



PAPER

Learning to adapt: student performance and adjustment in online open-book exams

Deborah Fetter 

University of California, Davis, USA

ABSTRACT

Open-book exams (OBEs) mimic the format of closed-book exams (CBEs) but allow the use of resources, and questions focus on reasoning/problem-solving. This quasi-experimental pilot investigates how format affected performance in an online introductory nutrition course from fall 2019 to spring 2021. Students who completed OBE Midterm #1 performed worse than the CBE ($83.1 \pm 0.2\%$ compared to $84.0 \pm 0.4\%$; $p < 0.05$) but 5% better for Midterm #2 ($85.6 \pm 0.2\%$ compared to $80.6 \pm 0.4\%$; $p < 0.05$). The overall course grade for OBEs was 3.5% higher ($p < 0.05$). The OBE format often requires adjusting expectations and alternative study techniques; thus, incorporating OBEs should be approached thoughtfully, with consideration for established learning frameworks. While this study was conducted prior to the widespread availability of generative AI tools, the findings offer insights for contemporary assessment design, emphasising the importance of considering resources needed to support OBEs and their influence on learning and academic integrity.

KEYWORDS: open-book examinations, online examinations, academic integrity, online learning.

ARTICLE HISTORY: Received 26 June 2025. Accepted 24 June 2026.

Introduction

Assessment in higher education has changed as a result of increased access to information, the emergency remote teaching era during the height of the COVID-19 pandemic, the availability of online learning options, and the adoption of generative artificial intelligence (GenAI) tools (Alier et al., 2024; Kerrigan et al., 2022; Luo, 2024; Mishra et al., 2020; Permzadian & Cho, 2025). These developments have prompted educators to reconsider traditional approaches to assessment, particularly those that emphasise memorisation and



information recall, as students can now readily access information and GenAI-powered tools. At the same time, online learning has steadily increased in popularity, with over half (54.6%) of undergraduate students in the United States enrolled in at least one online course, and nearly a quarter (23.7%) taking online courses exclusively (National Center for Education Statistics, 2026). Online education provides a flexible learning experience and attracts more diverse students such as student parents and students in full-time employment (Paechter & Maier, 2010; Palvia et al., 2018). Students also value the self-pacing format of online learning as it allows them to tailor their learning experience, such as revisiting complicated concepts according to their own schedule (Paechter & Maier, 2010).

Expansion of online learning during the COVID-19 pandemic

As a result of the COVID-19 pandemic, universities rapidly transitioned to offering courses virtually, and students were no longer able to self-select an online course (Berry & Jensen, 2023; Kerrigan et al., 2022; Mishra et al., 2020). Because online learning typically relies on a learner's ability to be self-directed, it can be difficult for all students to be successful in an online class (Kebritchi et al., 2017). There is also inconsistency in the development and implementation of online courses, such as whether course content is synchronous or asynchronous, the type of Learning Management System (LMS) used, communication policies, and degree of interactivity, which can impact student learning (Cavanaugh et al., 2022). Particularly during the pandemic, the abrupt shift to online learning meant that many instructors did not have the time or resources to implement evidence-based online learning pedagogical principles, which has contributed to the variability of student experiences in an online course (Cavanaugh et al., 2022). Furthermore, there is ongoing debate about the effectiveness of online education versus traditional in-person instruction (Fischer et al., 2020; Neuhauser, 2002; Tate & Warschauer, 2022). While multiple studies have observed poorer performance in the online version of a course (Fischer et al., 2020; Tate & Warschauer, 2022), students who opt for online classes are more likely to graduate, and community college students taking online courses are more likely to transfer to a four-year institution (Tate & Warschauer, 2022). As universities expand online course offerings, and GenAI becomes increasingly available, valid and reliable assessment design is critical to ensuring learning objectives are met, consistent with constructive alignment theory's emphasis on aligning learning outcomes, instructional activities, and assessments (Alier et al., 2024; Biggs, 2014; Luo, 2024; Martin et al., 2025; Palvia et al., 2018).



Open-book versus closed-book exams

Academic integrity continues to be a major issue, and concerns about honest assessments in online education have been raised (Alessio & Messinger, 2021; Berry & Jensen, 2023; Newton & Essex, 2023). While the higher education landscape has evolved considerably since the release of ChatGPT in 2022 and the subsequent availability of additional GenAI models (Alier et al., 2024; Luo, 2024), this study offers a useful baseline for understanding the pedagogical value of open-book versus closed-book exams and provides insights that may inform assessment design in the current GenAI era. Traditional assessment in higher education, especially in high-enrolment courses, often includes proctored, closed-book exams (CBEs), which typically require recall skills and rote memorisation (Agarwal & Roediger, 2011). One study that assessed student behaviour on CBEs versus OBEs found that, when preparing for CBEs, students focused more on memorising facts/information, reviewing only the assigned material, and procrastinating studying (Theophilides & Koutselini, 2010). Advocates for CBEs argue that these assessments are pivotal for evaluating a student's understanding and recall ability, which reflects what learners know (Zagury-Orly & Durning, 2021). Preparing for a CBE also involves more effortful retrieval from memory, contributing to a learning event known as the *testing effect* (Rowland, 2014; Zagury-Orly & Durning, 2021). This phenomenon refers to the finding that active retrieval of information from memory can improve memory of the concept (Rowland, 2014). However, in an age in which students have expansive access to information, largely due to the internet and, most recently, GenAI, there is a greater demand for critical thinking skills to differentiate between fact and misconception (Alier et al., 2024; Durning et al., 2016; Luo, 2024; Szmyd & Mitera, 2024).

There have been concerns about the adequacy of CBEs to measure understanding and real-world application of concepts (Durning et al., 2016). An alternative to CBEs is open-book exams (OBEs). These types of assessments mimic the format of a CBE, but students are allowed to use resources such as course notes and textbooks. Rather than focusing on memorisation or recall, OBEs focus on students' reasoning and problem-solving skills, thus providing the basis of an authentic assessment, which reflects how knowledge and skills are applied in real-world contexts (Berry & Jensen, 2023; Block, 2012; Permzadian & Cho, 2025; Wiggins, 1990). Although it has been postulated that OBEs lead to more passive studying and the perception that the exam will be easier, a systematic review by Durning, et al. found testing effects occurred in both CBEs and OBEs (Durning et al., 2016). This suggests that



learning outcomes may depend less on the exam format and more on the design, particularly the extent to which questions require application, analysis, and synthesis.

Assessment in a high-enrolment course

It has been well documented that large undergraduate courses present challenges for assessment design and implementation—including workload considerations and providing meaningful feedback—which can hinder the use of learner-centred assessments (Mulryan-Kyne, 2010; Xu & Harfitt, 2018). “Nutrition 10/10V: Discoveries and Concepts in Nutrition” is a three-unit, high-enrolment introductory nutrition course at the University of California, Davis that is taught in both in-person and online formats. The fully online version (Nutrition 10V) enrolls more than 700 undergraduate students each quarter, including students from other UC campuses. Because this course is taught in two modalities and enrolls a diverse population of undergraduate students, it serves as a living lab in which student data can be generated and analysed. This presents a unique opportunity to examine student learning and behaviour in a real-world instructional setting.

Course assessment structure and impact of the COVID-19 pandemic

Prior to emergency remote instruction, the assessment structure for Nutrition 10V included closed-book online proctored exams (OPEs) (two midterm examinations and one non-cumulative final exam) facilitated by ProctorU, an online proctoring service. In lieu of traditional in-person proctored exams, many universities initially adopted closed-book OPEs as a means of preserving academic integrity (Hartnett et al., 2023). However, there are significant equity concerns to consider with OPEs; they generally need access to a strong, stable internet connection, personal computer or laptop, and certain software capabilities, such as downloading the proctoring software programme before the exam and a working webcam and microphone (Milone et al., 2017). Some services also require a private room for the duration of the exam. Students also tend to be resistant to OPEs and report increased test anxiety, technical difficulties, and concerns over privacy (Hartnett et al., 2023).

Because of the challenges associated with OPEs, Nutrition 10V and other courses transitioned to online OBEs during the COVID-19 pandemic, thus reflecting the broader shifts in assessment practices in higher education (Kerrigan et al., 2022). Research conducted before the expansion of GenAI has shown that exam format can influence student performance, but findings are mixed and often dependent on course context and assessment design (Durning



et al., 2016). The transition to a different exam format in the same course provided an opportunity to evaluate its potential impact by enabling comparison of grades across both formats. As online education offerings and use of GenAI continue to increase, identifying assessment approaches that promote academic integrity and meaningful learning has become increasingly important. Thus, the purpose of this pilot study is to investigate how transitioning to an online OBE format impacts performance in the pre-GenAI era. Because the data were collected prior to the widespread use of GenAI tools, the findings provide a useful baseline for comparing student performance and perceptions of OBEs before AI-assisted learning became commonplace.

Methodology

Data was used from UC students enrolled in Nutrition 10V from fall quarter 2019 to spring quarter 2021. As this was a pilot study, and a power analysis was not calculated, to increase sample size, data were aggregated across quarters as course content and learning objectives remained consistent. Due to the study's sample size, even modest score differences may reach statistical significance; therefore, effect sizes and practical implications are emphasised in interpretation. These data included exam grades, overall course grade, and self-reported demographics. Potential confounding factors include variability in student cohorts across academic quarters, prior OBE exposure, academic preparation/study strategies, and familiarity with online learning environments. While these factors introduce potential sources of bias, the consistent instructional design across quarters provides some support for meaningful comparison of performance trends over time.

Although each iteration of the course was taught by the same instructor with the same learning objectives, there was variation in pedagogical approach. From fall 2019 to winter 2020 students took online CBEs as OPEs (two midterm exams and one final exam) using ProctorU. From spring 2020 onward, OBEs were implemented and consisted of reflective, open-ended questions based on the course's six learning objectives (Table 1). Multiple versions of a similar question were used across question banks to ensure each exam was different but tested the same content and learning objectives. The UC Davis Human Subjects Institutional Review Board approved the pilot study as exempt research.



Table 1. Sample of open-ended, written, reflective questions administered for the OBEs.

Exam	Question	Learning objective(s)
#1	(a) What are two specific risk factors that can be modified through diet and/or lifestyle changes to reduce the risk for heart disease?	(3) Apply knowledge of characteristics and functions of nutrients to understand the interactions between the body and food, particularly with diet-related chronic diseases.
	(b) Are you at risk for heart disease? Why or why not?	
	Your friend is looking for some dietary recommendations for a good protein source. Meanwhile, they are trying to lower the risk of cardiovascular disease (CVD). Give one specific food that you would recommend to your friend. In your answer, include two reasons for why you would recommend this food as a source of protein for your friend who is trying to reduce their risk for CVD.	(1) Explain the key features of nutrients and identify food sources of various nutrients. (3) Apply knowledge of characteristics and functions of nutrients to understand the interactions between the body and food, particularly with diet-related chronic diseases.
#2	What are three health consequences of elevated blood glucose in someone with diabetes? Describe one lifestyle modification that may improve insulin sensitivity.	(3) Apply knowledge of characteristics and functions of nutrients to understand the interactions between the body and food, particularly with diet-related chronic diseases.
	Your friend is considering taking a supplement. What are two (2) risks you would tell your friend to consider before deciding on whether to take a supplement? Risk 1: Risk 2:	(4) Analyse popular nutritional developments, trends, and fads to evaluate potential health impacts and inform healthful decision-making.
	(a) What is bioavailability and what are three (3) factors that alter bioavailability? (b) Choose one (1) mineral discussed in course and describe two (2) specific ways the bioavailability is altered for this specific mineral.	(2) Describe the basic functions of nutrients over the life cycle.
	(a) Describe a typical meal that you eat for dinner.	(5) Analyse personal health data to evaluate potential health impacts and inform healthful decision-making.



	(b) Based on what you have learned in this course, what are two (2) ways that you could modify this meal to shift it to a more healthful eating pattern? (c) Describe how one (1) of these modifications may reduce your risk for a specific diet-related chronic disease.	(6) Integrate understanding of nutritional concepts to enhance/advance/further personal, academic, and/or career goals.
--	--	--

Data analysis

Means and standard deviations (SDs) for each exam and overall course grade were calculated for each quarter, in addition to the aggregate data used to compare overall closed-book and open-book performance. The number of points allocated for each exam varied quarter to quarter (100 or 125 points), so exam scores were weighted evenly across quarters (out of 100 points). Distributions were examined for normality using histograms, skewness, and Kurtosis. Descriptive statistics were expressed as means and SDs for continuous variables and percentages for categorical variables. Demographics were compared between open-book and closed-book groups, as well as each quarter. Categorical variables were calculated into percentages and groups were compared using the chi-square test for homogeneity. Student's t tests were used for comparisons between aggregate closed-book and open-book exam scores and overall course grade with and without data from spring 2020. Stata 16 software (StataCorp, College Station, TX, 2019) was used for all statistical analyses, and significance was determined using $p \leq 0.05$.

Results

From fall 2019 to spring 2021, 5,585 students completed course assessments. Of these, 1,560 students completed CBEs, while 4,025 students completed OBEs. The average age was 19 years, and the majority of students identified as female and Asian/Pacific Islander (Table 2). There were significant differences in enrolment status with more first-year students taking OBEs than CBEs, and fewer transfer students and international students taking OBEs. Although there were no significant differences in gender and race/ethnicity when comparing overall CBE and OBE groups, there were significant differences between quarters in these characteristics, as well as for income, enrolment status, transfer student status, and international student status.



Table 2. Characteristics of students enrolled in Nutrition 10V (fully online course).

Characteristic		Proctored closed-book (n=1,560)	Open-book (n=4,025)	P value	Fall 2019 (n=801)	Winter 2020 (n=759)	Spring 2020 (n=1,108)	Fall 2020 (n=1,186)	Winter 2021 (n=882)	Spring 2021 (n=848)	P value
Gender, n (%)	Female	933 (60)	2,382 (59)	0.06	473 (59)	460 (61)	675 (61)	674 (57)	542 (61)	491 (58)	<0.01*
	Male	508 (32)	1,350 (34)		287 (36)	221 (29)	361 (32)	410 (34)	287 (53)	292 (34)	
	Transgender	1 (0.06)	2 (0.05)		1 (0.1)	0 (0)	1 (0.1)	1 (0.1)	0 (0)	0 (0)	
	Unreported ^a	123 (8)	291 (7)		40 (5)	78 (10)	71 (6)	101 (8)	53 (6)	65 (8)	
Race/ Ethnicity, n (%)	African American/ Black, not of Hispanic origin	25 (16)	82 (2)	0.3	11 (1)	14 (2)	27 (2)	22 (2)	21 (2)	12 (1)	0.002*
	American Indian/ Alaska native	2 (0.1)	8 (0.2)		1 (0.1)	1 (0.1)	2 (0.2)	3 (0.2)	1 (0.1)	2 (0.2)	
	Asian/Pacific Islander	835 (54)	1,979 (49)		471 (59)	364 (48)	467 (42)	654 (55)	432 (49)	430 (51)	
	White, not of Hispanic origin	210 (13)	570 (14)		87 (11)	123 (16)	190 (17)	146 (12)	124 (14)	110 (13)	
	Chicano/a/x	5 (0.3)	17 (0.4)		2 (0.2)	3 (0.4)	9 (0.8)	0 (0)	2 (0.2)	6 (0.7)	



	Latino/a/x/ Hispanic (Mexican- American, Puerto Rican, Cuban)	148 (9)	490 (12)		77 (10)	71 (9)	163 (15)	108 (9)	118 (13)	101 (12)	
	Multiple reported	159 (10)	463 (12)		83 (10)	76 (10)	135 (12)	113 (10)	112 (13)	103 (12)	
	Unreported ^a	176 (11)	416 (10)		69 (9)	107 (14)	115 (10)	140 (12)	72 (8)	84 (10)	
Income, n (%)	\$0 – \$39,999	365 (23)	903 (22)	<0.01*	202 (25)	163 (21)	303 (27)	228 (19)	190 (21)	182 (21)	<0.01*
	\$40,000 – \$79,999	257 (16)	658 (16)		135(17)	122 (16)	181 (16)	199 (17)	153 (17)	125 (15)	
	\$80,000 or more	346 (22)	1,028 (26)		177 (22)	169 (22)	269 (24)	290 (24)	233 (26)	236 (28)	
	Unreported ^a	592 (38)	1,436 (36)		287 (36)	305 (40)	355 (32)	469 (40)	306 (35)	305 (36)	
Enrolment status	UC Freshman (1 st year)	561 (36)	1,827 (45)	<0.01*	280 (35)	281 (37)	400 (36)	604 (51)	456 (52)	367 (43)	<0.01*
	UC Sophomore (2 nd year)	295 (19)	511 (13)		150 (19)	145 (19)	169 (15)	132 (11)	98 (11)	112 (13)	
	UC Junior (3 rd year)	291 (19)	622 (15)		182 (23)	109 (14)	195 (18)	182 (15)	129 (15)	116 (14)	



	UC Senior (4 th year)	231 (15)	680 (17)		116 (14)	115 (15)	234 (21)	142 (12)	132 (15)	172 (20)	
	UC 5 th or 6 th year	50 (3)	101 (2)		26 (3)	24 (3)	36 (3)	24 (2)	20 (2)	21 (2)	
	UC Graduate student	3 (0.2)	6 (0.1)		0 (0)	3 (0.4)	1 (0.1)	3 (0.2)	2 (0.2)	0 (0)	
	Not a UC Student or other	7 (0.4)	12 (0.3)		4 (0.5)	3 (0.4)	7 (0.6)	2 (0.2)	0 (0)	3 (0.4)	
	Unreported ^a	122 (8)	266 (7)		43 (5)	79 (10)	66 (6)	97 (8)	45	57 (7)	
Transfer student, n (%)	Yes	297 (19)	635 (16)	<0.01*	177 (22)	120 (16)	200 (18)	193 (16)	120 (14)	122 (14)	<0.01*
	No	1,144 (73)	3,132 (78)		584 (73)	560 (74)	843 (76)	898 (76)	718 (81)	673 (79)	
	Unreported ^a	119 (8)	258 (6)		40 (5)	79 (10)	65 (6)	95 (8)	44 (5)	53 (6)	
International student, n (%)	Yes	344 (22)	785 (20)	0.05*	215 (27)	129 (17)	131 (12)	339 (28)	162 (18)	153 (18)	<0.01*
	No	1,096 (70)	2,976 (74)		545 (68)	551 (72)	913 (82)	750 (63)	675 (76)	638 (75)	
	Unreported ^a	120 (8)	264 (6)		41 (5)	79 (10)	64 (6)	97 (8)	45 (5)	57 (7)	

^aUnreported includes those that preferred to not answer or those that left the question blank. * $p \leq 0.05$.



When comparing average scores on exams, students who completed the OBE version of Midterm #1 performed worse ($83.1 \pm 0.2\%$) than students who completed the CBE version ($84.0 \pm 0.4\%$; $p < 0.05$; Table 3). For Midterm #2, students who completed the OBE version performed 5% better than students who completed the CBE version ($85.6 \pm 0.2\%$ compared to $80.6 \pm 0.4\%$; $p < 0.05$). When comparing overall course grades, students who took the course with OBEs performed 3.5% better (Table 2). Because of the grade inflation observed in spring 2020, the analysis was conducted again without including data from this quarter and similar results were found ($81.8 \pm 0.2\%$ compared to $84.0 \pm 0.4\%$; $p < 0.01$ for Midterm #1 and $84.4 \pm 0.3\%$ compared to $80.6 \pm 0.4\%$; $p < 0.01$ for Midterm #2).

The final course grade for the first quarter of emergency remote education (spring 2020) was the highest average grade, overall (382.6 points [A+] out of 400 points). Prior to spring 2020, the average course grades were 363 points (A-) and 366 points (A) for fall 2019 and winter 2020, respectively. From fall 2020 onwards, the average course grades dropped from the high of spring 2020 but remained above the average course grade from before the pandemic. An examination of individual exam scores between each quarter showed that students performed better on the first exam for the CBE format and worse on the second exam. For the OBE format, students performed worse on the first exam, but better on the second exam.

Table 3. Comparison of mean scores on course grade and exams.

	Combined		Fall 2019	Winter 2020	Spring 2020 ¹	Fall 2020	Winter 2021	Spring 2021
	Proctored closed-book (n=1,560)	Open-book (n=4,025)	Proctored closed-book (n=801)	Proctored closed-book (n=759)	Open-book (n=1,108)	Open-book (n=1,186)	Open-book (n=882)	Open-book (n=848)
Overall grade (400 points)	364.5 (1.4)*	377.4 (0.8)*	363.0 (2.0)	366.0 (2.0)	382.6 (1.4)	376.7 (1.8)	372.9 (1.8)	376.3 (1.6)
Midterm exam #1	84.0 (0.4)*	83.1 (0.2)*	83.5 (0.5)	84.6 (0.4)	86.4 (0.4)	84.9 (0.4)	79.2 (0.4)	80.3 (0.4)
Midterm exam #2	80.6 (0.4)*	85.6 (0.2)*	79.7 (0.5)	81.6 (0.5)	88.8 (0.4)	88.9 (0.5)	81.7 (0.4)	81.0 (0.4)
Final exam (100 points)	82.9 (0.4) ^a	N/A	83.0 (0.6)	82.8 (0.6)	N/A	N/A	N/A	N/A



* $p \leq 0.05$; ¹Indicates the first quarter of emergency remote teaching due to the COVID-19 pandemic; ^a $n=811$. Exam scores weighted evenly (all out of 100).

Note: Fall 2019 to Spring 2020 exams were 100 points total; Fall 2020 to Spring 2021 exams were 125 points total.

Discussion

While the overall course grade ranged from A- to A+, exam scores for both formats averaged in the B and B- range. Students who took the course with the CBE structure performed better on the first exam than students who took the OBE. Yet, students who took the course with the OBE structure performed better on the second exam. Although the observed differences in exam scores were modest in magnitude, small shifts in performance can affect a substantial number of students, especially since the course grade is based on a point system instead of a curve. Student performance also improved across sequential OBEs, thus suggesting adaptation to the assessment format over time. This is in line with prior literature, as well as self-regulated learning (SRL) theory, which suggests that students probably needed to have an initial exposure to the OBE format to develop more suitable preparation methods for the second exam (Theophilides & Koutselini, 2010; Zimmerman, 2000). The initial lower performance on Midterm #1 may reflect a mismatch between students' prior study strategies and the cognitive demands of OBEs, while improved performance on Midterm #2 suggests adaptation through self-regulated learning processes (Theophilides & Koutselini, 2010; Zimmerman, 2000).

Furthermore, there often seems to be a misperception that OBEs are easier than CBEs, which could lead some students to forgo studying and expect to look up answers during the exam (Theophilides & Koutselini, 2010). The format of an OBE requires an adjustment of expectations and different preparation methods compared to a CBE (Berry & Jensen, 2023; Heijne-Penninga et al., 2008). This highlights the importance of providing early exposure to OBEs, such as by incorporating low-stakes, formative assessment opportunities and offering guidance on effective study strategies. Practitioners can help students understand that OBEs still require preparation and conceptual understanding, while also collaborating with instructors to ensure assessments are intentionally designed to align with intended learning outcomes.



The OBE structure requires educators to develop questions that do not directly assess recall ability, but rather problems that probe cognitive ability and critical thinking (Theophilides & Koutselini, 2010). Questions on an OBE are more aligned with real-world situations students may face in their future jobs, since this type of assessment requires synthesis and application of information instead of simple memorisation and recall (Berry & Jensen, 2023; Heijne-Penninga et al., 2008; Theophilides & Koutselini, 2010). From a practical perspective, OBEs should not be implemented as direct substitutes for closed-book assessments without intentional scaffolding. For instructors, particularly in high-enrolment courses, the results of this pilot study support the use of sequenced and constructively aligned assessments that allow students to develop and refine skills needed for success (Biggs, 2014). The effectiveness of OBEs can be strengthened if grounded in SRL theory and constructive alignment (Biggs, 2014; Zimmerman, 2000). Thus, students must adapt their study strategies to match assessment demands, while instructors ensure alignment between learning objectives and evaluation methods, with implications for scaffolding student exposure to this format and redesigning course assessments toward applied learning outcomes (Berry & Jensen, 2023; Biggs, 2014; Zimmerman, 2000).

Misalignment between perceived difficulty and actual performance

Similar to this study, it has been reported in the literature that students can perform worse on OBEs, even though this exam format is perceived to be easier (Boniface, 2006; Theophilides & Koutselini, 2010). Several reasons have been cited for seeing lower scores on OBEs compared to CBEs: the questions on the OBE may truly have been more difficult, students prepared less, and/or students did not properly use their study materials and spent most of the exam looking up answers (Boniface, 2006; Theophilides & Koutselini, 2010).

This provides supporting evidence that students may have misconceptions about the difficulty of OBEs, and it is recommended that courses implementing OBEs should outline expectations and study skills to help students meet the learning objectives. Although this study's analysis of grade variation between CBEs and OBEs in a single course provides insight into student performance, this pilot study is limited by its reliance on quantitative data. However, because the data collection occurred prior to the availability of GenAI, students were unable to use GenAI models to assist them with the OBEs. Results of this pilot could be strengthened by



including qualitative data to identify how students prepared for CBEs versus the OBEs, and to identify student perceptions of exam difficulty.

Academic integrity considerations with online exams

There is often concern about whether scores from online exams are meaningful, whether or not proctoring is involved (Chan & Ahn, 2023). Taking an exam online can create a temptation to breach academic integrity since students are not physically in a classroom being monitored while test-taking (Chan & Ahn, 2023). Exams are often high-stakes assessments influencing most of the overall course grade, which can further encourage students to take whatever means necessary to perform well (Chan & Ahn, 2023). Three factors seem to encourage cheating: incentive, opportunity, and rationalisation (Lee & Fanguy, 2022). Internal and external pressures creating motivation for breaching academic conduct (incentive), having the ability to cheat due to inadequate detection methods (opportunity), and feeling like these practices are still in line with the individual's ethics (rationalisation) can all contribute to academic dishonesty (Lee & Fanguy, 2022). However, previous studies have found scores on online exams to be similar to scores on the comparable in-person exams (Chan & Ahn, 2023; Holden et al., 2021). One recent study used data from 18 courses offered during spring 2020 from a large, Midwestern University in the United States and compared scores from the proctored in-person exams and unproctored online exams within the same courses. The results showed the scores from the online exams were similar to the in-person exam scores, thus suggesting online exams were still able to provide a meaningful assessment of student learning and, even if cheating occurred, it did not drastically increase exam scores (Chan & Ahn, 2023).

Considerations for student success in online learning

Although online OBEs have been found to be favourable for some students because of the increased test-taking time and flexibility, other students report concerns with the test-taking window (e.g., 24-hour period), harder questions, inexperience with OBE and expectations, and/or general technical issues (Slack & Priestley, 2022). This supports the need to properly prepare students for the format and expectations of OBEs prior to the test. Instructors could consider doing sample OBE test questions during class as a group or assigning them as



homework to familiarise students with the difficulty level and question style. Providing resources that outline preparation suggestions for the OBE can also be helpful for students.

Factors that contribute to academic success typically include regular class attendance, course participation, help-seeking behaviours (e.g., attending office hours, review sessions), and completing additional opportunities, such as extra credit (Moore & Jensen, 2007). In an online learning environment, there can be multiple barriers including lack of social connection, technical difficulties, motivation, and consistent and stable internet access. (Pei & Wu, 2019). Students also need to have proficient technological skills to navigate the fully online course platform (Summers et al., 2005). Successful online learning especially requires self-regulated strategies since it is up to the student to complete lectures and assignments on-time and be proactive with help-seeking strategies (Summers et al., 2005). Motivation can also be a major concern for online learning since this is directly tied to course performance (Summers et al., 2005). Motivated students are more likely to complete the coursework, while unmotivated students are less likely to engage with the course, affecting their overall grade (Kebritchi et al., 2017). Because online learning takes place remotely, helping foster motivation and engagement in the course can be difficult due to the structure of having few synchronous interactions with the instructional team and other students (Kebritchi et al., 2017).

While there is evidence to support the benefits of both CBEs and OBEs, a balanced approach, in which the choice of exam format depends on the specific goals and context of each course, may be necessary (Durning et al., 2016). While some universities are moving away from proctoring, there are concerns about maintaining academic integrity in open-book formats, particularly in high-stakes courses. In the health professions, it is often argued that CBEs help with information recall, and CBE performance has been linked to improved practice outcomes (Durning et al., 2016). There is also a belief that OBEs are less difficult and may be more prone to academic misconduct than CBEs, rendering their scores less meaningful (Durning et al., 2016). Yet, the stress associated with high-stakes CBEs may drive some students to consider cheating if given the opportunity, highlighting the need for carefully designed assessments (Chan & Ahn, 2023).

Design and application of open-book exams: implications in the age of GenAI

Although this study was conducted prior to the widespread availability of GenAI, the findings remain pertinent to the current discussion on assessment in higher education. While there



are practical challenges to implementing open-book exams, including increased preparation time for instructors and longer grading periods, well-designed OBEs can emphasise authentic learning and critical thinking skills, which is crucial in the current GenAI landscape (Alier et al., 2024; Al Malki et al., 2023). Authentic assessments that present novel scenarios, case studies, or personalised reflections encourage students to use content from the course material and can make it more difficult to use GenAI to answer these types of higher-order questions (Alier et al., 2024). However, it is important to note that creating questions to assess higher order thinking skills for an OBE can take considerable time and effort for the instructor (Al Malki et al., 2023; Er et al., 2023). It may be helpful for universities to provide training and resources to faculty on designing effective OBEs and authentic assessments (Er et al., 2023). Instructors also need to ensure the allowable resources for the OBE are available for all students since there may be inequities with access (Al Malki et al., 2023). Overall, assessments should evolve to reflect changing times and learning methods, focusing less on memorisation and more on application of knowledge. Educators could consider strategies for incorporating higher-level thinking into assessments, even within the constraints of traditional exam formats.

Limitations

While the present study finds evidence of differential performance on open-book exams compared to closed-book exams in a high-enrolment, online course, there are several limitations to consider. First, the data analysis focuses on exam and overall course grades; incorporating qualitative data to understand how students prepared for the exams and their perceptions of exam difficulty would strengthen this analysis. Follow-up mixed-methods research could be conducted to expand on these initial findings, such as by including an exam wrapper. This would be a brief, structured reflection activity following the release of exam grades that prompts students to consider how they prepared and performed, and what future strategies to maintain or improve. Second, the external validity is limited, as this study was done in a single class at one university. Finding other courses at different universities that have transitioned from CBEs to OBEs in an online setting and including these in the analysis could increase the generalisability of these findings. Further, students were not randomly assigned to an exam format, thus selection effects cannot be ruled out and observed differences may partially reflect cohort-level variation. However, this study includes more than 5,500 students over six academic quarters, which captures variability across cohorts and time,



thus partially strengthening the generalisability of these findings. Third, while course content remained consistent across quarters, variation in specific exam questions may have introduced differences in difficulty or emphasis, which could affect the comparability of performance between quarters. Yet, exams across cohorts all aligned with the course learning objectives and contained similar evidence of learning. For the open-ended questions, members of the instructional team completed thorough training on how to grade these based on a standardised rubric. Each question was graded by 1-2 members of the team, and the grading was spot-checked for consistency by the instructor and Head Teaching Assistants. With the exception of the grade inflation observed in spring 2020, which was the first quarter of emergency remote teaching (Tillinghast et al., 2023), the exam scores were similar between quarters. Lastly, this study was conducted prior to the widespread use and availability of GenAI; thus, the findings reflect student performance from a pre-GenAI assessment environment. Although the assessment landscape in higher education has considerably changed in the past few years, the results from this pilot can be used to inform assessment design in the current GenAI era as they provide a useful baseline for understanding the pedagogical value of OBEs versus CBEs.

Conclusion

In this pilot study, students taking the OBE performed worse on the first exam but better on the second. This suggests that there may have been a learning curve in figuring out how to approach this exam type. It is recommended that courses implementing OBEs outline expectations and study strategies to help students meet the learning objectives.

Future research should consider collecting qualitative feedback to identify common themes on how students perceive and prepare for OBEs. In addition, conducting sub-analyses based on demographic characteristics could help identify how certain student groups respond to OBEs. Furthermore, understanding if and how students are using GenAI in the OBE setting is essential for evaluating whether these assessments continue to measure intended learning outcomes. Investigation into how students engage with GenAI for OBEs could help instructors design improved assessments that emphasise authentic learning and critical thinking skills, as well as informing institutional policies related to academic integrity. Ultimately, the shift towards OBEs should be guided by careful alignment with learning objectives, intentional



question design that promotes critical thinking, and consideration for the institutional resources available for supporting academic integrity in online education.

Acknowledgements

The author is grateful to the UC Davis Scholarship of Teaching and Learning community for inspiring and supporting this work and extends sincere appreciation to the students who made this study possible.

Disclosure statement

The author did not use generative AI technologies in the creation of this manuscript.

The author reports there are no competing interests to declare.

Funding

No funding was received in support of the research contained in this manuscript.

References

- Agarwal, P. K., & Roediger, H. L. III. (2011). Expectancy of an open-book test decreases performance on a delayed closed-book test. *Memory*, 19(8), 836–852. <https://doi.org/10.1080/09658211.2011.613840>
- Alessio, H. M., & Messinger, J. D. (2021). Faculty and student perceptions of academic integrity in technology-assisted learning and testing. *Frontiers in Education*, 6, Article 629220. <https://doi.org/10.3389/feduc.2021.629220>
- Alier, M., García Peñalvo, F. J., & D. Camba, J. (2024). Generative artificial intelligence in education: From deceptive to disruptive. *International Journal of Interactive Multimedia and Artificial Intelligence*, 8(5), 5–14. <https://doi.org/10.9781/ijimai.2024.02.011>
- Al Malki, M., Al'Adawi, S., Al-Azzawi, N., & Al-Abri, K. (2023). Open-book-based assessment during COVID-19: Challenges and opportunities. *International Journal of Learning, Teaching and Educational Research*, 22(8), 146–168. <https://doi.org/10.26803/ijlter.22.8.9>
- Berry, L. R. & Jensen, P. A. (2023). Should open-book, open-web exams replace traditional closedbook exams in Science in higher education? An evaluation of their effectiveness in



different disciplines. *Journal of Learning Development in Higher Education*, (28).

<https://doi.org/https://doi.org/10.47408/jldhe.vi28.1030>

Biggs, J. (2014). Constructive alignment in university teaching. *HERDSA Review of Higher Education*, 1, 5–22. <https://herdsa.org.au/herdsa-review-higher-education-vol-1/5-22>

Block, R. M. (2012). A discussion of the effect of open-book and closed-book exams on student achievement in an Introductory statistics course. *Primus*, 22(3), 228–238.

<https://doi.org/10.1080/10511970.2011.565402>

Boniface, D. (2006). Candidates' use of notes and textbooks during an open-book examination. *Educational Research*, 27(3), 201–209.

<https://doi.org/10.1080/0013188850270307>

Cavanaugh, J., Jacquemin, S. J., & Junker, C. R. (2022). Variation in student perceptions of higher education course quality and difficulty as a result of widespread implementation of online education during the COVID-19 pandemic. *Technology, Knowledge and Learning*, 28(4), 1787–1802. <https://doi.org/10.1007/s10758-022-09596-9>

Chan, J. C. K., & Ahn, D. (2023). Unproctored online exams provide meaningful assessment of student learning. *PNAS USA*, 120(31), Article e2302020120.

<https://doi.org/10.1073/pnas.2302020120>

Durning, S. J., Dong, T., Ratcliffe, T., Schuwirth, L., Artino, A. R., Jr., Boulet, J. R., & Eva, K. (2016). Comparing open-book and closed-book examinations: A systematic review. *Academic Medicine*, 91(4), 583–599. <https://doi.org/10.1097/ACM.0000000000000977>

Er, H. M., Wong, P. S., & Nadarajah, V. D. (2023). Remote online open book examinations: Through the lenses of faculty and students in health professions programmes. *BMC Medical Education*, 23(1), Article 397. <https://doi.org/10.1186/s12909-023-04368-0>

Fischer, C., Xu, D., Rodriguez, F., Denaro, K., & Warschauer, M. (2020). Effects of course modality in summer session: Enrollment patterns and student performance in face-to-face and online classes. *The Internet and Higher Education*, 45, Article 100710

<https://doi.org/10.1016/j.iheduc.2019.100710>

Hartnett, M., Butler, P., & Rawlins, P. (2023). Online proctored exams and digital inequalities during the pandemic. *Journal of Computer Assisted Learning*, 39(4), 1103–1115.

<https://doi.org/10.1111/jcal.12813>

Heijne-Penninga, M., Kuks, J. B., Schonrock-Adema, J., Snijders, T. A., & Cohen-Schotanus, J. (2008). Open-book tests to complement assessment-programmes: Analysis of open and



- closed-book tests. *Advances in Health Sciences Education*, 13(3), 263–273.
<https://doi.org/10.1007/s10459-006-9038-y>
- Holden, O. L., Norris, M. E., & Kuhlmeier, V. A. (2021). Academic integrity in online assessment: A research review. *Frontiers in Education*, 6, Article 639814.
<https://doi.org/10.3389/feduc.2021.639814>
- Kebritchi, M., Lipschuetz, A., & Santiago, L. (2017). Issues and challenges for teaching successful online courses in higher education. *Journal of Educational Technology Systems*, 46(1), 4–29. <https://doi.org/10.1177/0047239516661713>
- Kerrigan, J., Cochran, G., Tabanli, S., Charnley, M., & Mulvey, S. (2022). Post-COVID changes to assessment practices: A case study of undergraduate STEM recitations. *Journal of Educational Technology Systems*, 51(2), 192–201.
<https://doi.org/10.1177/00472395221118392>
- Lee, K., & Fanguy, M. (2022). Online exam proctoring technologies: Educational innovation or deterioration? *British Journal of Educational Technology*, 53(3), 475–490.
<https://doi.org/10.1111/bjet.13182>
- Luo, J. (2024). A critical review of GenAI policies in higher education assessment: A call to reconsider the “originality” of students’ work. *Assessment & Evaluation in Higher Education*, 49(5), 651–664. <https://doi.org/10.1080/02602938.2024.2309963>
- Martin, F., Kim, S., Bolliger, D. U., & DeLarm, J. (2025). Assessment types, strategies, and feedback in online higher education courses in the age of artificial intelligence: Perspectives of instructional designers. *TechTrends*, 69(6), 1330–1346. <https://doi.org/10.1007/s11528-025-01115-8>
- Milone, A. S., Cortese, A. M., Balestrieri, R. L., & Pittenger, A. L. (2017). The impact of proctored online exams on the educational experience. *Currents in Pharmacy Teaching and Learning*, 9(1), 108–114. <https://doi.org/10.1016/j.cptl.2016.08.037>
- Mishra, L., Gupta, T., & Shree, A. (2020). Online teaching-learning in higher education during lockdown period of COVID-19 pandemic. *International Journal of Education Research Open*, 1, Article 100012. <https://doi.org/10.1016/j.ijedro.2020.100012>
- Moore, R. M. & Jensen, P. A. (2007). Do open-book exams impede long-term learning in introductory biology courses? *Journal of College Science Teaching*, 36(7), 46–49.
<https://www.jstor.org/stable/42992504>



- Mulryan-Kyne, C. (2010). Teaching large classes at college and university level: Challenges and opportunities. *Teaching in Higher Education*, 15(2), 175–185.
<https://doi.org/10.1080/13562511003620001>
- National Center for Education Statistics. (2026). *Table 311.15. Number and percentage of students enrolled in degree-granting postsecondary institutions, by distance education participation, location of student, level of enrollment, and control and level of institution: Fall 2023 and fall 2024*. U.S. Department of Education, Institute of Education Sciences.
https://nces.ed.gov/programs/digest/d25/tables/dt25_311.15.asp
- Neuhauser, C. (2002). Learning style and effectiveness of online and face-to-face instruction. *The American Journal of Distance Education*, 16(2), 99–113.
<https://doi.org/10.1207/S15389286AJDE1602>
- Newton, P. M., & Essex, K. (2023). How common is cheating in online exams and did it increase during the COVID-19 pandemic? A systematic review. *Journal of Academic Ethics*, 22(2), 323–343. <https://doi.org/10.1007/s10805-023-09485-5>
- Paechter, M., & Maier, B. (2010). Online or face-to-face? Students' experiences and preferences in e-learning. *The Internet and Higher Education*, 13(4), 292–297.
<https://doi.org/10.1016/j.iheduc.2010.09.004>
- Palvia, S., Aeron, P., Gupta, P., Mahapatra, D., Parida, R., Rosner, R., & Sindhi, S. (2018). Online education: Worldwide status, challenges, trends, and implications. *Journal of Global Information Technology Management*, 21(4), 233–241.
<https://doi.org/10.1080/1097198x.2018.1542262>
- Pei, L., & Wu, H. (2019). Does online learning work better than offline learning in undergraduate medical education? A systematic review and meta-analysis. *Medical Education Online*, 24(1), Article 1666538. <https://doi.org/10.1080/10872981.2019.1666538>
- Permzadian, V., & Cho, K. W. (2025). Does test format affect learning? A meta-analysis comparing the effect of closed-book and open-book examinations. *Assessment & Evaluation in Higher Education*, 50(7), 1098–1118. <https://doi.org/10.1080/02602938.2025.2502590>
- Rowland, C. A. (2014). The effect of testing versus restudy on retention: a meta-analytic review of the testing effect. *Psychological Bulletin*, 140(6), 1432–1463.
<https://doi.org/10.1037/a0037559>
- Slack, H. R., & Priestley, M. (2022). Online learning and assessment during the Covid-19 pandemic: Exploring the impact on undergraduate student well-being. *Assessment &*



Evaluation in Higher Education, 48(3), 333–349.

<https://doi.org/10.1080/02602938.2022.2076804>

Szmyd, K. & Mitera, E. (2024). The impact of artificial intelligence on the development of critical thinking skills in students. *European Research Studies Journal*, 27(2), 1022–1039.

<https://doi.org/10.35808/ersj/3876>

Summers, J. J., Waigandt, A., & Whittaker, T. A. (2005). A comparison of student achievement and satisfaction in an online versus a traditional face-to-face statistics class. *Innovative Higher Education*, 29(3), 233–250. <https://doi.org/10.1007/s10755-005-1938-x>

Tate, T., & Warschauer, M. (2022). Equity in online learning. *Educational Psychologist*, 57(3), 192–206. <https://doi.org/10.1080/00461520.2022.2062597>

Theophilides, C., & Koutselini, M. (2010). Study behavior in the closed-book and the open-book examination: A comparative analysis. *Educational Research and Evaluation*, 6(4), 379–393.

<https://doi.org/10.1076/edre.6.4.379.6932>

Tillinghast, J. A., Mjelde, D. J. W., & Yeritsyan, A. (2023). COVID-19 and grade inflation: Analysis of undergraduate GPAs during the pandemic. *Sage Open*, 13(4), 1–12.

<https://doi.org/10.1177/21582440231209110>

Wiggins, G. (1990). The case for authentic assessment. *Practical Assessment, Research & Evaluation*, 2(2), Article 2. <https://doi.org/10.7275/ffb1-mm19>

Xu, Y., & Harfitt, G. (2018). Is assessment for learning feasible in large classes? Challenges and coping strategies from three case studies. *Asia-Pacific Journal of Teacher Education*, 47(5), 472–486. <https://doi.org/10.1080/1359866x.2018.1555790>

Zagury-Orly, I., & Durning, S. J. (2021). Assessing open-book examination in medical education: The time is now. *Medical Teacher*, 43(8), 972–973.

<https://doi.org/10.1080/0142159X.2020.1811214>

Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13–39). Academic Press. <https://doi.org/https://doi.org/10.1016/B978-012109890-2/50031-7>

Author details

Deborah Fetter is an Associate Professor of Teaching in the Department of Nutrition at the University of California, Davis. She teaches large-enrolment undergraduate courses and conducts research in the scholarship of teaching and learning, with a focus on assessment,



educational technology, online learning, and student success. Her work evaluates evidence-based instructional practices that enhance student engagement and learning in higher education. Through her research and teaching, she aims to develop inclusive, transferable approaches that improve learning experiences and support diverse learners across higher education contexts.

ORCID

Deborah Fetter  <https://orcid.org/0000-0002-8531-9760>

