*, *3*, 60–95.

Delone, W. H., & McLean, E. R. (2003). The delone and mclean model of information systems success: a ten-year update. *Journal of management information systems*, *19*, 9–30.

Dias, S. B., & Diniz, J. A. (2014). Towards an enhanced learning management system for blended learning in higher education incorporating distinct learners' profiles. *Journal of Educational Technology & Society*, *17*, 307–319.

edutechnica (2018). 6th Annual LMS Data Update. URL: `https://edutechnica.com/2018/10/06/6th-annual-lms-data-update/`.

Elbadrawy, A., Rangwala, H., Johri, A., & Karypis, G. (2017). SDM 2017 Tutorials: Opportunities, challenges and methods for higher education data mining. URL: `https://www-users.cs.umn.edu/~elbad004/sdm2017_tutorial/higher_edm_SDM_2017.pdf`.

Elbadrawy, A., Studham, R. S., & Karypis, G. (2015). Collaborative multi-regression models for predicting students' performance in course activities. In *Proceedings of the Fifth International Conference on Learning Analytics And Knowledge* (pp. 103–107). ACM.

Fathema, N., Shannon, D., & Ross, M. (2015). Expanding the technology acceptance model (tam) to examine faculty use of learning management systems (lmss) in higher education institutions. *Journal of Online Learning & Teaching*, *11*.

Fenu, G., Marras, M., & Meles, M. (2017). A learning analytics tool for usability assessment in moodle environments. *Journal of e-Learning and Knowledge Society*, *13*.

Ferguson, R. (2012). Learning analytics: drivers, developments and challenges. *International Journal of Technology Enhanced Learning*, *4*, 304–317.

Hassan, T. (2019). On bias in social reviews of university courses. In *Companion Publication of the 10th ACM Conference on Web Science* (pp. 11–14). ACM.

Hassan, T., Edmison, B., Cox, L., Louvet, M., & Williams, D. (2019). Exploring the context of course rankings on online academic forums. In *2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)* (pp. 553–556). IEEE.

Hassan, T., Edmison, B., Cox II, L., Louvet, M., Williams, D., & McCrickard, D. S. (2020). Depth of use: an empirical framework to help faculty gauge the relative impact of learning management system tools. In *Proceedings of the 25th ACM Conference on Innovation and Technology in Computer Science Education (ITiCSE '20)*. ACM.

Hassan, T., & McCrickard, D. S. (2019). Trust and trustworthiness in social recommender systems. In *Companion Proceedings of the 2019 World Wide Web Conference (WWW '19 Companion)*. ACM.

He, J., Bailey, J., Rubinstein, B. I., & Zhang, R. (2015). Identifying at-risk students in massive open online courses. In *AAAI* (pp. 1749–1755).

Henrie, C. R., Bodily, R., Manwaring, K. C., & Graham, C. R. (2015). Exploring intensive longitudinal measures of student engagement in blended learning. *The International Review of Research in Open and Distributed Learning*, *16*.

INSTITUTION NAME OMITTED (2019). Upcoming Offerings, INSTITUTION NAME OMITTED. URL: `https://URL.OMITTED/`.

Kruskal, W. H., & Wallis, W. A. (1952). Use of ranks in one-criterion variance analysis. *Journal of the American statistical Association*, *47*, 583–621.

Macfadyen, L. P., & Dawson, S. (2010). Mining lms data to develop an "early warning system" for educators: A proof of concept. *Computers & education*, *54*, 588–599.

Manouselis, N., Drachsler, H., Vuorikari, R., Hummel, H., & Koper, R. (2011). Recommender systems in technology enhanced learning. In *Recommender systems handbook* (pp. 387–415). Springer.

Martin, L., Martínez, D. R., Revilla, O., Aguilar, M. J., Santos, O. C., & Boticario, J. G. (2008). Usability in e-learning platforms: heuristics comparison between moodle, sakai and dotlrn. In *The 7th Europian Conference on e-Learning, Agia Napa, Cyprus*. Citeseer.

Masapanta-Carrión, S., & Velázquez-Iturbide, J. Á. (2018). A systematic review of the use of bloom's taxonomy in computer science education. In *Proceedings of the 49th ACM Technical Symposium on Computer Science Education* (pp. 441–446). ACM.

Mazza, R., & Milani, C. (2004). Gismo: a graphical interactive student monitoring tool for course management systems. In *International Conference on Technology Enhanced Learning, Milan* (pp. 1–8).

McGill, T. J., & Klobas, J. E. (2009). A task–technology fit view of learning management system impact. *Computers & Education*, *52*, 496–508.

Mtebe, J. (2015). Learning management system success: Increasing learning management system usage in higher education in sub-saharan africa.

International Journal of Education and Development using ICT, *11*.
Ngai, E. W., Poon, J., & Chan, Y. (2007). Empirical examination of the adoption of webct using tam. *Computers & Education*, *48*, 250–267.
Ozkan, S., & Koseler, R. (2009). Multi-dimensional students' evaluation of e-learning systems in the higher education context: An empirical investigation. *Computers & Education*, *53*, 1285–1296.
Prasad, P. (2018). Developing students' cognitive processes required for software design verification. In *Proceedings of the 2018 ACM Conference on International Computing Education Research* (pp. 284–285). ACM.
Romero, C., & Ventura, S. (2010). Educational data mining: a review of the state of the art. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, *40*, 601–618.
Romero, C., Ventura, S., & García, E. (2008). Data mining in course management systems: Moodle case study and tutorial. *Computers & Education*, *51*, 368–384.
US Department of Homeland Security (2016). STEM Designated Degree Program List. URL: `https://www.ice.gov/sites/default/files/documents/Document/2016/stem-list.pdf`.
Warnick, B. R. (2009). *Imitation and education: A philosophical inquiry into learning by example*. SUNY Press.
West, R. E., Waddoups, G., & Graham, C. R. (2007). Understanding the experiences of instructors as they adopt a course management system. *Educational Technology Research and Development*, *55*, 1–26.
Wilcox, D., Thall, J., & Griffin, O. (2016). One canvas, two audiences: How faculty and students use a newly adopted learning management system. In *Society for Information Technology & Teacher Education International Conference* (pp. 1163–1168). Association for the Advancement of Computing in Education (AACE).
Wolff, A., Zdrahal, Z., Nikolov, A., & Pantucek, M. (2013). Improving retention: predicting at-risk students by analysing clicking behaviour in a virtual learning environment. In *Proceedings of the third international conference on learning analytics and knowledge* (pp. 145–149). ACM.

A. Hypothesis tests by semester

Tables 8 through 12 describe the hypothesis-tests for DOU and course metadata from fall 2017, spring 2018, fall 2018, spring 2019 and fall 2019 academic terms at [INSTITUTION NAME OMITTED].

Table 8
Fall 2017: Hypothesis-testing the relationship between DOU and key course attributes

Hypothesis	t, p	F, p	H, p
H1: Undergraduate	9.0, 2e-19*	81.1, 2e-19*	75.8, 3e-18*
H2: STEM	-4.5, 8e-6*	19.8, 8e-6*	22.4, 2e-6*
H3: Online	0.5, 6e-1	0.3, 6e-1	0.0, 9e-1
H4: App use	32.3, 2e-211*	1e3, 2e-211*	1e3, 2e-223*
H5: Enrollment	-	69.1, 2e-30*	974.8, 2e-212*
H6: Viewership	-	4.4, 1e-2*	3.6, 1e-1
H7: #TA	-	108.5, 5e-47*	1e3, 2e-250*
H8: Skills	6.1, 8e-10*	37.7, 8e-10*	52.8, 3e-13*
H9: GPA	-	3.8, 2e-2*	4.2, 1e-1
H10: DFW	-	5.3, 5e-3*	2.5, 2e-1

Table 9
Spring 2018: Hypothesis-testing the relationship between DOU and key course attributes

Hypothesis	t, p	F, p	H, p
H1: Undergraduate	9.0, 4e-19*	80.3, 4e-19*	75.3, 4e-18*
H2: STEM	-4.5, 6e-6*	20.4, 6e-6*	20.8, 5e-6*
H3: Online	4.6, 3e-6*	21.5, 3e-6*	19.5, 9e-6*
H4: App use	45.3, 1e-5*	2e3, 1e-5*	1e3, 1e-5*
H5: Enrollment	-	73.0, 6e-32*	969.5, 2e-211*
H6: Viewership	-	4.2, 1e-2*	0.5, 7e-1
H7: #TA	-	91.0, 1e-39*	1e3, 3e-242*
H8: Skills	3.7, 2e-4*	13.8, 2e-4*	20.9, 4e-6*
H9: GPA	-	3.6, 2e-2*	2.8, 2e-1
H10: DFW	-	3.0, 5e-2*	0.3, 8e-1

Table 10
Fall 2018: Hypothesis-testing the relationship between DOU and key course attributes

Hypothesis	t, p	F, p	H, p
H1: Undergraduate	5.8, 7e-9*	33.6, 7e-9*	33.1, 8e-9*
H2: STEM	0.9, 3e-1	0.7, 3e-1	0.6, 4e-1
H3: Online	2.2, 2e-2*	4.7, 2e-2*	5.3, 2e-2*
H4: App use	1.1, 2e-1	1.1, 2e-1	0.9, 3e-1
H5: Enrollment	-	9.7, 6e-5*	170.9, 7e-38*
H6: Viewership	-	5.9, 2e-3*	4.2, 1e-1
H7: Skills	-3.6, 3e-4*	12.7, 3e-4*	13.0, 3e-4*
H8: #TA	-	5.9, 2e-3*	76.8, 2e-17*
H9: GPA	-	7.8, 4e-4*	10.4, 5e-3*
H10: DFW	-	7.0, 8e-4*	1.3, 5e-1

Table 11
Spring 2019: Hypothesis-testing the relationship between DOU and key course attributes

Hypothesis	t, p	F, p	H, p
H1: Undergraduate	6.3, 2e-10*	40.0, 2e-10*	38.9, 4e-10*
H2: STEM	-0.8, 4e-1	0.6, 4e-1	0.7, 4e-1
H3: Online	1.5, 1e-1	2.4, 1e-1	2.9, 9e-2
H4: App use	1.4, 1e-1	2.0, 1e-1	1.8, 1e-1
H5: Enrollment	-	10.2, 3e-5*	181.9, 3e-40*
H6: Viewership	-	1.5, 2e-1	2.6, 2e-1
H7: Skills	-5.0, 6e-7*	25.0, 6e-7*	25.0, 5e-7*
H8: #TA	-	3.4, 3e-2*	61.7, 4e-14*
H9: GPA	-	1.5, 2e-1	0.4, 8e-1
H10: DFW	-	3.5, 2e-2*	2.7, 2e-1

Table 12
Fall 2019: Hypothesis-testing the relationship between DOU and key course attributes

Hypothesis	t, p	F, p	H, p
H1: Undergraduate	21.4, 4e-99*	4e2, 4e-99*	4e2, 4e-103*
H2: STEM	6.4, 1e-10*	40.6, 1e-10*	45.1, 1e-11*
H3: Online	-1.4, 0.14	2.1, 0.14	1.8, 0.18
H4: App use	36.3, 4e-267*	1e3, 4e-267*	1e3, 1e-280*
H5: Enrollment	-	95.1, 3e-41*	1e3, 8e-285*
H6: Viewership	-	1.5, 0.22	8.7, 0.01*
H7: #TA	-	146.5, 3e-63*	1e3, 0.0*
H8: Skills	6.9, 4e-12*	48.1, 4e-12*	73.2, 1e-17*
H9: GPA	-	0.7, 0.51	0.6, 0.74
H10: DFW	-	0.6, 0.52	0.8, 0.66