The Internet and Higher Education 56 (2023) 100882

Contents lists available at ScienceDirect

The Internet and Higher Education

journal homepage: www.elsevier.com/locate/iheduc

Credit hours is not enough: Explaining undergraduate perceptions of course workload using LMS records

Zachary A. Pardos [a,*], Conrad Borchers [b], Run Yu [c]

[a] University of California, Berkeley, 2121 Berkeley Way, Berkeley 94720, CA, USA
[b] Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh 15213, PA, USA
[c] New York University, 251 Mercer Street, New York 10012, NY, USA

ARTICLE INFO

Keywords:
Learning analytics
Course load
Course design
Learning management system
Enrollment data
Survey research

ABSTRACT

Credit hours traditionally quantify expected instructional time per week in a course, informing student course selection decisions and contributing to degree requirement satisfaction. In this study, we investigate course load measures beyond this metric, including determinants from course assignment structure and LMS interactions. Collecting 596 course load ratings on time load, mental effort, and psychological stress, we investigate to what extent course design decisions gleaned from LMS data explain students' perception of course load. We find that credit hours alone explain little variance compared to LMS features, specifically number of assignments and course drop ratios late in the semester. Student-level features (e.g., satisfied prerequisites and course GPA) exhibited stronger associations with course load than the credit hours of a course; however, they added only little explained variance when combined with LMS features. We analyze students' perceived importance and manageability of course load dimensions and argue in favor of adopting a construct of course load more holistic than credit hours.

1. Introduction

Credit hours is an attribute of a course at an institution of higher education expected to correlate with the amount of time a student can expect to spend on the course. Giving a thorough historical account of the credit hour, Heffernan (1973) describes how the introduction of the elective system at Harvard in 1869 through Charles Eliot brought the need for standardized quantification of student progress. Initially, measurements of credit hours were based on hours of weekly classroom contact. The idea of standardizing credit hours around classrtoom time further expanded in adoption via the Carnegie unit, which linked faculty eligibility for the Carnegie Foundation's retirement fund to satisfying a 120 instruction hours per year standard (Shedd, 2003).

Around the same time, the student hour was described in a report titled, "Academic and industrial efficiency: A report to the Carnegie Foundation for the Advancement of Teaching" (Cooke, 1910). It represented one hour of work of one student on lectures, lab work, or recitation and was adopted by public institutions to measure the cost and output of teaching and research (i.e., instructional cost per student hour). The Carnegie unit and the student hour are foundational to university administration today and have since spilled over to the European Credit Transfer and Accumulation System (ECTS), which similarly bases its accounting of credit hours on student time spent on courses (Silva, White, & Toch, 2015). In 2009, the U.S. Department of Education defined a credit hour as a minimum of one hour of classroom instruction and two hours of out-of-class work per week (Laitinen, 2012). The University of California at Berkeley (UC Berkeley), the location of our study, considers one credit hour to represent one hour of in-class instruction per week, with students given the guidance that this translates approximately to an additional two to three hours spent studying outside of class per credit hour.[1]

Although initially designed with instructor contact hours rather than student hours in mind, credit hours are highly relevant to students' course selection and academic outcomes. This attribute can be regarded as the single piece of official information about course workload conveyed to students when choosing their courses. Auxiliary sources include peers, advisers, and inferences made from reading the syllabus if available during registration. The credit hours attribute is therefore instrumental in students selecting a set of courses for a term that is manageable with respect to workload. Prior literature suggests that

* Corresponding author.
E-mail addresses: pardos@berkeley.edu (Z.A. Pardos), cborcher@cs.cmu.edu (C. Borchers), run.yu@nyu.edu (R. Yu).

[1] https://academic-senate.berkeley.edu/coci-handbook/2.3.1.

https://doi.org/10.1016/j.iheduc.2022.100882
Received 22 April 2022; Received in revised form 14 August 2022; Accepted 17 August 2022
Available online 28 August 2022

students with high GPAs tend to choose a higher term load compared to their peers (Cummings & Knott, 2001). Higher term load has also been observed at institutions with fixed-term tuition (i.e., as opposed to per-credit tuition), where students may be incentivized to exceed the normative number of courses taken per term in order to attempt to reduce their cost and time to degree (Bound, Lovenheim, & Turner, 2007).

Given the coarse-grain resolution of credit hours, typically ranging from 1–4 at UC Berkeley, and its focus on instruction time, the question arises how precisely credit hours can encapsulate students' course workload experiences. At the same time, the digital age allows the collection of several data sources that may account for student workload. The modern university poses a wide breadth of challenges and tasks for students to undertake (e.g., projects, lab reports, writing assignments), which may not adequately be represented by a course's credit hours. Next to discrepancies with regards to the actual time students spend on courses, credit hours only represent the dimension of *time* in course load. This precludes course dimensions such as the intrinsic difficulty of the material or stress factors induced by instructional decisions (e.g., the placing of deadlines). We ask how adequate the notion of credit hours is in describing students' university course load experiences. To investigate this question, we turn to psychological literature to define how workload is measured in different domains.

Beyond time load, we propose two additional components to student workload in higher education: Mental effort and psychological stress. Within psychology, workload has been routinely studied in industrial and organizational psychology. We find studies that investigate the link between workload and organizational commitment (Ahuja, Chudoba, George, Kacmar, & McKnight, 2002; Dee Cuyper & De Witte, 2006) as well as workload and well-being (Geurts, Kompier, Roxburgh, & Houtman, 2003; Ilies, Dimotakis, & De Pater, 2010). These studies measure workload via subjective survey assessment. Items from these scales often include references to overload and workload pressure inducing subjective stress (Spector & Jex, 1998; Janssen, 2001; Bowling & Kirkendall, 2012). Workload is nevertheless more than stress and multi-faceted (Bowling, Alarcon, Bragg, & Hartman, 2015).

Another variable that can increase workload is tasks requiring a high degree of *mental effort*, which may be defined as the complexity and cognitive difficulty of completing a task (Paas, Van Gog, & Sweller, 2010). Arguably, mental effort has been neglected in the literature due to its multiple facets and influence factors (e.g., practice, strategies to deal with tasks) that make it a complex construct to measure and quantify (Kantowitz, 1987). This has led to very different operationalizations and approaches to measuring mental effort, ranging from self-report (Hart & Staveland, 1988) to physiological data such as eye-tracking (Palinko, Kun, Shyrokov, & Heeman, 2010). Nevertheless, there is reason to assume that mental effort might be vital in higher education workload, as it was found to be in high school science learning (Kalyuga, Ayres, Chandler, & Sweller, 2003).

Interactions have been observed between workload and STEM degree persistence. For example, prerequisites of higher education courses indicate a degree of prior knowledge required to deal with the content of a course successfully and subsequently with the ability to persist in a STEM degree (Sithole et al., 2017). Furthermore, prior research has noted declines in STEM student motivation in the last week of the semester (i.e., finals week) in higher education that might relate to increased end-of-term workload (Young, Wendel, Esson, & Plank, 2018). In this study, we engineer features from the learning management system and enrollment data that, analogous to industrial and organizational psychology studies investigating the complexity and excess of tasks, are expected to be associated with student workload.

We draw from the Subjective Workload Assessment Technique (SWAT) by Reid and Nygren (1988) to measure time load, mental effort, and psychological stress and adapt its wording to the educational context. We ask: Is credit hours a reasonable reflection of this more general workload associated with a course that students will consider when balancing their term selections? Do individual students differ in their perceptions or ability to manage course workload? Finally, if the credit hours attribute of a course does not accurately represent its workload, then what is a more accurate representation?.

We explore potential determinants of course load, broken out by student self-reported time load, mental effort, and psychological stress for $N = 596$ courses taken by $N = 127$ students during the Spring '21 semester at UC Berkeley. We employ correlational analysis and linear mixed modeling to course design and student interaction features derived from learning management system (LMS) and enrollment data that may relate to course load. We extend on a previous study showing that the prediction of student course load leaves considerable room for improvement (Chockkalingam et al., 2021). Given these gaps regarding the factors contributing to student course load and the emerging possibility of mining LMS and enrollment data to investigate them, we defined two research questions:

RQ 1: Which course design choices contribute to students' perceptions of course load, broken out by time load, mental effort, and psychological stress? How does this compare to the credit hours of a course ascribed by the instructor?

RQ 2: How specific to the individual are student perceptions of course load? Are students' academic attributes associated with their perceptions of course load?

2. Literature review

There has been little research exploring and comparing contributors to student course load systematically. However, studies have focused on single determinants of course load and their correspondence to student outcomes. For example, a survey study linked test anxiety with students' perceptions of course load and their course work time management ability (Sansgiry & Sail, 2006). Other studies have investigated the effects of the COVID-19 pandemic on students' workload, academic outcomes, and well-being in higher education. A recent study by Huntington-Klein and Gill (2021) investigating longitudinal data from a four-year university with a high average time-to-degree suggested that students often avoid lower grades in high time load courses by reallocating free time otherwise planned for non-educational activities. Smith (2019), however, observed that high time load is not without other consequences, finding that students' self-reported time load and time pressures were significantly associated with negative well-being. A longitudinal study during the first and second wave of the COVID-19 pandemic later demonstrated the adverse effects of the pandemic on perceived student stress and mediating effects of coping styles between perceived stress and life satisfaction (Rogowska, Kuśnierz, & Ochnik, 2021). These studies also operate on a conceptualization of course workload experiences of students that extend beyond credit hours or time load measures alone. In particular, they emphasize that personal resources to deal with course load moderate the effects of course load, relating to our second research question regarding individual differences in course load perception.

In the following literature review, we survey research from educational psychology to describe previous attempts at operationalizing course load. In addition, we describe studies related to our methodology and research aims from the learning analytics literature. Finally, we take a broader look into work related to this study's implications for course design to contextualize our first research question.

2.1. Measuring course load

Hart and Staveland (1988) evaluated 10 workload-related factors from different experiments. They began with the assumption that workload is a hypothetical construct that represents the cost incurred by a human operator to achieve a particular level of performance. Therefore, they defined workload as human-centered rather than task-centered, and workload ratings were subject to a variety of task- and operator-specific sources of variability. They proposed the following sources: task difficulty, time pressure, performance, mental sensory effort, physical effort, frustration level, stress level, fatigue, and activity type. Justification for a subset of these constructs involved manual labor activities found in the workplace but not in an academic environment. Based on their suitability to evaluate workload in the college and university course context, we focused on time load, mental effort, and psychological stress.

2.1.1. Time load, mental effort, and psychological stress

Time load can refer to both the time a task requires with respect to the time available as well as task overlap as assessed via the Subjective Workload Assessment Technique (SWAT) by Reid and Nygren (1988). Task overlap can impose additional time load through the necessity to prioritize some task over another at the expense of the performance in the task with lower priority. Task overlap is applicable to courses in the form of parallel assignments.

Mental effort, or cognitive load, can be measured through self-report items, for example, through the NASA Task Load Index (Hart & Staveland, 1988) whose subscales (e.g., effort) have been modified and routinely used in educational psychology (Brünken, Seufert, & Paas, 2010). However, mental effort is frequently assessed via physiological responses such as pupil dilation in eye-tracking (Palinko et al., 2010), EEG activity (Zarjam, Epps, & Chen, 2011). Conventionally, mental effort is measured during or immediately after a task. Hence, it may be challenging to operationalize mental effort through features from LMS or enrollment data. However, the number of satisfied prerequisites of students may represent their ability to manage more advanced courses.

Psychological stress is typically assessed through self-report (Cohen, Kamarck, & Mermelstein, 1994; Matthews et al., 1999). Psychological stress may depend on the characteristics of a course (e.g., the placement of assignment deadlines). Previous studies found the time students spent on courses regularly increased right before assignment deadlines (Ruiz-Gallardo, Castaño, Gómez-Alday, & Valdés, 2011; Breslow et al., 2013). A recent study surveying students regarding instructor behavior associated with stress found that multiple, effective strategies to mitigate stress in higher education are exam-related, for example, allowing for a makeup exam day (Meredith, Liu, & Frazier, 2021).

2.2. Learning analytics and learning management system data

The advent of big data in educational contexts has enabled new data-driven approaches to improve educational processes and decision-making at scale (Fischer et al., 2020). This has led to nascent research communities, such as education data science (McFarland, Khanna, Domingue, & Pardos, 2021), offering methodologies that enable new lenses with which we can study facets of higher education.

The learning analytics community has explored time management, cognitive load, and student emotional states with respect to learning outcomes. Uzir et al. (2020) used undergraduate trace data from a learning management system to derive time management tactics that correlated with academic performance. Similarly, Li, Baker, and Warschauer (2020) investigated to what degree self-regulation (e.g., studying in advance) can be automatically assessed through clickstream data to predict learning outcomes beyond traditional self-report measures (e.g., Motivated Strategies for Learning Questionnaire). Investigating the relationship between procrastination and learning outcomes through LMS data, Park et al. (2018) related the regularity of procrastination behavior to course performance. Relating to mental effort, Larmuseau, Cornelis, Lancieri, Desmet, and Depaepe (2020) link skin temperature to probabilistic reasoning task performance and self-reported cognitive load. Yen, Chen, Lai, and Chuang (2015) leverage log data from a learning management system and social messenger to adjust instructional strategies and cognitive load to improve learning. Meanwhile, affect detection has been used to predict college attendance from middle school interaction data collected in the ASSISTments system (Pedro, Baker, Bowers, & Heffernan, 2013) and learning outcomes in intelligent tutoring systems (Joshi et al., 2019). Similarly, there has been evidence that self-reported emotional states predict dropout in MOOCs (Dillon et al., 2016). In summary, as time management, cognitive load, and student emotional states have been found to correlate with learning outcomes, we ask whether course load is a mediating factor that might be influenced via course design decisions.

Learning management system data offer course structure and learning behavior (Conijn, Snijders, Kleingeld, & Matzat, 2016) from a wide breadth of courses. These may include clickstream measures (Cicchinelli et al., (2018)), data about assignments and quizzes (Zacharis, 2015), or discussion forums (Mwalumbwe & Mtebe, 2017). Early work leveraging these data included an early warning detection system to predict adverse course grades (Macfadyen & Dawson, 2010). Recent studies also revealed that engineering features from LMS forum posts are not trivial. Top-level, rather than reply comments in for-credit online courses predict grades, one study found (Almeda et al., 2018). In addition, a study showed that students contributing to *content-related* threads earned slightly higher grades on average (Wise & Cui, 2018).

2.3. Instructional design and course design

Instructional design may be defined as a systematic and knowledge-driven approach to planning, managing, and evaluating instruction to improve learning and retention (Baturay, 2008) and has been a concern of university-specific centers for teaching and learning (Lieberman, 2005). In educational psychology, the guiding principle of instructional design is to design learning materials that consider how the human cognitive architecture can process information best to foster long-term retention (Sweller, 2021). For example, design principles of multimedia learning environments take into account cognitive load (Mayer, 2005; Brünken et al., 2010). However, instructional design requires practitioners to effectively translate instructional design principles to in-person and online courses.

The advent of online courses such as massive open online courses (MOOCs) has put new importance on instructional design, or course design decisions, more specifically (Matcha et al., 2020). We may study course design features associated with successful course completion or other outcomes or measures through learning analytics. These features may pertain to broader considerations such as the development of task complexity within the course or detailed features like unlocking a problem set only by watching a video until the end (Shukor & Abdullah, 2019). The evaluation of course design in MOOCs is particularly relevant given the high dropout rates observed in MOOCs (Kloft, Stiehler, Zheng, & Pinkwart, 2014; Goopio & Cheung, 2021) which have been attributed to learners' perceptions of the course and their time management (Eriksson, Adawi, & Stöhr, 2017).

Methods applied to MOOC content can be relevant to similar educational settings. Davis, Seaton, Hauff, and Houben (2018) cluster MOOCs based on learning design sequences (e.g., transitioning from video to HTML) and link the resulting clusters to learning outcomes. In our study, we similarly aggregate instructional behavior within Canvas LMS courses, linking it to student workload perceptions in higher education. Notably, course design decisions in online courses (e.g., the number of assignments, frequency of assessment, and degree of scaffolding in lectures) can be idiosyncratic. Work by Margaryan, Bianco, and Littlejohn (2015) found that, in a sample of 76 MOOCs, the majority were neglecting effective instructional design principles while being well organized in terms of the presentation of learning material. We put forward that the systematic study of student workload, particularly in remote instruction, may advance educational effectiveness by considering course design aspects associated with workload and learners' individual ability to deal with workload effectively.

PHYSICS C21: Physics and Music

Course Subject: Physics

Course Catalog Description: What can we learn about the nature of reality and the ways that we humans have invented to discover how the world works? An exploration of these questions through the physical principles encountered in the study of music. The applicable laws of mechanics, fundamentals of sound, harmonic content, principles of sound production in musical instruments, musical scales. Numerous illustrative lecture demonstrations will be given. Only the basics of high school algebra and geometry will be used.

1) About how many hours per week did you spend on this course, on average, including instruction time?

- ○ 0-5 hours per week
- ○ 6-10 hours per week
- ○ 11-15 hours per week
- ○ 16-20 hours per week
- ○ 21-25 hours per week
- ○ 26+ hours per week

2) How often did assignments (e.g., homework, projects and study for exams) for this course overlap with one another?

- ○ Nearly never
- ○ Seldom
- ○ Sometimes
- ○ Frequently
- ○ Nearly always

3) How manageable were the hours required per week and assignment overlaps for this course?

- ○ Nearly always manageable
- ○ Mostly manageable
- ○ Sometimes manageable
- ○ Mostly unmanageable
- ○ Nearly always unmanageable

4) How much concentration and attention did assignments and studying for this course require?

- ○ A very low amount
- ○ A low amount
- ○ A moderate amount
- ○ A high amount
- ○ A very high amount

5) How manageable was the mental effort required to learn course material or complete course assignments for this course?

- ○ Nearly always manageable
- ○ Mostly manageable
- ○ Sometimes manageable
- ○ Mostly unmanageable
- ○ Nearly always unmanageable

6) How confused, frustrated, or anxious were you while learning course material or completing course assignments?

- ○ Nearly never
- ○ Seldom
- ○ Sometimes
- ○ Frequently
- ○ Nearly always

7) How manageable was the confusion, frustration, or anxiety experienced while learning course material or completing course assignments?

- ○ Nearly always manageable
- ○ Mostly manageable
- ○ Sometimes manageable
- ○ Mostly unmanageable
- ○ Nearly always unmanageable

Fig. 1. Course load survey questions for each course a respondent took in Spring 2021.

3. Methods

The aim of this study is threefold. First, we designed a survey to collect students' perceptions of course load on the dimensions of time load, mental effort, and psychological stress for their most recent term. Second, we investigated which features from the learning management system and enrollment data are associated with the different types of course load (i.e., time load, mental effort, and psychological stress) surveyed from students. Our data set of remote instruction in the Spring 2021 semester at UC Berkeley, a large public liberal arts university, offers a lens into which elements of *course design and delivery* relate to the perceived workload of students that is unprecedented in its coverage of courses, as all courses were mandated to be online during this semester, whereas only a subset may register considerable LMS activity in semesters prior to the pandemic. Third, we conducted an inferential analysis to investigate which *classes of features* (i.e., LMS data, enrollment data, personalized survey features) explain what level of variance in course load. Specifically, we considered course-level LMS features compared to individualized features constructed from historic enrollment and grade data (e.g., prior GPA of students). The following sections will elaborate on our data sources and analysis methods.

3.1. Overview on source data sets

Our analysis data set comprises three different data sources. These include (1) institutional learning management system (LMS) data for the Spring 2021 semester, (2) institutional enrollment data, and (3) survey data of individual student course pairs comprising course load ratings.

Learning management system (LMS) data. We engineered features from Spring 2021 LMS data collected throughout the semester (January 18th to May 13th, 2021) for all 332 courses rated in our survey. We engineered all features from a larger database based on the Canvas suite, including data on 3,724 courses. Furthermore, the database entailed 61,723 assignments, 3,992,638 submissions to assignments, 423,380 submission comments, and 225,732 forum posts. Further intel on our LMS data preprocessing, analytic code, and synthetic data availability can be found in Section 3.3.2.

Enrollment data. We sourced institutional enrollment data collected between the Fall 2011 and Spring 2021 semester to create course-level (e.g., number of credit hours, historic course GPA) and student-level variables (e.g., prior GPA, number of satisfied prerequisites) for course load inference. The total sample size of our enrollment data set amounted to 193,980, with the unit of analysis being student course enrollment pairs across individual semesters.

Survey data. Our survey data was collected between June 2nd and June 15th, 2021, and included 596 course load ratings from 128 students. We asked students to rate the courses they enrolled in the past semester on time load, mental effort, and psychological stress. The following section details the processing of our survey design and data collection.

3.2. Collecting student perceptions of course load

We first designed a survey instrument to collect student perceptions of different course load types in each of their courses taken in Spring 2021.

3.2.1. Survey design

Reid and Nygren (1988) designed a survey instrument to measure time load, mental effort load, and psychological stress load in workplace settings. They named their instrument the Subjective Workload Assessment Technique (SWAT). We adapted questions from this instrument to be appropriate for students in educational settings with the rationale for our adaptations detailed in A. Our survey questions consisted of two questions for each of the three load constructs, one question on the magnitude of the construct perceived in the course and the second on the manageability of that construct. These sets of questions were designed to be answered per course. At the end of the survey, participants were asked to rate how important each of the constructs was for them, in general, when making course selection decisions.

Our survey (Fig. 1) was designed for each student participant to answer the set of load questions for each course they were enrolled in for the Spring 2021 semester. All questions were five-point Likert scale options, with the exception of the first time load question which represents an objective assessment of the hours students spent per week on the rated course. For this question, compared to the subjective assessment scales, we added an additional scale point to accommodate ratings beyond the common range of weekly hours seen in the distribution of credit hours in our database, which translated to between 1 and 36 h per week per course. Since the amount of overlap in tasks was a component of the SWAT survey's time load measure, we included it as an additional time load question; however, during our feature engineering process, we found that we could create an objective measure of assignment overlap based on release dates and deadlines seen in the learning management system data. Therefore, we decided to omit this task overlap question from our analyses.

3.2.2. Survey implementation and participant recruitment

We developed a web-based system to deploy our survey connected to campus enrollment records. The advantage of this approach compared to off-the-shelf solutions (e.g., SurveyMonkey) was that our personalized system could provide students with the collection of all courses they took in the Spring 2021 semester directly through our database. This left less room for error by omission on the part of respondents and less cross-referencing error when connecting courses rated by students to their respective data in the LMS. Students logged into the system using their standard authentication system credentials for the institution.

To recruit students to participate in our study, we posted a link with a description of the study on student social media groups pages (e.g., UC Berkeley class of '23) and sent emails to listservs of interested student survey participants.

Once the student visited the survey page, they could see the requirements and benefits. The requirements were that participants must be registered for Spring 2021 and take at least three courses. After taking the survey, each participant received a $15 gift card. We set this reward amount based on a time expectation of 1-2 min to rate each course, with four as the median number of courses enrolled per semester. Students were only asked to rate lecture courses, filtering out non-lecture courses such as seminars, independent and group study courses.

3.3. Feature engineering

3.3.1. Individual student features from enrollment data

Research question 2 asks if course load perceptions depend on individual student attributes. To create student attributes, we utilized features that prior literature deemed potentially relevant to course load perception and used campus historic enrollment data from 2011 Fall to 2021 Spring, which contains 4.2 million enrollments from 193,980 anonymized undergraduate students.

Features based on course and student prerequisites. The presence of prerequisites to enroll in a course suggests higher knowledge or skill necessary to successfully complete it. For example, students need to recall the contents of prerequisite courses to master the course material, leading to higher mental effort. Hence, we expect the number of course prerequisites to be associated with course load. We calculated the total number of course prerequisites and the number of prerequisites a student fulfilled for the course. Prerequisite satisfaction is generally not enforced at UC Berkeley. For the latter, we created two separate variables for the number of prerequisites satisfied in the Spring 2021 semester and before the Spring 2021 semester.

Features based on prior student GPA. One recent study found a small, positive association between course load and GPA (as measured on a 4.0

scale) in a US higher education institution (Huntington-Klein & Gill, 2021). This was substantiated by a related study quantifying course load as the number of courses students took each semester (Boumi & Vela, 2021). We created a feature representing the student's overall GPA and another feature representing their within-major GPA.

Features based on prior course GPA. The average GPA of a university course may be regarded as a heuristic proxy for course difficulty as students who tend to register for courses with a relatively low average grade also tend to receive a lower GPA, according to a study by Szafran (Szafran, 2001). Other studies argued that course GPA represents grade leniency to obtain positive course evaluations (Marsh & Roche, 2000) or increase enrollments (Achen & Courant, 2009). We created features for both the mean and standard deviation of grades in courses in the Spring 2021 semester. Importantly, students may apply less effort to courses taken as Pass/No-pass, potentially leading to lower course load. Therefore, for courses where non-letter grades are an option, we both included the percentage of non-letter grades and the percentage of pass or satisfactory grades among non-letter grades of courses as additional features.

3.3.2. LMS data preprocessing

We received all learning management system data for the Spring 2021 semester at UC Berkeley from campus. UC Berkeley uses the Canvas LMS. In the following, we refer to Canvas as simply "LMS", as the data we engineer features from are, in principle, also available in other learning management systems that grant institutions the ability to export data. Due to privacy concerns, we did not link student-level data across our survey and LMS data, but rather aimed at predicting our course load survey measures through course-level variables derived from LMS data. The technical details of our analysis (e.g., Python version) can be found in the supplementary GitHub repository, including reproducible analysis code and synthetic data.[2]

To ascertain the dropout status of students, we took the latest enrollment status in our records (i.e., deleted rather than active or completed). By default, we based our features on students that did not drop out throughout the semester. In addition, we created separate features for students that dropped out where appropriate (see Appendix B). Since we could not ascertain the dropout status of students that rated our courses in our survey, we considered the dropout versions of LMS features as additional course-level variables. Therefore, we did not fit separate models for students who dropped out but investigated the utility of features based on dropout students in a combined inference of course load with the other LMS features. We further describe how we used the separate set of dropout features for modeling in Section 3.5.4. In addition, we divided the Spring 2021 semester instruction time (2021–01-18 until 2021–05-13) into four equally-sized parts. We estimated the frequency with which students who initially enrolled in the course dropped out of a course each semester quarter.

We sampled all assignments with at least one student submission for filtering assignments. Not all of these assignments had due dates and unlock dates (i.e., an optional, visible date of an upcoming assignment at which students would be able to access it). Therefore, we created a separate subset for assignments that included deadlines and deadlines *and* unlock dates. This ensured that we based our LMS features on the most assignment data possible for each feature. Since we could not create features based on deadlines for 23.49% and deadlines and unlock dates for 35.54% of courses in our sample due to missing data, we created two binary control variables for modeling representing whether features could be created based on these subsets or not.

Out of all rated courses, 44.56% of courses had multiple LMS courses attached to them. This is because different course sections (e.g., discussion groups) may be assigned their own LMS course. For these courses, we always created features based on concatenated data (e.g., combining all forum posts across LMS courses into one data set). We also included a control variable representing whether data were concatenated or not for modeling. Furthermore, if courses did not use the assignment (i.e., no assignment was posted), forum (i.e., no forum post was published), or submission (i.e., students did not submit any assignment) feature at all, we included additional control variables for modeling as described in Section 3.5.3 and omitted these cases for our correlational analysis. We refer the reader to B for an overview of all control variables.

3.3.3. LMS data feature engineering

Our feature engineering process was guided by prior LMS learning analytics literature (Li et al., 2020; Zacharis, 2015; Macfadyen & Dawson, 2010), literature review, papers on discussion-based features in forums, and a deliberate engagement with LMS features. We categorized our LMS features in three categories: Features based on *assignments* (13 features), *submission comments* (4 features), and *forum posts* (14 features) made by students, teaching assistants and instructors.

Features based on assignments. We computed the standard deviation of all due dates of course assignments. As prior research found university students' workload to increase close to assignment deadlines (Ruiz-Gallardo et al., 2011), we expected assignment deadlines being evenly spread across the semester to be negatively associated with psychological stress. Similarly, we calculated the *number of parallel assignments* by adding either 24 or 72-h timeframes (or a flexible timeframe based on whether the assignment was graded) to deadlines. We also included the maximum number of graded assignments per week and *timely submission* as the mean timeframe students submitted before the deadline. Similarly, we calculated the mean difference between assignment availability and due dates, including the ratio with which deadlines and assignments were within the first two weeks of instruction. We conjectured that a large number of parallel assignments, late assignment availability, and late submissions would all be conducive to psychological stress, potentially due to increased procrastination caused by high workload (Shokeen, 2018). Finally, we computed the total number of assignments and graded assignments (per week).

Features based on forum posts. We considered the number of original (i.e., not being a reply to another post) posts students created to be associated with time load. In addition, we considered the average size of original posts of students in bytes to be associated with mental effort load since longer posts might point to more complex forum assignments. We also created variables indicating the number of forum posts written by instructors *per student* and their average reply time. Similarly, we investigated the percentage of original forum posts made by students that received a reply (by either other students or instructional staff). We conjectured that receiving timely and frequent instructor posts on the forum correlates with psychological stress as studies are pointing to the importance of student-instructor interactions in online courses for educational outcomes. Jaggars and Xu (2016) found positive associations between high-quality and frequent student-instructor interaction in online courses and course grades through survey data and qualitative analysis. Similarly, the perceived quality of online discussion forums was positively associated with students' participation and perceived learning in another study (Balaji & Chakrabarti, 2010).

Features based on submission comments. We considered the number of submission comments per student and their average size in bytes to be associated with mental effort as we expect more complex courses and assignments to yield, on average, more and longer submission comments. Similar to forum responsivity, we considered the percentage of submissions that received at least one submission comment to be negatively associated with psychological stress. Notably, we did not create separate feature versions for dropout students for this category due to a high frequency of missing values.

Other LMS features. We divided the semester into four equally-sized quarters to calculate course dropout ratios since we considered the reasons for dropping out throughout the semester to differ. While early dropout might signal selection of courses after over-enrolling in courses,

[2] https://github.com/CAHLR/credit-hours-IHE.

Table 1
Example analysis data set for five (synthetic) student responses and sample variables from each sourced data set.

IDs		LMS			Enrollment				Responses		
Student ID	Course ID	N Course Assignments	% Posts with Replies	N Prereqs Course	Course Credit Hours	Historic Course GPA	Indiv: Student GPA	Indiv: N Satisfied Prereqs	Time Load	Mental Effort	Psychological Stress
1	Physics 123A	0	0.000	2	3	3.271	3.878	2	2	3	4
1	Mathematics 1A	17	0.113	0	2	3.668	3.878	0	1	2	1
2	Music 120	32	0.017	4	3	3.214	3.473	3	3	3	3
2	Education 131	64	0.000	1	4	3.411	3.473	1	2	3	2
2	History 100	6	0.000	0	1	3.051	3.473	0	2	5	5
3	Psychology 290	22	0.500	1	4	3.767	3.689	0	1	3	4

late dropout might signal prioritizing other courses to achieve a higher grade overall (i.e., students expecting to earn a worse grade than they planned). Note that at UC Berkeley, there is a deadline for dropping courses without a fee in December, while each student is allocated a budget of dropping two courses late in the semester without paying a fee until the last day of instruction.[3]

Notably, some of these variables might fluctuate each time the course is carried out. While we might assume characteristics of assignments and the responsiveness of instructors and TAs to remain similar over time, variables based on forum posts and assignment submissions likely vary by cohort.

An overview of all features and their mathematical definitions can be found in Appendix B.

3.4. Analysis data set

The unit of analysis for our study is student course pairs, with students generally rating multiple courses regarding course load. Our analysis data set includes 596 course ratings from 128 students. Data sets with repeated observations (i.e., students rating multiple courses) are particularly suited for *linear mixed models* as we detail in Section 3.5.3. As our independent variables, we join enrollment data on the student level and course level as well as LMS data on the course level to our student course pairs.

Note that student and course features are on different levels of aggregation. We combine both variable types by creating all features for every row of our student survey response data (i.e., course student pairs). We thus joined the student-level features of each student rater with course-level features of each rated course. For illustration, we include an analysis data set preview including example variables and survey responses in Table 1.

3.5. Analysis methods

3.5.1. Confirmatory correlational analysis

This correlational analysis relates to our first research question of which course design choices contribute to students' perceptions of course load, broken out by time load, mental effort, and psychological stress. In particular, we pre-select a set of course-level LMS features representing course design decisions (e.g., placement of deadlines and number of assignments) for each course load type based on our expectations regarding which features would likely correlate with which course load type. We then tested our expected correlations of LMS course-level features by estimating 95% confidence intervals ($\alpha = 0.05$) based on Pearson correlations (using pairwise complete observations) in the R statistical software (R Core Team, 2021).

We motivate the particular hypotheses regarding which LMS features correlate with which course load type via prior literature. These include previous studies on the relationship between forum use in online courses and course grades (Jaggars & Xu, 2016) as well as previous survey research on the relationship between assignment deadlines and psychological stress (Ruiz-Gallardo et al., 2011). Our hypothesizing is further detailed in Section 3.3.3. A tabular overview of our hypotheses and a list of all course-level LMS features are in Appendix B.

3.5.2. Exploratory correlational analysis

We extend our confirmatory correlational analysis to speak to our first research question by correlating all of our LMS and enrollment features with each course load type and reporting significant correlations. In particular, we conduct this analysis to speak to the latter part of our first research question, that is, how the correlation of credit hours with students' course load perceptions compares to other LMS and enrollment features. While we have pre-selected variables for our confirmatory correlational analysis, unexpected associations might exist across LMS and enrollment variables that are distinct or common across course load types.

We hand-code the STEM status of all courses and majors based on the U.S. Department of Homeland Security (DHS) STEM Designated Degree Program List.[4] in order to explore if there are significant differences in the determinants of course load between STEM and non-STEM courses.

3.5.3. Mixed models

We employed linear mixed models to model time load, mental effort, and psychological stress. We checked model assumptions (normal distribution of target, homoscedasticity, random effects normality, residual normality, and linearity) and excluded variables with a variance inflation factor greater than 10 which is a commonly applied threshold (Midi & Bagheri, 2010).

Linear mixed models allow for the representation of multilevel data, most notably data obtained through repeated measurements. Using the package lme4 (Bates, Mächler, Bolker, & Walker, 2015) for the R statistical software (R Core Team, 2021), we conceptualized students as *random effects* since they were measured repeatedly (i.e., students rated multiple courses; $M = 4.69$, $SD = 1.22$), potentially introducing bias to estimates based on their frequency in the data.

Next to fitting random model intercepts per student, we *did not* fit random slopes since we considered the effect of, for example, the number of assignments on course load to not differ across individuals. In contrast, we considered students to have different baselines (i.e., proneness to psychological stress), represented by random intercepts. Notably, linear mixed models allowed us to estimate how much variance in course load perceptions these individual baselines explain.

Overall, these decisions resulted in the following model formula per type of course load, where $\beta_1, \ldots, \beta_j$ denote features predicting course

[3] https://lsadvising.berkeley.edu/policies/late-change-class-schedule.

[4] https://www.ice.gov/sites/default/files/documents/stem-list.pdf.

Table 2
Mean and standard deviation for all survey question responses (5-point Likert scales except a 6-point Likert scale for Q1).

Question	*M*	*SD*
Q1 Time Load	2.25	1.18
Q2 Parallel Work	2.62	1.29
Q3 Time Load Manageability	3.91	1.07
Q4 Mental Effort	3.24	1.23
Q5 Mental Effort Manageability	3.70	1.13
Q6 Psychological Stress	2.70	1.27
Q7 Psychological Stress Manageability	3.83	1.14
Q8 Time Load Importance	3.85	1.00
Q9 Mental Effort Importance	3.64	1.08
Q10 Psychological Stress Importance	3.70	1.18

load Y_i while υ_k denotes student random effects:

$$Y_i = \beta_0 + \beta_1 + \ldots + \beta_j + \upsilon_k + \epsilon_i \quad (1)$$

3.5.4. *Model comparison*

Our second research question relates to the specificity of perceptions of course load to each individual student. To investigate this research question, we ran incremental likelihood-ratio model comparisons of linear mixed models for different variables sets and time load, mental effort, psychological stress, and combined course load as dependent variables. We classified our features into (1) course-level LMS features conjectured to be associated with the different course load types as outlined in Section 3.5.1, (2) the personalized student-level survey rating of how important the student finds the different course load types when choosing courses, and (3) individualized student-level features based on enrollment records (e.g., prior student GPA, number of satisfied prerequisites). Given that the personalized importance rating might be more feasibly collected in practice (e.g., in a course recommendation system) compared to the individualized enrollment features, we treated these two variable sets separately. We used a model only featuring credit hours as our baseline.

Within our LMS feature set, we considered variables based on students that dropped out of the course to be a separate set of features. We conjectured that the signal from students that did not complete the course might be particularly useful for inferring course load. Note, that our data precluded us from linking an individual student's course load

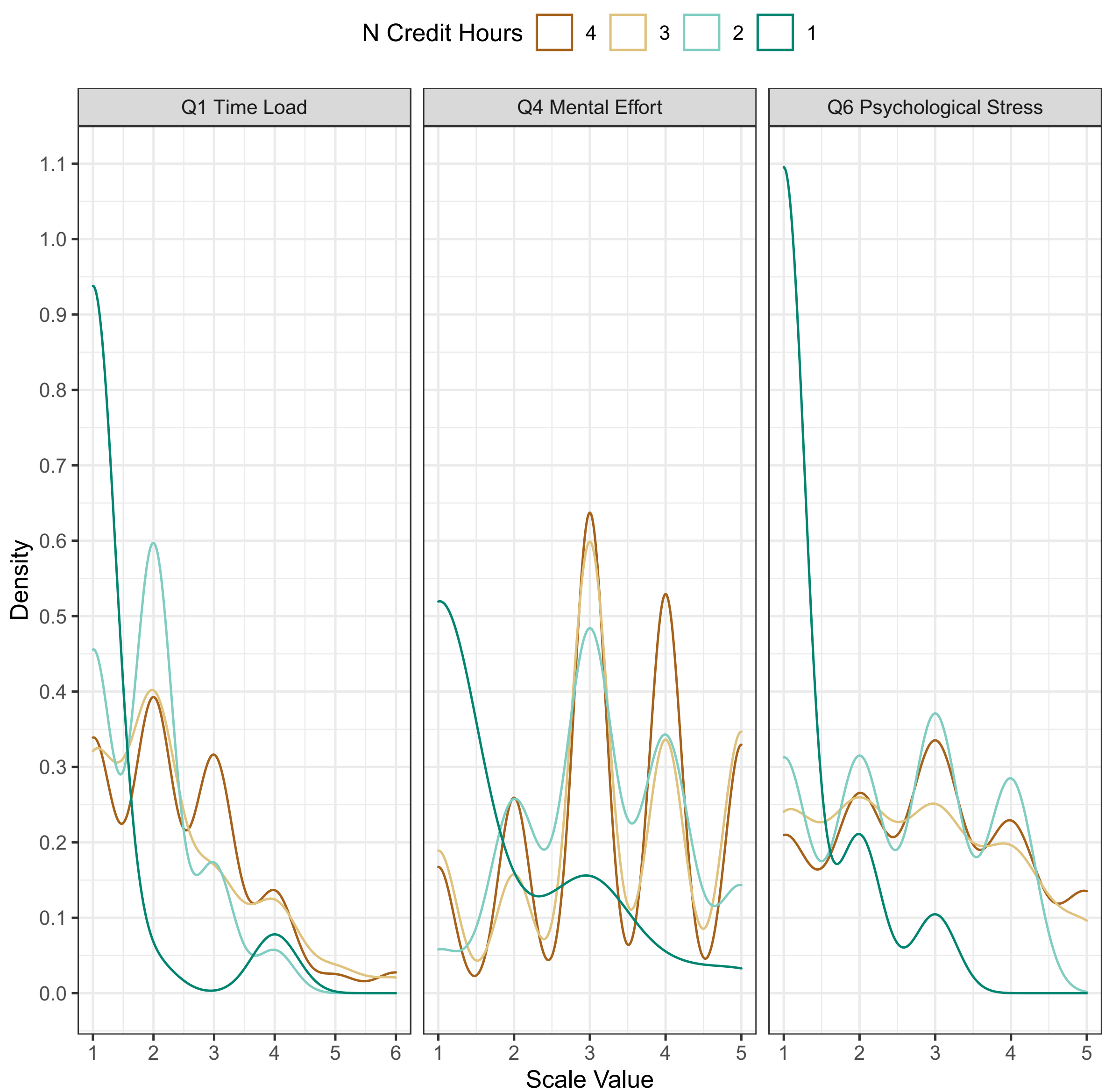

Fig. 2. Survey responses by course load type and credit hours, 13 responses with more than 4 credit hours omitted.

Table 3
Correlation tests for course-level LMS features expected to correlate with time load, significant correlations with asterisk.

Feature	r	95% CI
Forum Post Size (Bytes), Dropout Students*	−0.17	[−0.32, −0.01]
Forum Post Size (Bytes), Students	−0.12	[−0.26, 0.02]
N Course Assignments*	0.18	[0.08, 0.26]
N Graded Assignments per Week*	0.13	[0.03, 0.23]
N Graded Course Assignments*	0.17	[0.08, 0.26]
N Original Forum Posts by Students	−0.07	[−0.21, 0.07]
N Original Forum Posts by Students (Dropout)	−0.06	[−0.22, 0.10]
Original Forum Posts per Student	−0.07	[−0.21, 0.07]
Original Forum Posts per Student (Dropout)	−0.08	[−0.24, 0.08]

Table 4
Correlation tests for course-level LMS features expected to correlate with mental effort, significant correlations with asterisk.

Feature	r	95% CI
% Assignments Available in 1st 2 Weeks of Semester	−0.04	[−0.14, 0.07]
Dropout Ratio Q1	−0.04	[−0.12, 0.04]
Dropout Ratio Q2*	0.15	[0.07, 0.23]
Dropout Ratio Q3*	0.25	[0.17, 0.32]
Dropout Ratio Q4*	0.22	[0.14, 0.30]
Forum Post Size (Bytes), Dropout Students	−0.15	[−0.31, 0.01]
Forum Post Size (Bytes), Students	−0.02	[−0.16, 0.13]
Submission Comments per Student	−0.02	[−0.11, 0.07]
Submission Comments Size (Bytes)	−0.04	[−0.14, 0.06]

Table 5
Correlation tests for course-level LMS features expected to correlate with psychological stress, significant correlations with asterisk.

Feature	r	95% CI
% Assignments Available in First 2 Weeks of Sem.	0.02	[−0.08, 0.13]
% Posts by Students (Dropout) with Replies	−0.12	[−0.28, 0.04]
% Posts by Students with Replies	0.00	[−0.14, 0.14]
% Submissions that Received Comments	−0.10	[−0.19, 0.00]
Avg Diff Assignment Availability to Deadline*	−0.17	[−0.27, −0.06]
Avg Instructor/ TA Forum Reply Time	−0.12	[−0.26, 0.02]
Avg Instructor/ TA Forum Reply Time (Dropout)*	−0.21	[−0.36, −0.05]
Avg Submission Time to Deadline	−0.07	[−0.17, 0.03]
Instructor/ TA Posts per Student	0.05	[−0.09, 0.19]
N Assignments in Week with Most Due Assignments	0.06	[−0.04, 0.16]
N Parallel Assignments (1 day timeframe)*	0.13	[0.03, 0.22]
N Parallel Assignments (3 day timeframe)*	0.13	[0.03, 0.22]
N Parallel Assignments (flexible timeframe)*	0.14	[0.04, 0.23]
Spread of Assignment Due Dates	−0.03	[−0.13, 0.07]

rating from our survey to the dropout status of the student that gave the rating. Therefore, we compared a model with the standard LMS features of students that completed the course to a model *additionally featuring* LMS features based on students that dropped out. We amended the LMS model to include the additional dropout features based on the p-value of the model comparison to the standard LMS model.

In summary, our model comparison procedure proceeded as follows: (1) Compare the baseline model against the curated LMS model including, if useful, an additional set of dropout features, (2) compare the curated LMS model to a model additionally featuring the personalized importance rating of each student regarding the respective course load type when choosing courses, and (3) compare the personalized model to a model additionally including all individualized enrollment features.

4. Results

In our results, we first report descriptive statistics of our survey on course load perceptions, confirmatory correlation testing, and exploratory correlation analysis. We then report inferential model comparisons, which prompt an additional exploratory analysis of individual student differences in sensitivity to course load.

4.1. Descriptive survey results

This entire survey took, on average, 7 min. All students participated in this survey between June 2nd and June 15th, which was after their grades were released. Finally, we collected 596 course ratings from 128 students.

We show the mean and standard deviation of all survey questions in Table 2 assuming that the majority of the scales would be centered around a rating of 2.5 (with the exception of time load). Given that credit hours is conventionally used to represent course load, we expected that perceived course load would primarily covary with credit hours. To visually investigate this assumption, we exported the distribution of course load ratings by credit hours in Fig. 2. We can observe three notable trends in this figure.

First, the course load distribution for one credit hour courses is more left-skewed than for courses with two, three, and four credit hours. Second, the distributions of course load for two, three, and four credit hour courses are somewhat similar for time load and close to indistinguishable for mental effort and psychological stress. Third, the distribution of time load is left-skewed across all credit hours. Given our response categories, higher response categories may be too high and unrealistic (e.g., 25 + hours per week on one course). This descriptive analysis speaks to our first research question. Given that there seems to be no strong association between the number of credit hours and the distribution of course load (except for one credit hour courses) based on Fig. 2, credit hours might be a relatively imprecise measure for students' perceptions of course load.

The relative frequency of 1, 2, 3, 4, 5, and 6 credit hours in our survey data were 4.53%, 7.55%, 26.01%, 59.73%, 1.85%, and 0.34% (note that we omitted ratings with credit hours larger than 4 in Fig. 2). In the full database of all courses at UC Berkeley in the Spring 2021 semester ($N =$ 3,002) the relative frequency from 0 to 12 credit hours were 0.23% (0), 15.02% (1), 14.29% (2), 27.44% (3), 38.70% (4), 3.53% (5), 0.23% (6), 0.23% (7), 0.06% (8), 0.09% (9), 0.06% (10), 0.06% (12).

4.2. Confirmatory correlational analysis of course design choices to course load

Our first research question pertains to the course design choices relating to students' perceptions of course load (i.e., time load, mental effort, and psychological stress). We conjectured different sets of course-level LMS features, representing course design choices, to correlate with the three course load constructs as detailed in Section 3.3.3 and report the resulting correlations in Tables 3–5.

Observing the correlations of our LMS features with course load in Tables 3–5, we find small to moderate positive associations between the number of assignments ($r = 0.18$), graded assignments ($r = 0.17$), and graded assignments per week ($r = 0.13$) in LMS courses and perceived time load. In addition, the average size of forum posts in bytes by students that dropped out throughout the semester was significantly negatively correlated with perceived time load ($r = -0.17$). While the correlation estimate was similar for students that did not drop out (i.e., longer posts were related to lower perceived time load), that estimate was not significantly different from 0. Similarly, the number of parallel assignments was positively associated with perceived psychological stress ($r = 0.13$ for overlap counts with a 1, 3, and flexible timeframe before assignment deadlines). This means that the more deadlines overlapped during the semester (given a timeframe of 1 or 3 days before the deadline), the higher students perceived psychological stress. Notably, assignments' general spread (i.e., the standard deviation of deadlines) was not significantly correlated with psychological stress. On average, we also find that courses that made assignments available earlier before their deadline were rated lower on our psychological

Table 6
Course load type intercorrelations.

	r	95% CI
Time Load, Mental Effort	0.61	[0.56, 0.66]
Time Load, Psych. Stress	0.54	[0.48, 0.59]
Mental Effort, Psych. Stress	0.67	[0.62, 0.71]

stress scale ($r = -0.17$). Furthermore, courses with *longer* reply times in the LMS forum by instructors to posts by students that dropped out during the semester were also rated lower in psychological stress ($r = -0.21$). This means that the shorter the average course reply time was in the LMS forum, the higher students perceived psychological stress. Finally, the ratio of students that dropped out in the second ($r = 0.15$), third ($r = 0.25$), and fourth ($r = 0.22$) quarter of the Spring 2021 semester was positively associated with perceived mental effort. All remaining correlations were not significant, see Tables 3–5.

4.3. Exploratory correlational analysis and comparison to credit hours

The second part of our first research question asks how course design choices correlate with credit hours. Given the constrained subsets of variables tested for each course load type in our confirmatory analysis, we filtered significant correlations between course-level LMS and enrollment features. Before doing so, we investigated the intercorrelations of our three course load scales in Table 6. We find correlations between 0.54 for time load and psychological stress and 0.67 for mental effort and psychological stress. Heavily consulted literature (Costello & Osborne, 2005) suggests that, given these correlations, the scales may be combined into a unidimensional construct but do not necessarily have to be combined (e.g., by averaging course load ratings across scale types).

To further investigate this issue of unidimensionality, we computed correlations for all features for all course load types and a combined, averaged course load score. We then tested whether the correlation estimates of the different scales would be significantly different from one another, which was not the case and is visualized in Fig. 3.

We could not significantly distinguish the correlation estimated across course load types based on our data. Therefore, we only performed correlation-based filtering of features based on a *combined* course load measures by averaging across our three course load scales. In this step, we also investigated individualized features on the student

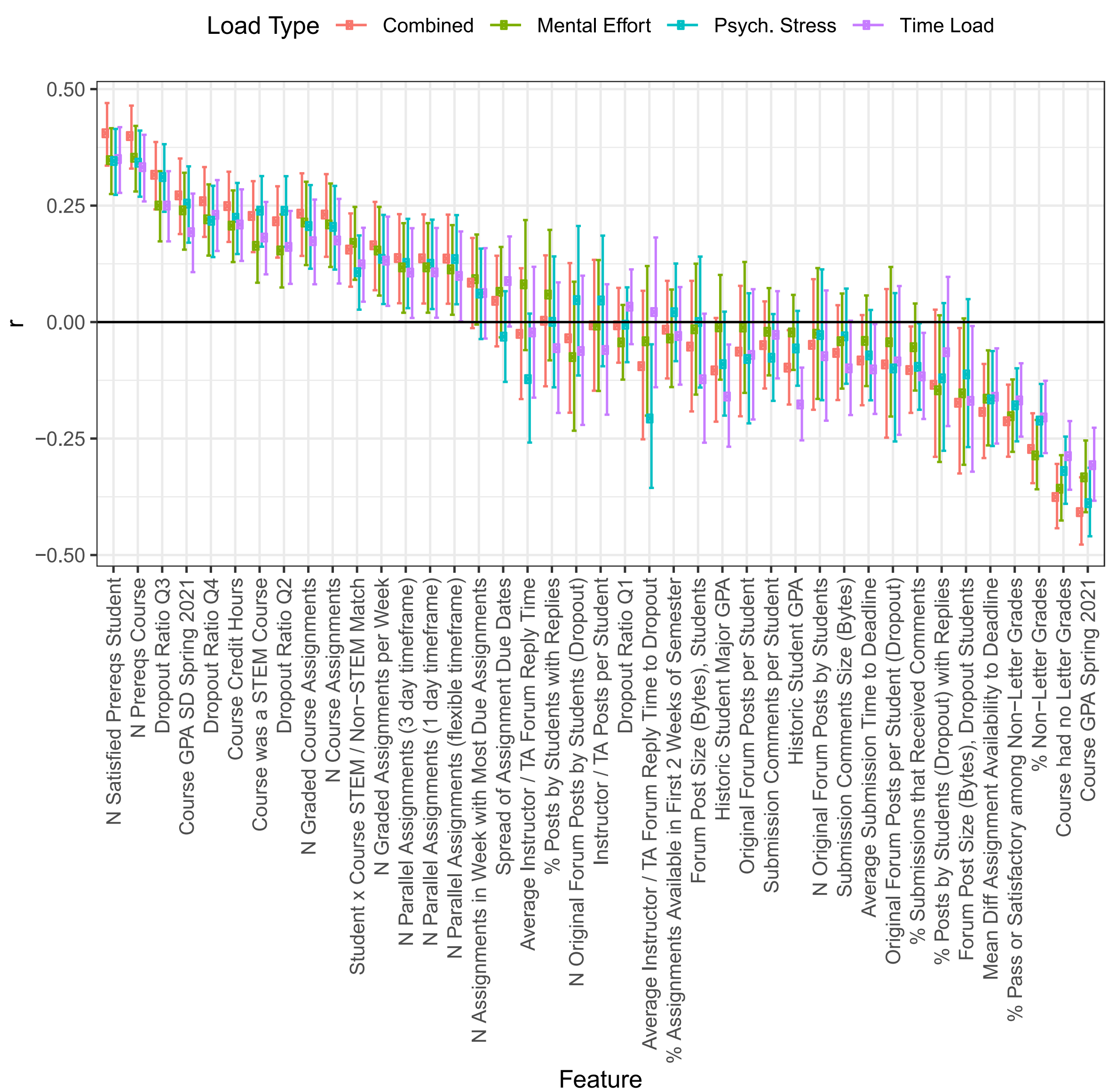

Fig. 3. Correlations of LMS and individualized enrollment features with different course load types, including 95% confidence intervals.

Table 7
Significant correlations of LMS features and individualized enrollment features (indicated via "indiv" prefix) with combined course load.

Feature	r	95% CI
Indiv: N Satisfied Prereqs of Student	0.41	[0.34, 0.47]
N Prereqs Course	0.40	[0.33, 0.46]
Dropout Ratio Q3	0.32	[0.24, 0.39]
Course GPA SD Spring 2021	0.27	[0.19, 0.35]
Dropout Ratio Q4	0.26	[0.18, 0.33]
Course Credit Hours	0.25	[0.17, 0.32]
N Graded Course Assignments	0.23	[0.14, 0.32]
N Course Assignments	0.23	[0.14, 0.32]
Course was a STEM Course	0.23	[0.15, 0.30]
Dropout Ratio Q2	0.22	[0.14, 0.29]
N Graded Assignments per Week	0.16	[0.07, 0.26]
Indiv: Student x Course STEM Match	0.16	[0.08, 0.23]
N Parallel Assignments (3 day timeframe)	0.14	[0.04, 0.23]
N Parallel Assignments (1 day timeframe)	0.14	[0.04, 0.23]
N Parallel Assignments (flexible timeframe)	0.14	[0.04, 0.23]
Indiv: Historic Student GPA	−0.10	[−0.18, −0.02]
% Submissions that Received Comments	−0.10	[−0.19, −0.01]
Forum Post Size (Bytes), Dropout Students	−0.17	[−0.32, −0.01]
Avg Diff Assignment Availability to Deadline	−0.19	[−0.29, −0.09]
% Pass or Satisfactory among Non-Letter Grades	−0.21	[−0.29, −0.13]
% Non-Letter Grades	−0.27	[−0.35, −0.20]
Course had no Letter Grades	−0.38	[−0.44, −0.30]
Course GPA Spring 2021	−0.41	[−0.48, −0.33]

Table 8
Likelihood-ratio test for combined course load comparing credit hours to LMS features, the importance item, and individualized features. Note: A direct comparison between the second and last model also yielded significance.

Model	*BIC*	$R^2_{marg./cond.}$	*deviance*	χ^2	*df*	*p*
Credit Hours	1712.87	0.06/0.12	1687.37			
LMS	1585.80	0.42/0.54	1388.12	299.24	27	<.001
LMS + Import.	1589.57	0.42/0.54	1385.51	2.61	1	.106
LMS + Indiv.	1584.19	0.45/0.57	1361.01	24.51	3	<.001

level (e.g., prior student GPA).

We present all significant correlations with combined course load in Table 7. Speaking to our first research question, we find eight features that exhibited stronger correlations with combined course load than the number of credit hours of the course ($r = 0.25$): First, the number of satisfied prerequisites of the student and the number of prerequisites of the course. The more prerequisites were required and fulfilled, the higher the perceived course load. Note, however, that in courses with more prerequisites, the number of prerequisites students fulfilled was also higher ($r = 0.74$) while 62.35% of courses in our sample had no prerequisites. Second, the course GPA in the Spring 2021 semester, whether the course was a non-letter grade only course, and the percentage of non-letter grades in the course. This indicates that the higher the average grade of the course and the more no letter grades were given, the lower was perceived course load. Third, the dropout ratio in the third and fourth quarter of the semester, meaning that students perceived courses with high late dropout ratios as being high in course load. Finally, the *standard deviation* of letter grades in the course in Spring 2021, indicating that the more grades varied in courses, the higher their perceived course load was.

4.4. *Model comparison of the inferential utility of course-level and student-level attributes for course load*

Our second research question related to the relative utility of individual student academic attributes to infer course load. In other words, we investigated how useful an individualized inference of course load beyond course-level features is. We operationalized this question by classifying our features into different types (see Section 3.5.4) and computing how much variance in course load ratings they explain. In particular, we used model comparisons to speak to the utility of load *importance* ratings and *individualization* (i.e., student-level enrollment features such as GPA) to infer course load. As this may differ across course load types, we ran our model selection procedure for all types of course load, including combined course load, individually.

The individualized model was chosen for all four model selection procedures (i.e., inferring time load, mental effort, psychological stress, and combined course load). Recall that the individualized model included, next to the LMS features and credit hours, enrollment features (e.g., prior GPA) tailored to each student rater as well as the rating of how important each student generally perceived the relevant course load type when choosing courses. In addition, the separate set of LMS features based on students that dropped out did not improve the LMS models significantly and were, therefore, not included in any LMS model. We observed differences between course load types in the marginal (i.e., variance explained by features) and conditional (i.e., variance explained by features and random intercept) R^2 of the chosen model. First, the conditional R^2 was lowest for psychological stress (33%) and highest for time load (50%). Second, marginal R^2 was highest for combined course load (45%) compared to 32% for mental effort and 27% for both time load and psychological stress. Third, the variance explained by individual student-level intercepts was highest for time load (23%) and lowest for psychological stress (6%) which can either be interpreted as different levels of within-student variance or different efficacy levels of individualized features to capture within-student variance. For reference, this metric was 12% for combined course load. In summary, speaking to our second research question, while individualization of course load prediction significantly improved model fit, the relative inferential utility compared to course-level LMS features is small, as indicated by R^2 values in Table 8.

We report the model selection results for combined course load in Table 8 with all other model comparison tables included in Appendix D. Full model tables including model coefficients are in Appendix E.

4.5. *Additional exploratory analyses*

We report on three exploratory analyses relating to our two research questions. With respect to the first research question, we investigated if any course design choices correlate differently with course load depending on if the course was STEM or non-STEM. Regarding our second research question regarding the utility of individualization for course load inference, we conducted additional analyses to investigate student-level differences in course load manageability and sensitivity.

4.5.1. *Differences in course load correlations between STEM and non-STEM courses*

We extend inquiry into our first research question regarding which course design choices contribute to perceptions of course load by examining if these contributions differ between STEM and non-STEM courses. Past work has shown that different factors do play a role when it comes to STEM and non-STEM *persistence* at the degree level. Perception of math preparedness was found to contribute to STEM persistence, whereas perceived social fit was found to be a significant factor in non-STEM preparedness (Dika & D'Amico, 2016; Sithole et al., 2017). Different correlates at the course level may lead to different implications for course design for STEM vs. non-STEM courses.

We summarize significant differences in correlation estimates of our course-level LMS and enrollment features between STEM courses and non-STEM courses for all types of course load (i.e., time load, mental effort, psychological stress, and combined course load) in Fig. 4.

We find that courses without any letter grade (i.e., taken as Pass/No-Pass) are less indicative of low course load for non-STEM courses. In addition, the number of prerequisites correlate only with course load for STEM courses. This is also true for the spread of grades in STEM courses, which was positively correlated to mental effort. Notably, for all

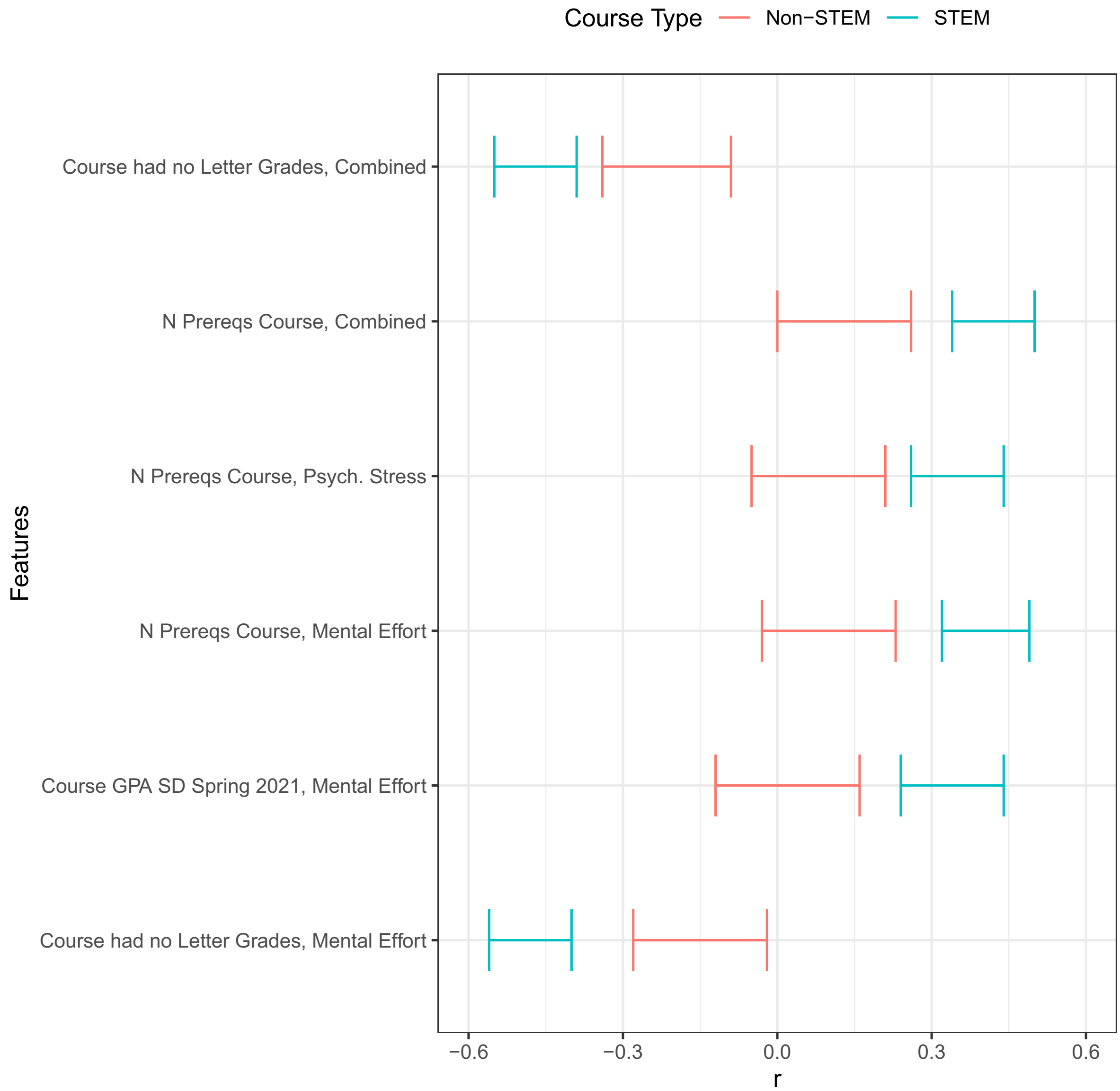

Fig. 4. Significantly different correlation estimates with course load types between STEM and non-STEM courses.

Table 9
Course load correlations with manageability responses by type.

Feature	*r* (Manageability Rating)	95% CI
Time Load	−0.56	[−0.61, −0.50]
Mental Effort	−0.61	[−0.66, −0.56]
Psych. Stress	−0.79	[−0.82, −0.76]

Table 10
Importance item correlations with student-level course load sensitivity as a function of the random student intercept taken from linear mixed models.

Feature	*r* (Random Effect)	95% CI
Time Load Importance	−0.00	[−0.18, 0.18]
Mental Effort Importance	−0.00	[−0.18, 0.18]
Psych. Stress Importance	0.05	[−0.13, 0.22]

observed differences, the direction of the association did not change between STEM and non-STEM courses but was amplified for STEM courses. Our observation that prerequisites are a significant factor only

Table 11
Importance item correlations with student-level course load sensitivity as a function of the average difference between course load and manageability responses.

Feature	*r* (Δ Load, Manage)	95% CI
Time Load Importance	−0.03	[−0.20, 0.15]
Mental Effort Importance	0.22	[0.04, 0.38]
Psych. Stress Importance	0.11	[−0.06, 0.28]

in STEM courses squares with similar observations made in related work with respect to STEM major persistence (Dika & D'Amico, 2016; Sithole et al., 2017).

4.5.2. Associations between course load and course load manageability across students

Speaking to our second research question, our main analysis indicated that the relative utility of individualization for course load inference is low. As another feature that may systematically differ across students, we investigated the correlation of our course load manageability items (i.e., how manageable students perceived each course load type) with our course load scales across all three types of course load. We find that higher load was associated with lower perceived manageability for all three course load types. This association was significantly higher for psychological stress than time load and mental effort (Table 9).

4.5.3. The role of student course load sensitivity and perceived importance for individualization

As another lens on individualization and our second research question, individual sensitivity or ability to deal with a high course load may differ. For example, prior research pointed out that resources to deal with psychological stress differ across individuals (Rabenu & Yaniv, 2017). In this exploratory analysis, we explore ways of gauging course load sensitivity at the student level.

At the end of our survey, we asked students to rate how important they find time load, mental effort, and psychological stress when choosing courses. We hypothesized post hoc that students that are particularly prone (i.e., sensitive, leading to lower perceived manageability) to particular course load types also place high importance on them during course planning. Hence, we created two metrics representing sensitivity to course load on the student level. Notice that the course load ratings were collected *after* students completed the course, while the importance measure alludes to the general course selection strategy of students. Hence, if there is a dissonance between the perception of what type of course load is important and what is actually manageable for students, this speaks to a miscalculation on the side of students and raises an opportunity for intervention and advising.

We considered two ways of quantifying student sensitivity to different types of course load. First, we extracted the estimated individual intercepts of the linear mixed models for each student for the models we chose during our model selection procedure. We may interpret these intercepts as the baseline or proclivity of each student to rate time load, mental effort, or psychological stress as high or low. Second, we calculated the average difference between the course load ratings and their perceived manageability for each student. For example, a student who tends to rate course load high and manageability low will be assigned a high sensitivity score. Conversely, a student that tends to rate course load low and manageability high will receive a low sensitivity score. We then correlated these two sensitivity measures for each student and course load type with the reported importance ratings for the respective course load type.

We find no significant correlations of importance scores at the end of the survey with student sensitivity as represented as the random intercepts in the respective linear mixed models (Table 10).

We find that students that place particular importance on the mental effort of a course when choosing it were also significantly more likely to find the mental effort of courses less manageable (Table 11).

5. Discussion

Our first research question pertained to the course design choices relating to students' perceptions of course load (i.e., time load, mental effort, and psychological stress) and how they compare to course credit hours. We found that eight features exhibited stronger associations with perceived course load than the number of credit hours of the course. These findings may have implications for pedagogy, course design, and possibly program design. Specifically, reducing the number of (parallel) assignments and increasing the time assignments are available in the LMS ahead of their deadline might reduce course load. This is in line with a recent study finding that students perceived stress and anxiety to be lower for courses making course materials available throughout the semester (Meredith et al., 2021) and adds emphasis to prior research proposing novel ways instructors can make study material available for students during university remote instruction (Majumdar, Flanagan, & Ogata, 2021). It might also be beneficial to deeply engage students in the forum as we found longer forum posts to be associated with lower perceptions of course load.

The number of prerequisites a course has most heavily correlated with perceived course load. Departments could consider spreading out the material of high prerequisite courses so as to more evenly distribute prerequisite requirements, and thus workload, across a program of study. Courses with a high percentage of students enrolled as Pass/No-Pass were associated with a notable decrease in perceived workload. This observation supports a policy enacted at UC Berkeley to allow students to satisfy more degree requirements as Pass/No-Pass during the pandemic to reduce additional anxieties caused by emergency remote instruction.[5] Notably, our exploratory comparison of STEM and non-STEM courses indicated that the associations of non-letter grades and prerequisites with course load differ across course types. Future research may investigate how these course-level enrollment features can be integrated into course advising to help students plan their semester and balance their course load more effectively.

Our second research question related to the association of individual students' academic attributes with course load and the specificity of individual student perceptions of course load. Crucially, our model comparisons indicate that LMS features explain a considerably larger amount of variance in course load than credit hours, approximately four times more for time load and six times more for combined course load. Credit hours alone explained only 6% of course load variance; however, the best single feature was under double this value, showing that a multitude of features may be necessary to capture the nuanced dimensions of course load. Student-level random intercepts explained 12% of the variance for combined course load, though our importance and individualized features explained only 3%, a small but still significant fraction of the variance beyond LMS features. The within-student variance might represent student-level calibration of how students interpreted the course load scales (e.g., what a moderate amount of psychological stress is) as captured by the random intercept in the mixed models.

Other factors may contribute to student-level differences in course load perception. For example, student-level variance might primarily reside in students' sensitivity and perceived manageability of course load. In our exploratory analysis, we found that the importance students place on mental effort when choosing courses correlated with the tendency to perceive high mental effort as less manageable. This association may further be studied through a construct from personality psychology called need for cognition, which is the individual proclivity to engage with and enjoy demanding cognitive tasks (Cacioppo & Petty, 1982; Cacioppo, Petty, & Feng Kao, 1984). Need for cognition has been found to have behavioral correlates in customer decisions and nudging effectiveness, among others (Ingendahl, Hummel, Maedche, & Vogel, 2021). Similarly, we found a specifically pronounced inverse correlation between psychological stress and manageability compared to time load and mental effort, pointing towards the importance of stress coping styles in student course load (Endler & Parker, 2008; Rogowska et al., 2021).

While students perceived time load as most important when choosing courses, this importance rating was least related to perceived manageability compared to mental effort and psychological stress. This may mean that while students are conditioned to focus on time load when choosing courses (i.e., with credit hours being the only analytic often provided about a course's load), they might be least affected by this dimension of load. This is in accordance with prior research arguing that students may deal with time load by subtracting time from non-education activities (Huntington-Klein & Gill, 2021) while conditions during the pandemic may have affected students' perceived stress (Rogowska et al., 2021) and is more difficult to deal with than time load.

[5] https://lsadvising.berkeley.edu/policies/covid-policy-modifications/fall-2020-and-spring-2021-policy-modifications-and-faq.

This speaks to the need to view student available resources to deal with course load more holistically.

Given that we primarily collected our sample during the Spring 2021 semester and emergency remote instruction during the pandemic, the question of how specific our findings are to the pandemic context arises. There are different ways the pandemic context might have influenced our results. First, the pandemic might have increased the general psychological stress of students. While a qualitative study noted increases in psychological stress in university students due to the nature of online courses (Kara, 2021), this scenario would not have altered the relative feature correlations with course load and inferences drawn from our study. In addition, we collected our data during the second year of remote instruction at UC Berkeley, which precludes effects caused by the *emergency* switch to remote instruction. Second, the pandemic might have amplified individual differences in course load experience. Consequently, our individualized model might be less useful in inferring course load outside the pandemic than this study reported. For example, Wang, Zhang, Wang, and Li (2021) found that individual differences in digital competence among university studies indirectly influenced academic burnout via its counteracting effect on cognitive load during the pandemic. Similarly, students' coping styles may moderate the impact of perceived stress during the pandemic (Rogowska et al., 2021). Finally, course-unrelated pandemic experiences, for example, isolation, distinctly affected different student populations (Maleku et al., 2021). This might have led to more variance captured by our individualized model compared to our model based on LMS features only. Future studies may further explore how individual differences might influence students' course load and how they need to be accommodated during remote instruction.

Our findings bear implications for research and practice. Credit hours appeared to be an insufficient measure of course load. At the same time, higher education institutions can now capitalize on more data and analytics that may speak to course load. Our results suggest that basing course load analytics on these data is fruitful. Therefore, the question arises if course workload analytics, such as the ones we presented, might serve students better. What context would be appropriate for such analytics to be presented to students? The course catalog is where credit hours can be found; however, the catalog traditionally contains only factual information provided by the instructor or registrar. Including workload analytics in the catalog would be an innovation, if not norm-breaking. It's an open question as to what academic unit, or more likely, what software vendors might provide these analytics and how they will be integrated. Institutions may like to consider where those analytics should be placed and how they should be framed.

Furthermore, our results speak to *course design*. Associations between course design elements of online courses and course load may have implications for course design that mitigates online and offline psychological stress. We demonstrated how institutional data allow us to investigate which types of data and course features contribute most to students' perceptions of course load (i.e., course-level characteristics such as the number of assignments or student-level factors such as GPA). The mining of these data for their ability to represent course information relevant to students is also valuable as it can help institutions decide if the retrieval of these data is worth the cost. These costs can otherwise hinder adoption (Arroway, Morgan, O'Keefe, & Yanosky, 2015) and come with additional ethical considerations (Slade & Prinsloo, 2013).

Finally, our analytical methods may contribute to *course advising* with respect to the various aims of students (e.g., completing their degree, maintaining GPA, and balancing effort with the attractiveness of courses to employers or graduate programs). This is particularly the case in the U.S., where students have considerable freedom in choosing courses to enroll in. A common issue in course selection may be choosing too many courses at once, as excessive course load can be regarded as an additional stressor for students next to social, emotional, and financial stress (Pariat, Rynjah, Joplin, & Kharjana, 2014). Hence, helping students balance their course load is a paramount priority of academic advisors. However, traditional means to quantify course load and guide students' course selection process may be insufficient to take into account students' diverse aims and resources appropriately.

5.1. Limitations

Our study has notable limitations. First, there are additional likely contributors to course load that were not part of our data set. As an example, we did not have access to the content of any lecture material nor did we have access to the chats and voice discussions in any synchronous remote sessions (i.e., Zoom). The U.S. Department of Education's definition of the credit hour implies that in-class time comprises approximately one-third of student workload (Laitinen, 2012). Therefore, future work is encouraged to engineer features from in-class interactions to infer student course load (e.g., frequency and quality of in-class engagements).

Second, students rated the course load of the courses they took after completing them and might have had difficulties accurately recalling the time load, mental effort, and psychological stress they experienced throughout the semester. Prior literature pointed out that retroactive evaluations are chiefly influenced by significant events and the endpoint of events, called peak-end rule (Geng, Chen, Lam, & Zheng, 2013). For example, ratings may have been biased by the grades students received since our data collection between June 2nd and June 15th took place a few weeks after final course grades were determined.[6] Future studies may elect to continuously assess course load throughout the semester as a potential extension of this work.

Third, additional work is required to investigate how our findings regarding course load determinants in higher education generalize to other cultural contexts. For example, the number of (satisfied) prerequisites might capture less variance in higher education contexts outside of the U.S., where there is less elective course choice and fewer ways to fulfill prerequisites. As another source of variability, prerequisites, unlike at many other universities in the U.S., are not enforced at UC Berkeley. Similarly, we did not investigate moderating effects of learner demographics on course load. Demographic variables (e.g., gender) may moderate the interest or participation in subject areas via sense of belonging (Cheryan, Plaut, Davies, & Steele, 2009; Cheryan et al., 2020). Brooks, Gardner, and Chen (2018) find the gender in instructional videos in MOOCs influences student engagement across genders which, according to our findings, may transfer to course load in the context of formal higher education courses. Furthermore, identity threat may cause psychological stress (Major & O'brien, 2005) and has been studied with respect to student–teacher race congruence in K-12 (Joshi, Doan, & Springer, 2018). Our particular data set did not include demographic data on instructors; however, future research with access to demographic data may look at how the distribution of student demographics in classes and a student's relationship to that distribution may correspond to perceived psychological stress.

Lastly, our study data sample is somewhat unique since the Spring 2021 semester was entirely comprised of remote instruction during the pandemic which may limit generalizability. This likely increased LMS data availability and some LMS features may require calibration in post-pandemic replication studies. For example, an institutional study of LMS data at UC Berkeley found that forum activity spiked while the number of assignment comments decreased during pandemic-impacted semesters (Pardos, 2022). However, findings may be more generalizable if a greater percentage of courses persist in being offered in the online modality.

[6] https://registrar.berkeley.edu/wp-content/uploads/2021/03/UCB_AcademicCalendar_2020-21_V3.pdf.

6. Conclusion

We can conclude that course credit hours alone is an inadequate attribute for students to use to anticipate the course load they will experience, as it only explained 6% of course load variance compared to a collection of LMS course-level features, which explained 36%. Course design features that correlated with higher student perceptions of course load included the number of total assignments and the number of overlapping assignment deadlines in a course. Several course design features correlated with lower course load, including courses being offered as Pass/No Pass instead of a letter grade and the average amount of time assignments were made available on the LMS ahead of their due date. The three constructs of course load (i.e., mental effort and psychological stress) were highly correlated with one another. Thus, we did not find any LMS feature correlations significantly distinct to each; however, there were differences in how manageable students reported them to be. High psychological stress was reported to be significantly less manageable than both time load and mental effort.

We find that perceptions of course load were somewhat specific to the individual, with LMS features capturing 42% of course load variance without student intercepts and 54% with student intercepts. Our three individual student-level features, combined with LMS features, provided a small but significant increase in variance captured, improving to 45% without intercepts.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Our complete feature generation and analysis code, including synthetic data generator, can be found at: https://github.com/CAHLR/credit-hours-IHE

Acknowledgements

We thank Shruthi Chockkalingam for her related work suggestions and feedback on survey design, Research Teaching and Learning at UC Berkeley for their efforts with data retrieval, and Weijie Jiang for her open source jupyter notebooks which were of help in processing our LMS data. This work was supported in part by the UC Berkeley Office of the Executive Vice Chancellor and Provost.

Appendix A. Survey questions adapted from SWAT

A.1. Time load questions

Reid and Nygren (1988) described in their Subjective Workload Assessment Technique (SWAT) instrument that time load depends on the availability of spare time and overlap of task activities. They defined the following three levels of time load:

1. Often have spare time. Interruptions or overlap among activities occur infrequently or not at all.
2. Occasionally have spare time. Interruptions or overlap among activities occur frequently.
3. Almost never have spare time. Interruptions or overlap among activities are frequent or occur all the time.

To measure the spare time and the overlap of task activities separately, we split these three levels into two questions. The first question was about how many hours per week a student spent on the course per week, on average, including instructional time. The more time students reported spending on homework, projects, exams, and other activities, the less spare time they would have each week. For this time spent question, we defined six choices of time intervals. For all other questions, we used a 5-point Likert scale. We switched from detailed level selection to a Likert scale because this has become the more common and valid way to conduct social research since the SWAT study was published 30 years ago. We also switched from three options to five for more granularity, with the hope of being able to assess the time load of students more precisely.

The second question was how often assignments (e.g., homework and projects) for a course overlap with one another, corresponding to the part of SWAT levels referring to overlapping tasks and interruptions. We removed the language of "interruptions" from the question and focused on assignment overlap. We then defined five levels of frequency from low to high as nearly never, seldom, sometimes, frequently, and nearly always to replace the three levels given by SWAT.

A.2. Mental effort questions

Mental effort was defined in SWAT as the amount of attention or concentration that is required to perform a task. Activities such as performing calculations, making decisions, remembering or storing information, and problem-solving are all examples of mental effort. SWAT defined the following three levels of mental effort:

1. Very little conscious mental effort or concentration required. Activity is almost automatic, requiring little or no attention.
2. Moderate conscious mental effort or concentration required. Complexity of activity is moderately high due to uncertainty, unpredictability, or unfamiliarity. Considerable attention required.
3. Extensive mental effort and concentration are necessary. Very complex activity requiring total attention.

The fourth question of our survey asked how much concentration and attention assignments for a course required, as derived from the SWAT levels above. The answers to this question were defined as a Very Low amount, a Low amount, a Moderate amount, a High amount, and a Very High amount.

A.3. Psychological stress questions

Psychological stress was defined by SWAT as the presence of confusion, frustration, and/or anxiety associated with task performance. SWAT defined the following three levels of psychological stress:

1. Little confusion, risk, frustration, or anxiety exists and can be easily accommodated.
2. Moderate stress due to confusion, frustration, or anxiety noticeably adds to workload. Significant compensation is required to maintain adequate performance.
3. High to very intense stress due to confusion, frustration, or anxiety. High to extreme determination and self-control required.

The sixth question of our survey was about how confused, frustrated, or anxious a student was while learning course material or completing course assignments. The only change we made was to remove the word "risk" from the language of the SWAT levels because there is no physical danger involved in normal course activities as there is in some workplace settings. We again designed five options as frequency-related terms; Nearly Never, Seldom, Sometimes, Frequently, and Nearly Always.

A.4. Course load manageability questions

After each course load type (i.e., time load, mental effort, psychological stress) question, we asked students to rate the degree to which they found the respective course load *manageable*. We included this measure as prior research on coping strategies of college students noted differences across students in coping with academic stress (Kariv & Heiman, 2005). The five options provided were: Nearly Always manageable, Mostly manageable, Sometimes manageable, Mostly unmanageable, Nearly Always unmanageable.

A.5. Importance items

At the end of our survey, we asked students to rate the general importance (irrespective of courses) of time load, mental effort, and psychological stress. We asked, "How important are each of these factors for you when choosing courses to take for a semester?" We then provided the following five options for each load type; Not Important at all, Slightly Important, Moderately Important, Important, Very Important.

Appendix B. Variable overview

B.1. LMS variables

Table B.12
Full LMS feature overview including hypothesized correlations (time load, mental effort, or psychological stress) and if separate dropout versions were created for testing the utility of dropout features in a combined inference with the default feature set. Note: Features based on submission comments did not include a dropout version due to a high frequency of missing values.

Name	Explanation	Mathematical Formula	Dropout Version?	Hypothesized Correlation
Average Size of Original Posts	Averaging the byte size of forum posts made by students throughout the semester.	$\frac{\sum postsize}{n_{posts}}$	Yes	TL
Number of Original Forum Posts	Counting the total number of posts in the LMS forum made by students throughout the semester.	$count(n_{posts})$	Yes	TL
Number of Original Forum Posts	Counting the total number of top-level posts in the LMS forum.	$count(n_{posts})$	Yes	TL
Original Forum Posts per Student	The number of top-level posts in the LMS forum divided by the number of students.	$\frac{count(n_{posts})}{count(n_{students})}$	Yes	TL
Number of TA/Instructor Posts per Student	Counting the total number of posts in the LMS by TAs/instructors and dividing it by the number of originally enrolled students.	$\frac{count(n_{posts})}{count(n_{students})}$	No	PS
Average Forum Reply Time	Average difference in minutes between student posts and TA/instructor replies.	$\frac{\sum replytime}{n_{replies}}$	Yes	PS
% Posts by Students with Replies	Ratio of forum posts by students with at least one reply from any course user.	$\frac{count(n_{postswithreplies})}{count(n_{posts})}$	Yes	PS
Average Size Of Submission Comments	Averaging the byte size of submission comments.	$\frac{\sum submissionsize}{n_{submissions}}$	No	ME
Number of Submission Comments per Student	Counting the total number of submission comments and dividing it by the number of students.	$\frac{count(n_{submcomments})}{count(n_{students})}$	No	ME
% Submissions with Comments	Dividing the number of submissions with comments by the number of submissions.	$\frac{count(n_{submw/comments})}{count(n_{submissions})}$	No	PS
Assignment Spread	Standard deviation of due dates represented in integers.	$SD(int(t_{duedatedatetime}))$	No	PS
Number of Parallel Assignments	Number of pair overlaps of assignment due dates timeframes (1 day, 3 day, flexible depending on assignment type).	$count(n_{assign.-overlap})$	No	PS
Number of Course Assignments	Counting the total number of assignments in the LMS.	$count(n_{assignments})$	No	TL
Number of Graded Course Assignments	Counting the total number of graded assignments in the LMS.	$count(n_{gradedassignments})$	No	TL
Number of Graded Course Assignments per Week	Counting the total number of graded assignments in the LMS and dividing it by the number of weeks of instruction time.	$\frac{count(n_{gradedassignments})}{n_{weeksofinstruction}}$	No	TL
Most Graded Course Assignments perWeek	Taking the number of graded assignments from the week with the most assignments.	$max(n_{asgn_{w1}}, \ldots, n_{asgn_{wi}})$	No	PS
Average Submission Time to Deadline	Average difference in minutes between student submissions and assignment deadlines.	$\frac{\sum(deadline - submission)}{n_{assignments}}$	Yes	PS
Early Assignment Availability Ratio	Ratio of assignments visible to students within the first two weeks of instruction.	$\frac{count(n_{assignments_{avail}})}{count(n_{assignments})}$	No	PS
Average Assignment Availability Time	Average difference in minutes between assignment unlock and due dates.	$\frac{\sum t_{due} - t_{available}}{n_{assignments}}$	No	PS
Dropout ratio Q1, Q2, Q3, and Q4	Ratio of enrolled students in primary section that dropped out each quarter.	$\frac{\sum t_{due} - t_{available}}{n_{assignments}}$	No	ME

B.2. Control variable and enrollment-based variables

Table B.13
Full individual and control variable overview, including whether separate versions based on students that dropped out were created.

Name	Category	Explanation	Mathematical Formula
Number of Satisfied Prerequisites	Student	Total number of satisfied prerequisites of student.	$count(n_{satisfiedprereqs})$
Student Course STEM Match	Student	Binary variable representing whether student major and course had the same STEM status (STEM/non-STEM).	$m(c,s) = \begin{cases} 1 & c_{stem} = s_{stem} \\ 0 & c_{stem} \neq s_{stem} \end{cases}$
Historic Student Major GPA	Student	Average student GPA in all past semesters in their major.	$\frac{\sum grade}{n_{courses_{major}}}$
Course Credit Hours	Course	Number of credit hours of course.	$count(n_{credithours})$
Course GPA Spring 2021	Course	Average GPA of students in course in Spring 2021.	$\frac{\sum grade}{n_{students}}$
Course GPA Standard Deviation Spring 2021	Course	Variation of GPA of students in course in Spring 2021 expressed as standard deviation.	$\sqrt{\frac{\sum (grade - GPA_{course})^2}{n_{students}}}$
Course was a STEM Course	Course	STEM status of the course.	$status(c) = \begin{cases} 1 & yes \\ 0 & no \end{cases}$
Number of Course Prerequisites	Course	Number or prerequisites of the course.	$count(n_{satisfiedprereqs})$
% Pass and Satisfactory among Non-Letter Grades	Course	Ratio of pass and satisfactory grades of all non-letter grades of course.	$\frac{count(grade_{pass/satis})}{n_{grades_{non-letter}}}$
Course had no Letter Grades	Course	Binary variable representing whether the course did not have any letter grades.	$noletter(c) = \begin{cases} 1 & yes \\ 0 & no \end{cases}$
No LMS Assignments	Control	Binary variable representing whether the course did not use the LMS assignment feature.	$f(c) = \begin{cases} 1 & count(asgn) = 0 \\ 0 & count(asgn) > 0 \end{cases}$
No LMS Assignments with Due	Control	Binary variable representing whether the course did not use the LMS assignment feature with due dates.	$f(c) = \begin{cases} 1 & count(asgn_d) = 0 \\ 0 & count(asgn_d) > 0 \end{cases}$
No LMS Assignments with Due + Unlock	Control	Binary variable representing whether the course did not use the LMS assignment feature with due *and* unlock dates.	$f(c) = \begin{cases} 1 & count(asgn_{d+u}) = 0 \\ 0 & count(asgn_{d+u}) > 0 \end{cases}$
No LMS Forum	Control	Binary variable representing whether the course did not use the LMS forum feature	$f(c) = \begin{cases} 1 & count(posts) = 0 \\ 0 & count(posts) > 0 \end{cases}$
No LMS Submissions	Control	Binary variable representing whether the course did not use the LMS submission feature.	$f(c) = \begin{cases} 1 & count(subm) = 0 \\ 0 & count(subm) > 0 \end{cases}$
Multiple LMS Courses Concatenated	Control	Binary variable representing whether the course had multiple LMS courses via sections.	$f(c) = \begin{cases} 1 & count(sections) = 0 \\ 0 & count(sections) > 0 \end{cases}$

Appendix C. Full correlation tables

Table C.14
Correlations of all features with time load.

Feature	*r*	95% CI	*t*	*df*	*p*
Indiv: N Satisfied Prereqs of Student	0.35	[0.28, 0.42]	9.10	594	<.001
N Prereqs Course	0.33	[0.26, 0.40]	8.59	594	<.001
Dropout Ratio Q3	0.25	[0.17, 0.32]	6.29	594	<.001
Dropout Ratio Q4	0.23	[0.15, 0.30]	5.76	594	<.001
Course Credit Hours	0.21	[0.13, 0.28]	5.22	594	<.001
Course GPA SD Spring 2021	0.19	[0.11, 0.28]	4.39	500	<.001
Course was a STEM Course	0.18	[0.10, 0.26]	4.49	594	<.001
N Course Assignments	0.18	[0.08, 0.26]	3.71	436	<.001
N Graded Course Assignments	0.17	[0.08, 0.26]	3.68	436	<.001
Dropout Ratio Q2	0.16	[0.08, 0.24]	3.98	594	<.001
N Graded Assignments per Week	0.13	[0.03, 0.23]	2.67	403	.008
Indiv: Student x Course STEM Match	0.12	[0.04, 0.20]	3.03	590	.003
N Parallel Assignments (1 day timeframe)	0.11	[0.01, 0.20]	2.15	403	.032
N Parallel Assignments (3 day timeframe)	0.11	[0.01, 0.20]	2.14	403	.033
N Parallel Assignments (flexible timeframe)	0.10	[0.00, 0.19]	2.00	403	.047
Spread of Assignment Due Dates	0.09	[−0.01, 0.18]	1.77	403	.077
N Assignments in Week with Most Due Assignments	0.06	[−0.04, 0.16]	1.25	403	.211
Dropout Ratio Q1	0.03	[−0.05, 0.11]	0.81	594	.421
Avg Instructor/ TA Forum Reply Time to Dropout	0.02	[−0.14, 0.18]	0.26	147	.796
Avg Instructor/ TA Forum Reply Time	−0.02	[−0.16, 0.12]	−0.31	193	.757
Submission Comments per Student	−0.03	[−0.12, 0.07]	−0.57	436	.566
% Assignments Available in First 2 Weeks of Semester	−0.03	[−0.13, 0.08]	−0.56	348	.577

(*continued on next page*)

Table C.14 (*continued*)

Feature	*r*	95% CI	*t*	*df*	*p*
% Posts by Students with Replies	−0.06	[−0.20, 0.09]	−0.78	193	.436
Instructor/ TA Posts per Student	−0.06	[−0.20, 0.08]	−0.83	193	.405
N Original Forum Posts by Students (Dropout)	−0.06	[−0.22, 0.10]	−0.76	147	.451
% Posts by Students (Dropout) with Replies	−0.06	[−0.22, 0.10]	−0.79	147	.432
Original Forum Posts per Student	−0.07	[−0.21, 0.07]	−0.98	193	.326
N Original Forum Posts by Students	−0.07	[−0.21, 0.07]	−1.02	193	.308
Original Forum Posts per Student (Dropout)	−0.08	[−0.24, 0.08]	−1.03	147	.305
Submission Comments Size (Bytes)	−0.10	[−0.20, 0.00]	−1.91	366	.057
Avg Submission Time to Deadline	−0.10	[−0.20, 0.00]	−2.05	403	.041
% Submissions that Received Comments	−0.12	[−0.21, −0.02]	−2.45	436	.015
Forum Post Size (Bytes), Students	−0.12	[−0.26, 0.02]	−1.72	193	.087
Indiv: Historic Student Major GPA	−0.16	[−0.27, −0.05]	−2.81	302	.005
Avg Diff Assignment Availability to Deadline	−0.16	[−0.26, −0.06]	−3.03	348	.003
% Pass or Satisfactory among Non-Letter Grades	−0.17	[−0.25, −0.09]	−4.12	584	<.001
Forum Post Size (Bytes), Dropout Students	−0.17	[−0.32, −0.01]	−2.08	147	.039
Indiv: Historic Student GPA	−0.18	[−0.25, −0.10]	−4.37	590	<.001
% Non-Letter Grades	−0.20	[−0.28, −0.13]	−5.06	584	<.001
Course had no Letter Grades	−0.29	[−0.36, −0.21]	−7.33	594	<.001
Course GPA Spring 2021	−0.31	[−0.38, −0.23]	−7.31	512	<.001

Table C.15
Correlations of all features with mental effort.

Feature	*r*	95% CI	*t*	*df*	*p*
N Prereqs Course	0.35	[0.28, 0.42]	9.18	594	<.001
Indiv: N Satisfied Prereqs of Student	0.35	[0.27, 0.42]	9.03	594	<.001
Dropout Ratio Q3	0.25	[0.17, 0.32]	6.29	594	<.001
Course GPA SD Spring 2021	0.24	[0.16, 0.32]	5.52	500	<.001
Dropout Ratio Q4	0.22	[0.14, 0.30]	5.51	594	<.001
N Graded Course Assignments	0.21	[0.12, 0.30]	4.57	436	<.001
N Course Assignments	0.21	[0.12, 0.30]	4.48	436	<.001
Course Credit Hours	0.21	[0.13, 0.28]	5.16	594	<.001
Indiv: Student x Course STEM Match	0.17	[0.09, 0.25]	4.19	590	<.001
Course was a STEM Course	0.16	[0.08, 0.24]	4.04	594	<.001
Dropout Ratio Q2	0.15	[0.07, 0.23]	3.79	594	<.001
N Graded Assignments per Week	0.15	[0.06, 0.25]	3.12	403	.002
N Parallel Assignments (1 day timeframe)	0.12	[0.02, 0.21]	2.37	403	.018
N Parallel Assignments (3 day timeframe)	0.12	[0.02, 0.21]	2.37	403	.018
N Parallel Assignments (flexible timeframe)	0.11	[0.02, 0.21]	2.28	403	.023
N Assignments in Week with Most Due Assignments	0.09	[−0.01, 0.19]	1.86	403	.064
Avg Instructor/ TA Forum Reply Time	0.08	[−0.06, 0.22]	1.13	193	.261
Spread of Assignment Due Dates	0.06	[−0.03, 0.16]	1.30	403	.193
% Posts by Students with Replies	0.06	[−0.08, 0.20]	0.82	193	.413
Instructor/ TA Posts per Student	−0.01	[−0.15, 0.13]	−0.10	193	.917
Indiv: Historic Student Major GPA	−0.01	[−0.12, 0.10]	−0.20	302	.843
Original Forum Posts per Student	−0.01	[−0.15, 0.13]	−0.16	193	.870
Forum Post Size (Bytes), Students	−0.02	[−0.16, 0.13]	−0.22	193	.830
Submission Comments per Student	−0.02	[−0.11, 0.07]	−0.44	436	.663
Indiv: Historic Student GPA	−0.02	[−0.10, 0.06]	−0.54	590	.588
N Original Forum Posts by Students	−0.03	[−0.16, 0.12]	−0.35	193	.728
% Assignments Available in First 2 Weeks of Semester	−0.04	[−0.14, 0.07]	−0.66	348	.511
Avg Submission Time to Deadline	−0.04	[−0.14, 0.06]	−0.81	403	.416
Submission Comments Size (Bytes)	−0.04	[−0.14, 0.06]	−0.79	366	.431
Avg Instructor/ TA Forum Reply Time to Dropout	−0.04	[−0.20, 0.12]	−0.50	147	.617
Original Forum Posts per Student (Dropout)	−0.04	[−0.20, 0.12]	−0.53	147	.599
Dropout Ratio Q1	−0.04	[−0.12, 0.04]	−1.06	594	.287
% Submissions that Received Comments	−0.05	[−0.15, 0.04]	−1.13	436	.259
N Original Forum Posts by Students (Dropout)	−0.08	[−0.23, 0.09]	−0.91	147	.362
% Posts by Students (Dropout) with Replies	−0.15	[−0.30, 0.01]	−1.80	147	.075
Forum Post Size (Bytes), Dropout Students	−0.15	[−0.31, 0.01]	−1.88	147	.062
Avg Diff Assignment Availability to Deadline	−0.16	[−0.26, −0.06]	−3.11	348	.002
% Pass or Satisfactory among Non-Letter Grades	−0.20	[−0.28, −0.12]	−4.99	584	<.001
% Non-Letter Grades	−0.29	[−0.36, −0.21]	−7.22	584	<.001
Course GPA Spring 2021	−0.33	[−0.41, −0.25]	−8.00	512	<.001
Course had no Letter Grades	−0.36	[−0.43, −0.29]	−9.34	594	<.001

Table C.16
Correlations of all features with psychological stress.

Feature	*r*	95% CI	*t*	*df*	*p*
Indiv: N Satisfied Prereqs of Student	0.35	[0.27, 0.41]	8.98	594	<.001
N Prereqs Course	0.34	[0.27, 0.41]	8.87	594	<.001
Dropout Ratio Q3	0.31	[0.24, 0.38]	7.98	594	<.001
Course GPA SD Spring 2021	0.25	[0.17, 0.33]	5.88	500	<.001
Course was a STEM Course	0.24	[0.16, 0.31]	6.00	594	<.001
Dropout Ratio Q2	0.24	[0.16, 0.31]	5.99	594	<.001
Course Credit Hours	0.22	[0.15, 0.30]	5.59	594	<.001
Dropout Ratio Q4	0.22	[0.14, 0.29]	5.43	594	<.001
N Graded Course Assignments	0.21	[0.11, 0.29]	4.40	436	<.001
N Course Assignments	0.20	[0.11, 0.29]	4.36	436	<.001
N Graded Assignments per Week	0.14	[0.04, 0.23]	2.75	403	.006
N Parallel Assignments (flexible timeframe)	0.14	[0.04, 0.23]	2.74	403	.006
N Parallel Assignments (3 day timeframe)	0.13	[0.03, 0.22]	2.57	403	.011
N Parallel Assignments (1 day timeframe)	0.13	[0.03, 0.22]	2.53	403	.012
Indiv: Student x Course STEM Match	0.11	[0.03, 0.19]	2.61	590	.009
N Assignments in Week with Most Due Assignments	0.06	[−0.04, 0.16]	1.23	403	.219
N Original Forum Posts by Students (Dropout)	0.05	[−0.11, 0.21]	0.57	147	.569
Instructor/ TA Posts per Student	0.05	[−0.09, 0.19]	0.64	193	.521
% Assignments Available in First 2 Weeks of Semester	0.02	[−0.08, 0.13]	0.39	348	.695
% Posts by Students with Replies	0.00	[−0.14, 0.14]	0.01	193	.995
Forum Post Size (Bytes), Students	0.00	[−0.14, 0.14]	0.00	193	.999
Dropout Ratio Q1	−0.01	[−0.09, 0.07]	−0.14	594	.887
N Original Forum Posts by Students	−0.03	[−0.17, 0.11]	−0.39	193	.699
Submission Comments Size (Bytes)	−0.03	[−0.13, 0.07]	−0.59	366	.558
Spread of Assignment Due Dates	−0.03	[−0.13, 0.07]	−0.63	403	.529
Indiv: Historic Student GPA	−0.06	[−0.14, 0.02]	−1.38	590	.168
Avg Submission Time to Deadline	−0.07	[−0.17, 0.03]	−1.44	403	.150
Submission Comments per Student	−0.08	[−0.17, 0.02]	−1.60	436	.109
Original Forum Posts per Student	−0.08	[−0.22, 0.06]	−1.11	193	.270
Indiv: Historic Student Major GPA	−0.09	[−0.20, 0.02]	−1.58	302	.116
% Submissions that Received Comments	−0.10	[−0.19, 0.00]	−2.02	436	.044
Original Forum Posts per Student (Dropout)	−0.10	[−0.26, 0.06]	−1.21	147	.227
Forum Post Size (Bytes), Dropout Students	−0.11	[−0.27, 0.05]	−1.37	147	.172
% Posts by Students (Dropout) with Replies	−0.12	[−0.28, 0.04]	−1.48	147	.142
Avg Instructor/ TA Forum Reply Time	−0.12	[−0.26, 0.02]	−1.72	193	.088
Avg Diff Assignment Availability to Deadline	−0.17	[−0.27, −0.06]	−3.14	348	.002
% Pass or Satisfactory among Non-Letter Grades	−0.18	[−0.26, −0.10]	−4.39	584	<.001
Avg Instructor/ TA Forum Reply Time to Dropout	−0.21	[−0.36, −0.05]	−2.56	147	.011
% Non-Letter Grades	−0.21	[−0.29, −0.13]	−5.23	584	<.001
Course had no Letter Grades	−0.32	[−0.39, −0.25]	−8.23	594	<.001
Course GPA Spring 2021	−0.39	[−0.46, −0.31]	−9.55	512	<.001

Table C.17
Correlations of all features with combined course load.

Feature	*r*	95% CI	*t*	*df*	*p*
Indiv: N Satisfied Prereqs of Student	0.41	[0.34, 0.47]	10.80	594	<.001
N Prereqs Course	0.40	[0.33, 0.46]	10.61	594	<.001
Dropout Ratio Q3	0.32	[0.24, 0.39]	8.12	594	<.001
Course GPA SD Spring 2021	0.27	[0.19, 0.35]	6.32	500	<.001
Dropout Ratio Q4	0.26	[0.18, 0.33]	6.54	594	<.001
Course Credit Hours	0.25	[0.17, 0.32]	6.26	594	<.001
N Graded Course Assignments	0.23	[0.14, 0.32]	4.99	436	<.001
N Course Assignments	0.23	[0.14, 0.32]	4.95	436	<.001
Course was a STEM Course	0.23	[0.15, 0.30]	5.70	594	<.001
Dropout Ratio Q2	0.22	[0.14, 0.29]	5.40	594	<.001
N Graded Assignments per Week	0.16	[0.07, 0.26]	3.35	403	.001
Indiv: Student x Course STEM Match	0.16	[0.08, 0.23]	3.82	590	<.001
N Parallel Assignments (3 day timeframe)	0.14	[0.04, 0.23]	2.78	403	.006
N Parallel Assignments (1 day timeframe)	0.14	[0.04, 0.23]	2.77	403	.006
N Parallel Assignments (flexible timeframe)	0.14	[0.04, 0.23]	2.76	403	.006
N Assignments in Week with Most Due Assignments	0.08	[−0.01, 0.18]	1.70	403	.089
Spread of Assignment Due Dates	0.05	[−0.05, 0.14]	0.91	403	.361
% Posts by Students with Replies	0.00	[−0.14, 0.14]	0.04	193	.970
Dropout Ratio Q1	−0.01	[−0.09, 0.07]	−0.17	594	.865
Instructor/ TA Posts per Student	−0.01	[−0.15, 0.13]	−0.10	193	.923
% Assignments Available in First 2 Weeks of Semester	−0.02	[−0.12, 0.09]	−0.31	348	.759
Avg Instructor/ TA Forum Reply Time	−0.03	[−0.17, 0.12]	−0.35	193	.725
N Original Forum Posts by Students (Dropout)	−0.03	[−0.19, 0.13]	−0.42	147	.673
N Original Forum Posts by Students	−0.05	[−0.19, 0.09]	−0.68	193	.496

(*continued on next page*)

Table C.17 (*continued*)

Feature	r	95% CI	t	df	p
Submission Comments per Student	−0.05	[−0.14, 0.04]	−1.04	436	.301
Forum Post Size (Bytes), Students	−0.05	[−0.19, 0.09]	−0.73	193	.465
Original Forum Posts per Student	−0.06	[−0.20, 0.08]	−0.88	193	.378
Submission Comments Size (Bytes)	−0.07	[−0.17, 0.04]	−1.27	366	.205
Avg Submission Time to Deadline	−0.08	[−0.18, 0.02]	−1.66	403	.097
Original Forum Posts per Student (Dropout)	−0.09	[−0.25, 0.07]	−1.11	147	.270
Avg Instructor/ TA Forum Reply Time to Dropout	−0.09	[−0.25, 0.07]	−1.15	147	.251
Indiv: Historic Student GPA	−0.10	[−0.18, −0.02]	−2.39	590	.017
% Submissions that Received Comments	−0.10	[−0.19, −0.01]	−2.16	436	.031
Indiv: Historic Student Major GPA	−0.10	[−0.21, 0.01]	−1.81	302	.071
% Posts by Students (Dropout) with Replies	−0.13	[−0.29, 0.03]	−1.65	147	.101
Forum Post Size (Bytes), Dropout Students	−0.17	[−0.32, −0.01]	−2.13	147	.035
Avg Diff Assignment Availability to Deadline	−0.19	[−0.29, −0.09]	−3.67	348	<.001
% Pass or Satisfactory among Non-Letter Grades	−0.21	[−0.29, −0.13]	−5.27	584	<.001
% Non-Letter Grades	−0.27	[−0.35, −0.20]	−6.84	584	<.001
Course had no Letter Grades	−0.38	[−0.44, −0.30]	−9.88	594	<.001
Course GPA Spring 2021	−0.41	[−0.48, −0.33]	−10.11	512	<.001

Appendix D. Model comparison tables for time load, mental effort, and psychological stress

Table D.18
Likelihood-ratio test for time load comparing credit hours to LMS features, the importance item, and individualized features. Note: A direct comparison between the second and last model also yielded significance.

Model	BIC	$R^2_{marg./cond.}$	$deviance$	χ^2	df	p
Credit Hours	1825.65	0.05/0.25	1800.14			
LMS	1747.42	0.23/0.48	1658.15	142.00	10	<.001
LMS + Import.	1753.74	0.23/0.48	1658.09	0.05	1	.816
LMS + Indiv.	1745.88	0.27/0.50	1631.10	26.99	3	<.001

Table D.19
Likelihood-ratio test for mental effort comparing credit hours to LMS features, the importance item, and individualized features.

Model	BIC	$R^2_{marg./cond.}$	$deviance$	χ^2	df	p
Credit Hours	1893.07	0.04/0.15	1867.57			
LMS	1792.71	0.30/0.44	1671.56	196.01	15	<.001
LMS + Import.	1794.12	0.31/0.44	1666.58	4.97	1	.026
LMS + Indiv.	1803.32	0.32/0.46	1656.66	9.93	3	.019

Table D.20
Likelihood-ratio test for psychological stress comparing credit hours to LMS features, the importance item, and individualized features.

Model	BIC	$R^2_{marg./cond.}$	$deviance$	χ^2	df	p
Credit Hours	1947.82	0.05/0.05	1922.31			
LMS	1916.46	0.26/0.31	1776.18	146.13	18	<.001
LMS + Import.	1920.94	0.26/0.31	1774.28	1.90	1	.168
LMS + Indiv.	1929.30	0.27/0.33	1763.51	10.77	3	.013

Appendix E. Full model tables for all selected inferential models

Table E.21
Model table for time load with numerical features scaled to a mean of 0 and a standard deviation of 1.

Feature	β	95% CI	p
Intercept	2.09	[1.82 – 2.37]	<.001
Is Stem Course	0.27	[0.05 – 0.48]	.015
Indiv: N Satisfied Prereqs Student	0.25	[0.12 – 0.37]	<.001
No LMS Assignments	0.23	[0.02 – 0.45]	.033
Course Credit Hours	0.21	[0.12 – 0.29]	<.001
Course GPA Spring 2021	0.19	[0.09 – 0.29]	<.001
N Prereqs Course	0.15	[0.03 – 0.27]	.012
N Graded Assignments per Week	0.13	[0.05 – 0.22]	.002
Indiv: Student x Course STEM Match	0.04	[−0.16 – 0.23]	.704
Multiple LMS Courses Concatenated	−0.02	[−0.23 – 0.19]	.866
Indiv: Reported Time Load Importance	−0.02	[−0.15 – 0.10]	.718
N Original Forum Posts by Students	−0.02	[−0.15 – 0.10]	.733
Original Forum Posts per Student	−0.03	[−0.16 – 0.11]	.687
Forum Post Size (Bytes), Students	−0.04	[−0.14 – 0.05]	.370
No LMS Forum	−0.12	[−0.34 – 0.10]	.286
Indiv: Historic Student GPA	−0.20	[−0.32 – −0.08]	.001
Random Effects			
σ^2	0.76		
τ_{00anon}	0.35		
ICC	0.32		
N_{anon}	125		
Observations	588		
Marginal R^2 / Conditional R^2	0.27/	0.50	

Table E.22
Model table for mental effort with numerical features scaled to a mean of 0 and a standard deviation of 1.

Feature	β	95% CI	p
Intercept	3.11	[2.82 – 3.39]	<.001
Course GPA Spring 2021	0.27	[0.16 – 0.38]	<.001
Dropout Ratio Q4	0.18	[0.09 – 0.27]	<.001
N Prereqs Course	0.17	[0.04 – 0.30]	.008
Indiv: Satisfied Prereqs Student	0.17	[0.04 – 0.30]	.009
Course Credit Hours	0.16	[0.07 – 0.25]	<.001
No LMS Assignments with Due + Unlock	0.15	[−0.10 – 0.39]	.236
Dropout Ratio Q3	0.14	[0.05 – 0.23]	.003
Indiv: Student x Course STEM Match	0.14	[−0.07 – 0.34]	.199
Is Stem Course	0.12	[−0.11 – 0.35]	.302
Indiv: Reported Mental Effort Importance	0.11	[0.00 – 0.23]	.049
Dropout Ratio Q2	0.07	[−0.02 – 0.15]	.149
No LMS Submissions	0.05	[−0.22 – 0.32]	.721
Submission Comments per Student	0.03	[−0.07 – 0.13]	.536
Forum Post Size (Bytes), Students	0.02	[−0.08 – 0.12]	.719
Indiv: Historic Student GPA	0.00	[−0.11 – 0.11]	.983
% Assignments Available in 1st 2 Weeks	−0.03	[−0.11 – 0.06]	.548
Submission Comments Size (Bytes)	−0.06	[−0.16 – 0.04]	.233
No LMS Forum	−0.08	[−0.31 – 0.15]	.488
Multiple LMS Courses Concatenated	−0.13	[−0.36 – 0.09]	.243
Dropout Ratio Q1	−0.20	[−0.29 – −0.11]	<.001
Random Effects			
σ^2	0.86		
τ_{00anon}	0.22		
ICC	0.20		
N_{anon}	125		
Observations	588		
Marginal R^2 / Conditional R^2	0.32/	0.46	

Table E.23
Model table for psychological stress with numerical features scaled to a mean of 0 and a standard deviation of 1.

Feature	β	95% CI	p
Intercept	2.48	[2.16 – 2.80]	<.001
No LMS Assignments with Due + Unlock	0.51	[0.17 – 0.85]	.003
Is Stem Course	0.47	[0.22 – 0.72]	<.001
Course Credit Hours	0.24	[0.14 – 0.34]	<.001
Indiv: N Satisfied Prereqs Student	0.23	[0.09 – 0.37]	.002
Course GPA Spring 2021	0.22	[0.10 – 0.34]	<.001
N Parallel Assignments (1 day timeframe)	0.18	[0.06 – 0.30]	.002
N Prereqs Course	0.13	[−0.01 – 0.27]	.079
Indiv: Reported Psychological Stress Importance	0.07	[−0.04 – 0.17]	.217
% Assignments Available in 1st 2 Weeks	0.06	[−0.04 – 0.17]	.251
% Posts by Students with Replies	0.00	[−0.11 – 0.11]	.978
Instructor/ TA Posts per Student	−0.01	[−0.12 – 0.09]	.775
No LMS Submissions	−0.01	[−0.38 – 0.36]	.958
% Submissions that Received Comments	−0.02	[−0.13 – 0.09]	.762
Spread of Assignment Due Dates	−0.03	[−0.21 – 0.15]	.748
Avg Submission Time to Deadline	−0.04	[−0.14 – 0.05]	.379
Indiv: Historic Student GPA	−0.04	[−0.14 – 0.06]	.447
Avg Instructor/ TA Forum Reply Time	−0.06	[−0.17 – 0.04]	.243
Indiv: Student x Course STEM Match	−0.07	[−0.30 – 0.16]	.560
Multiple LMS Courses Concatenated	−0.09	[−0.34 – 0.17]	.503
No LMS Forum	−0.09	[−0.35 – 0.17]	.512
Avg Diff Assignment Availability to Due	−0.11	[−0.22 – −0.00]	.047
N Assignm. in Week with Most Deadlines	−0.11	[−0.25 – 0.02]	.088
No LMS Assignments with Due	−0.44	[−0.99 – 0.11]	.120
Random Effects			
σ^2	1.14		
τ_{00anon}	0.10		
ICC	0.08		
N_{anon}	125		
Observations	588		
Marginal R^2 / Conditional R^2	0.27/	0.33	

Table E.24
Model table for combined course load with numerical features scaled to a mean of 0 and a standard deviation of 1.

Feature	β	95% CI	p
Intercept	2.67	[2.42 – 2.92]	<.001
No LMS Assignments with Due + Unlock	0.30	[0.05 – 0.54]	.018
N Course Assignments	0.27	[0.16 – 0.37]	<.001
Indiv: N Satisfied Prereqs Student	0.23	[0.13 – 0.34]	<.001
Course GPA Spring 2021	0.19	[0.11 – 0.28]	<.001
Course Credit Hours	0.17	[0.10 – 0.25]	<.001
Dropout Ratio Q3	0.17	[0.10 – 0.25]	<.001
Dropout Ratio Q4	0.15	[0.08 – 0.22]	<.001
N Prereqs Course	0.15	[0.05 – 0.25]	.003
No LMS Submissions	0.15	[−0.11 – 0.42]	.253
Is Stem Course	0.13	[−0.05 – 0.32]	.162
Dropout Ratio Q2	0.07	[−0.00 – 0.14]	.055
Indiv: Reported Combined Importance	0.06	[−0.04 – 0.15]	.234
Spread of Assignment Due Dates	0.06	[−0.07 – 0.19]	.335
% Submissions that Received Comments	0.05	[−0.06 – 0.15]	.375
Student x Course STEM Match	0.05	[−0.12 – 0.21]	.577
% Assignments Available in 1st 2 Weeks	0.04	[−0.03 – 0.12]	.261
Forum Post Size (Bytes), Students	0.03	[−0.06 – 0.11]	.563
% Posts by Students with Replies	−0.00	[−0.08 – 0.08]	.935
Avg Instructor/ TA Forum Reply Time	−0.01	[−0.09 – 0.07]	.873
N Original Forum Posts by Students	−0.01	[−0.12 – 0.10]	.829
Avg Submission Time to Deadline	−0.04	[−0.11 – 0.03]	.285
Instructor/ TA Posts per Student	−0.05	[−0.12 – 0.02]	.170
Submission Comments per Student	−0.05	[−0.14 – 0.04]	.296
Submission Comments Size (Bytes)	−0.06	[−0.15 – 0.04]	.228
Original Forum Posts per Student	−0.06	[−0.17 – 0.06]	.335
Indiv: Historic Student GPA	−0.07	[−0.16 – 0.02]	.121
No LMS Assignments with Due	−0.08	[−0.46 – 0.30]	.680
Avg Diff Assignment Availability to Due	−0.09	[−0.16 – −0.01]	.027
Multiple LMS Courses Concatenated	−0.09	[−0.27 – 0.10]	.345
N Assignm. in Week with Most Deadlines	−0.12	[−0.22 – −0.02]	.018
Dropout Ratio Q1	−0.14	[−0.21 – −0.07]	<.001
No LMS Forum	−0.23	[−0.43 – −0.03]	.027
Random Effects			
σ^2	0.52		
τ_{00anon}	0.15		
ICC	0.22		
N_{anon}	125		
Observations	588		
Marginal R^2 / Conditional R^2	0.45/	0.57	

References

Achen, A. C., & Courant, P. N. (2009). What are grades made of? *Journal of Economic Perspectives, 23*(3), 77–92.

Ahuja, M., Chudoba, K. M., George, J. F., Kacmar, C., & McKnight, H. (2002). Overworked and isolated? Predicting the effect of work-family conflict, autonomy, and workload on organizational commitment and turnover of virtual workers. In *Proceedings of the 35th Annual Hawaii International Conference on System Sciences* (pp. 3586–3593). IEEE.

Almeda, M., et al. (2018). Comparing the factors that predict completion and grades among for-credit and open/MOOC students in online learning. *Online Learning, 22* (1), 1–18.

Arroway, P., Morgan, G., O'Keefe, M., & Yanosky, R. (2015). *Learning analytics in higher education. Technical report, Research report*. Louisville, CO: ECAR. March 2016. 2016 EDUCAUSE. CC by-nc-nd.

Balaji, M., & Chakrabarti, D. (2010). Student interactions in online discussion forum: Empirical research from'media richness theory' perspective. *Journal of Interactive Online Learning, 9*(1).

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48.

Baturay, M. H. (2008). Characteristics of basic instructional design models. *Ekev Academic Review, 12*(34), 471–482.

Boumi, S., & Vela, A. E. (2021). Quantifying the impact of students' semester course load on their academic performance. In *2021 ASEE Virtual Annual Conference Content Access*.

Bound, J., Lovenheim, M., & Turner, S. (2007). Understanding the decrease in college completion rates and the increased time to the baccalaureate degree. *Population Studies Center Research Report, 7*, 626.

Bowling, N. A., Alarcon, G. M., Bragg, C. B., & Hartman, M. J. (2015). A meta-analytic examination of the potential correlates and consequences of workload. *Work & Stress, 29*(2), 95–113.

Bowling, N. A., & Kirkendall, C. (2012). Workload: A review of causes, consequences, and potential interventions. *Contemporary Occupational Health Psychology: Global Perspectives on Research and Practice, 2*, 221–238.

Breslow, L., Pritchard, D. E., DeBoer, J., Stump, G. S., Ho, A. D., & Seaton, D. T. (2013). Studying learning in the worldwide classroom research into edx's first mooc. *Research & Practice in Assessment, 8*, 13–25.

Brooks, C., Gardner, J., & Chen, K. (2018). *How gender cues in educational video impact participation and retention*. International Society of the Learning Sciences, Inc. [ISLS].

Brünken, R., Seufert, T., & Paas, F. (2010). *Measuring Cognitive Load*. Cambridge University Press. pp. 181–202.

Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology, 42*(1), 116.

Cacioppo, J. T., Petty, R. E., & Feng Kao, C. (1984). The efficient assessment of need for cognition. *Journal of Personality Assessment, 48*(3), 306–307.

Cheryan, S., Lombard, E. J., Hudson, L., Louis, K., Plaut, V. C., & Murphy, M. C. (2020). Double isolation: Identity expression threat predicts greater gender disparities in computer science. *Self and Identity, 19*(4), 412–434.

Cheryan, S., Plaut, V. C., Davies, P. G., & Steele, C. M. (2009). Ambient belonging: How stereotypical cues impact gender participation in computer science. *Journal of Personality and Social Psychology, 97*(6), 1045.

Cicchinelli, A., Veas, E., Pardo, A., Pammer-Schindler, V., Fessl, A., Barreiros, C., et al. (2018). Finding traces of self-regulated learning in activity streams. In *Proceedings of the Eighth International Conference on Learning Analytics and Knowledge* (pp. 191–200).

Chockkalingam, S., Yu, R., & Pardos, Z. A. (2021). Which one's more work? Predicting effective credit hours between courses. *Proceedings of the Eleventh International Conference on Learning Analytics and Knowledge*, 599–605.

Cohen, S., Kamarck, T., Mermelstein, R., et al. (1994). Perceived stress scale. *Measuring Stress: A Guide for Health and Social Scientists, 10*(2), 1–2.

Conijn, R., Snijders, C., Kleingeld, A., & Matzat, U. (2016). Predicting student performance from LMS data: A comparison of 17 blended courses using Moodle LMS. *IEEE Transactions on Learning Technologies, 10*(1), 17–29.

Cooke, M. L. (1910). *Academic and industrial efficiency: A report to the Carnegie Foundation for the Advancement of Teaching. Number 5-6*.

Costello, A. B., & Osborne, J. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research, and Evaluation, 10*(1), 7.

Cummings, M., & Knott, T. (2001). How many credit hours should first semester engineering freshmen take? In *2001 ELD/ASEE Annual Conference* (pp. 6–544).

Davis, D., Seaton, D., Hauff, C., & Houben, G.-J. (2018). Toward large-scale learning design: Categorizing course designs in service of supporting learning outcomes. In *Proceedings of the Fifth Annual ACM Conference on Learning at Scale* (pp. 1–10).

De Cuyper, N., & De Witte, H. (2006). Autonomy and workload among temporary workers: Their effects on job satisfaction, organizational commitment, life satisfaction, and self-rated performance. *International Journal of Stress Management, 13*(4), 441.

Dika, S. L., & D'Amico, M. M. (2016). Early experiences and integration in the persistence of first-generation college students in stem and non-stem majors. *Journal of Research in Science Teaching, 53*(3), 368–383.

Dillon, J., Bosch, N., Chetlur, M., Wanigasekara, N., Ambrose, G. A., Sengupta, B., et al. (2016). *Student emotion, co-occurrence, and dropout in a mooc context*. International Educational Data Mining Society.

Endler, N., & Parker, J. D. (2008). *Coping inventory for stressful situations*. APA PsycTests.

Eriksson, T., Adawi, T., & Stöhr, C. (2017). "Time is the bottleneck": A qualitative study exploring why learners drop out of MOOCs. *Journal of Computing in Higher Education, 29*(1), 133–146.

Fischer, C., Pardos, Z. A., Baker, R. S., Williams, J. J., Smyth, P., Yu, R., et al. (2020). Mining big data in education: Affordances and challenges. *Review of Research in Education, 44*(1), 130–160.

Geng, X., Chen, Z., Lam, W., & Zheng, Q. (2013). Hedonic evaluation over short and long retention intervals: The mechanism of the peak–end rule. *Journal of Behavioral Decision Making, 26*(3), 225–236.

Geurts, S. A., Kompier, M. A., Roxburgh, S., & Houtman, I. L. (2003). Does work–home interference mediate the relationship between workload and well-being? *Journal of Vocational Behavior, 63*(3), 532–559.

Goopio, J., & Cheung, C. (2021). The MOOC dropout phenomenon and retention strategies. *Journal of Teaching in Travel & Tourism, 21*(2), 177–197.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (task load index): Results of empirical and theoretical research. In *Advances in psychology* (Vol. 52, pp. 139–183). Elsevier.

Heffernan, J. M. (1973). The credibility of the credit hour: The history, use, and shortcomings of the credit system. *The Journal of Higher Education, 44*(1), 61–72.

Huntington-Klein, N., & Gill, A. (2021). Semester course load and student performance. *Research in Higher Education, 62*(5), 623–650.

Ilies, R., Dimotakis, N., & De Pater, I. E. (2010). Psychological and physiological reactions to high workloads: Implications for well-being. *Personnel Psychology, 63*(2), 407–436.

Ingendahl, M., Hummel, D., Maedche, A., & Vogel, T. (2021). Who can be nudged? Examining nudging effectiveness in the context of need for cognition and need for uniqueness. *Journal of Consumer Behaviour, 20*(2), 324–336.

Jaggars, S. S., & Xu, D. (2016). How do online course design features influence student performance? *Computers & Education, 95*, 270–284.

Janssen, O. (2001). Fairness perceptions as a moderator in the curvilinear relationships between job demands, and job performance and job satisfaction. *Academy of Management Journal, 44*(5), 1039–1050.

Joshi, A., Allessio, D., Magee, J., Whitehill, J., Arroyo, I., Woolf, B., et al. (2019). Affect-driven learning outcomes prediction in intelligent tutoring systems. In *2019 14th IEEE international Conference on Automatic Face & Gesture Recognition (FG 2019)* (pp. 1–5). IEEE.

Joshi, E., Doan, S., & Springer, M. G. (2018). Student-teacher race congruence: New evidence and insight from tennessee. *AERA Open, 4*(4), Article 2332858418817528.

Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The Expertise Reversal Effect. *Educational Psychologist, 38*(1), 23–31.

Kantowitz, B. H. (1987). 3. Mental workload. In *Advances in psychology* (Vol. 47, pp. 81–121). Elsevier.

Kara, N. (2021). Enablers and barriers of online learning during the COVID-19 pandemic: A case study of an online university course. *Journal of University Teaching & Learning Practice, 18*(4), 11.

Kariv, D., & Heiman, T. (2005). Task-oriented versus emotion-oriented coping strategies: The case of college students. *College Student Journal, 39*(1), 72–85.

Kloft, M., Stiehler, F., Zheng, Z., & Pinkwart, N. (2014). Predicting MOOC dropout over weeks using machine learning methods. In *Proceedings of the EMNLP 2014 workshop on analysis of large scale social interaction in MOOCs* (pp. 60–65).

Laitinen, A. (2012). *Cracking the credit hour*. New America Foundation.

Larmuseau, C., Cornelis, J., Lancieri, L., Desmet, P., & Depaepe, F. (2020). Multimodal learning analytics to investigate cognitive load during online problem solving. *British Journal of Educational Technology, 51*(5), 1548–1562.

Li, Q., Baker, R., & Warschauer, M. (2020). Using clickstream data to measure, understand, and support self-regulated learning in online courses. *The Internet and Higher Education, 45*, Article 100727.

Lieberman, D. (2005). Beyond faculty development: How centers for teaching and learning can be laboratories for learning. *New Directions for Higher Education, 131*, 87–98.

Macfadyen, L. P., & Dawson, S. (2010). Mining LMS data to develop an "early warning system" for educators: A proof of concept. *Computers & Education, 54*(2), 588–599.

Major, B., & O'brien, L. T. (2005). The social psychology of stigma. *Annual Review of Psychology, 56*(1), 393–421.

Majumdar, R., Flanagan, B., & Ogata, H. (2021). ebook technology facilitating university education during COVID-19: Japanese experience. *Canadian Journal of Learning and Technology, 47*(4).

Maleku, A., Kim, Y. K., Kirsch, J., Um, M. Y., Haran, H., Yu, M., et al. (2021). *The hidden minority: Discrimination and belonging among international students in the us during the COVID-19 pandemic*. Health & Social Care in the Community.

Margaryan, A., Bianco, M., & Littlejohn, A. (2015). Instructional quality of massive open online courses (MOOCs). *Computers & Education, 80*, 77–83.

Marsh, H. W., & Roche, L. A. (2000). Effects of grading leniency and low workload on students' evaluations of teaching: Popular myth, bias, validity, or innocent bystanders? *Journal of Educational Psychology, 92*(1), 202.

Matcha, W., Gasevic, D., Uzir, N. A., Jovanovic, J., Pardo, A., Lim, L., et al. (2020). Analytics of learning strategies: Role of course design and delivery modality. *Journal of Learning Analytics, 7*(2), 45–71.

Matthews, G., Joyner, L., Gilliland, K., Campbell, S., Falconer, S., & Huggins, J. (1999). Validation of a comprehensive stress state questionnaire: Towards a state big three. *Personality Psychology in Europe, 7*, 335–350.

Mayer, R. E. (2005). Cognitive theory of multimedia learning. *The Cambridge Handbook of Multimedia Learning, 41*, 31–48.

McFarland, D. A., Khanna, S., Domingue, B. W., & Pardos, Z. A. (2021). Education data science: Past, present, future. *AERA Open, 7*, Article 23328584211052055.

Meredith, L., Liu, Y., & Frazier, P. (2021). Student perceptions of instructor behaviors that impact stress. *Scholarship of Teaching and Learning in Psychology*.

Midi, H., & Bagheri, A. (2010). Robust multicollinearity diagnostic measure in collinear data set. In *Proceedings of the 4th international conference on applied mathematics* (pp. 138–142). World Scientific and Engineering Academy and Society (WSEAS).

Mwalumbwe, I., & Mtebe, J. S. (2017). Using learning analytics to predict students' performance in moodle learning management system: A case of Mbeya University of Science and Technology. *The Electronic Journal of Information Systems in Developing Countries, 79*(1), 1–13.

Palinko, O., Kun, A. L., Shyrokov, A., & Heeman, P. (2010). Estimating cognitive load using remote eye tracking in a driving simulator. In *Proceedings of the 2010 symposium on eye-tracking research & applications* (pp. 141–144).

Paas, F., Van Gog, T., & Sweller, J. (2010). Cognitive load theory: New conceptualizations, specifications, and integrated research perspectives. *Educational Psychology Review, 22*(2), 115–121.

Pardos, Z. A. (2022). Resilience or Loss? Learning Under Emergency Remote Instruction at Cal. *Berkeley, CA: Senate Undergraduate Council Report*. https://academic-senate.berkeley.edu/sites/default/files/student_learning_report_final-complete.pdf.

Pariat, L., Rynjah, A., Joplin, M., & Kharjana, M. (2014). Stress levels of college students: Interrelationship between stressors and coping strategies. *Journal of Humanities and Social Science, 19*(8), 40–46.

Park, J., Yu, R., Rodriguez, F., Baker, R., Smyth, P., & Warschauer, M. (2018). *Understanding student procrastination via mixture models*. International Educational Data Mining Society.

Pedro, M. O., Baker, R., Bowers, A., & Heffernan, N. (2013). Predicting college enrollment from student interaction with an intelligent tutoring system in middle school. In *Educational Data Mining 2013*.

R Core Team. (2021). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Rabenu, E., & Yaniv, E. (2017). Psychological resources and strategies to cope with stress at work. *International Journal of Psychological Research, 10*(2), 8–15.

Reid, G. B., & Nygren, T. E. (1988). The subjective workload assessment technique: A scaling procedure for measuring mental workload. In *Advances in psychology* (Vol. 52, pp. 185–218). Elsevier.

Rogowska, A. M., Kuśnierz, C., & Ochnik, D. (2021). Changes in stress, coping styles, and life satisfaction between the first and second waves of the COVID-19 pandemic: A longitudinal cross-lagged study in a sample of university students. *Journal of Clinical Medicine, 10*(17), 4025.

Ruiz-Gallardo, J.-R., Castaño, S., Gómez-Alday, J. J., & Valdés, A. (2011). Assessing student workload in problem based learning: Relationships among teaching method, student workload and achievement. A case study in natural sciences. *Teaching and Teacher Education, 27*(3), 619–627.

Sansgiry, S. S., & Sail, K. (2006). Effect of students' perceptions of course load on test anxiety. *American Journal of Pharmaceutical Education, 70*(2).

Shedd, J. M. (2003). The history of the student credit hour. *New Directions for Higher Education, 2003*(122), 5–12.

Shokeen, A. (2018). Procrastination, stress and academic achievement among the B.Ed. students. *Educational Quest, 9*(1), 125–129.

Shukor, N. A., & Abdullah, Z. (2019). Using learning analytics to improve MOOC instructional design. *International Journal of Emerging Technologies in Learning (iJET), 14*(24), 6–17.

Silva, E., White, T., & Toch, T. (2015). *The Carnegie unit: A century-old standard in a changing education landscape*. Carnegie Foundation for the Advancement of Teaching.

Sithole, A., Chiyaka, E. T., McCarthy, P., Mupinga, D. M., Bucklein, B. K., & Kibirige, J. (2017). Student attraction, persistence and retention in stem programs: Successes and continuing challenges. *Higher Education Studies, 7*(1), 46–59.

Slade, S., & Prinsloo, P. (2013). Learning analytics: Ethical issues and dilemmas. *American Behavioral Scientist, 57*(10), 1510–1529.

Smith, A. P. (2019). Student workload, wellbeing and academic attainment. In *International symposium on human mental workload: Models and applications* (pp. 35–47). Springer.

Spector, P. E., & Jex, S. M. (1998). Development of four self-report measures of job stressors and strain: Interpersonal conflict at work scale, organizational constraints scale, quantitative workload inventory, and physical symptoms inventory. *Journal of occupational health psychology, 3*(4), 356.

Sweller, J. (2021). Instructional design. In *Encyclopedia of Evolutionary Psychological Science* (pp. 4159–4163). Springer.

Szafran, R. F. (2001). The effect of academic load on success for new college students: Is lighter better? *Research in Higher Education, 42*(1), 27–50.

Uzir, N. A., Gašević, D., Jovanović, J., Matcha, W., Lim, L.-A., & Fudge, A. (2020). Analytics of time management and learning strategies for effective online learning in blended environments. In *Proceedings of the Tenth International Conference on Learning Analytics and Knowledge* (pp. 392–401).

Wang, X., Zhang, R., Wang, Z., & Li, T. (2021). How does digital competence preserve university students' psychological well-being during the pandemic? An investigation from self-determined theory. *Frontiers in Psychology, 12*, Article 652594.

Wise, A. F., & Cui, Y. (2018). Learning communities in the crowd: Characteristics of content related interactions and social relationships in MOOC discussion forums. *Computers & Education, 122*, 221–242.

Yen, C.-H., Chen, I.-C., Lai, S.-C., & Chuang, Y.-R. (2015). An analytics-based approach to managing cognitive load by using log data of learning management systems and footprints of social media. *Journal of Educational Technology & Society, 18*(4), 141–158.

Young, A. M., Wendel, P. J., Esson, J. M., & Plank, K. M. (2018). Motivational decline and recovery in higher education STEM courses. *International Journal of Science Education, 40*(9), 1016–1033.

Zacharis, N. Z. (2015). A multivariate approach to predicting student outcomes in web-enabled blended learning courses. *The Internet and Higher Education, 27*, 44–53.

Zarjam, P., Epps, J., & Chen, F. (2011). Spectral EEG features for evaluating cognitive load. In *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (pp. 3841–3844). IEEE.