Comparing the predictive validity of English placement scores by proctor modality

Cindy L. James
Thompson Rivers University, Canada

Veronika Larsen
University of British Columbia, Canada

Entrance exams in higher education typically are proctored onsite by testing professionals or faculty members, but over the past two decades the utilization of remote proctoring services has become more common. To ensure remote proctoring is as effective as onsite proctoring, numerous studies investigating the comparability of test scores by proctor modality have been conducted, but very few have explored the predictive validity of test scores based on proctoring type. This study explored both issues by comparing the impact of onsite live and remote live proctoring on English entry assessment scores and subsequent performance in first-year English courses. The results revealed no significant differences in test scores on three English placement tests by proctor type, nor were there any significant differences in student performance in first year English courses. Factors related to active proctoring and features of the entrance exam provide some insight into these findings.

Keywords: first-year English, language testing, predictive validity, remote proctoring

In higher education, entrance tests are commonly utilized to place students into courses or programs based on their assessed skills. To ensure the reliability and validity of the scores, it is vital that testing practices and conditions are consistent and regulated for all candidates (American Educational Research Association et al., 2014). Proctors play a crucial role in this process as they ensure standardized procedures and guidelines are

Email address for correspondence: cjames@tru.ca

© The Author(s) 2026. This is an open access article distributed under the terms of the Creative Commons Attribution 4.0 International License, which permits the user to copy, distribute, and transmit the work provided that the original authors and source are credited. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

followed. Conventionally, proctoring involves testing professionals or faculty administering exams onsite in-person, but the growth of online education, related technologies, and the COVID 19 pandemic, have led to the establishment and expansion of remote proctoring. Remote proctoring, also known as virtual, online, or digital proctoring, utilizes technology to administer and monitor the exam session because the proctor is not onsite with the candidate.

Although onsite proctoring and remote proctoring of entrance tests both adhere to standardized procedures, they do differ in practice and function. The primary difference is the testing environment as onsite testing is usually conducted in testing centers equipped with modern technology and quiet spaces to minimize disruptions, distractions, and security risks. Remote testing usually takes place in the candidate's home, which provides more flexibility and perhaps comfort, but does not guarantee the optimum conditions in terms of technology, distractions, and privacy. Testing in a remote location also necessitates some adjustments to proctoring activities: in most cases the proctor can only utilize a webcam and microphone to verify the identity of the candidate and monitor the activities of the candidate during the test session. Onsite testing allows the proctor to physically authenticate the candidate's identity as well as view the candidate and the testing environment from many angles to detect unusual behavior or restricted items throughout the exam.

One other significant difference between onsite and remote is that onsite proctoring exclusively involves proctors supervising a test session in real time, a practice known as *Live* proctoring. Live proctoring is also available remotely, with the proctor supervising constantly, in real time, but from a different location. Most of the other remote proctoring options do not include live proctoring. For instance, another common practice is *Remote Recorded,* in which the session is recorded and reviewed after the fact by a proctor to identify if there are any incidents to be addressed. *Automated* proctoring in which artificial intelligence is used to monitor the test session is another method of remote proctoring, although it can and often is used in conjunction with Remote Live and Remote Recorded options.

Given these differences, it is important to confirm that proctoring mode does not affect test performance. This is especially true for high-stakes tests, such as admission and placement testing in higher education, as the test scores determine access to higher education and possibly length of program if remedial coursework is required. Due to the proliferation of remote proctoring, there has been substantial research exploring the efficacy of the various modes of proctoring, as evident by the existence of studies involving reviews (Han et al., 2024; Kuleva & Miladinov, 2024; Öncül, 2021) and meta-analyses on the topic (Steger et al., 2020). The topics explored in these studies range from technological advancements, especially those related to detection of cheating, students' performance by proctor mode, user perceptions of remote versus onsite proctoring, transition issues, and problems related to digital proctoring services (Han et al., 2024; Kuleva & Miladinov, 2024; Öncül, 2021; Steger et al., 2020). According to the review by Han et al. (2024) of the assorted topics researched, students' performance in different proctoring environments has garnered the most attention. The majority of these performance studies have focused on comparing test scores by proctor modality. Typically, this involves comparing test scores for sessions proctored onsite live, the well-established practice, to sessions proctored remotely or not at all.

The results from these studies have varied somewhat yet are unified in their expressed need for more studies. In addition to supplementing the current research, such studies could expand their research agenda by also exploring the predictive validity of entrance test scores by proctor modality. This is crucial since the purpose of entrance testing is to ensure students have the skills necessary to succeed in their academic studies. If proctor type affects the validity of test scores, there is a strong possibility it could also impact student performance in academic courses. Hence, future studies could still compare test scores by proctor mode, but to enhance their contribution to this topic they also could explore the academic performance of students who were admitted or placed into higher education courses or programs based on the entrance scores by proctor mode. To date, there is a paucity of such predictive validity studies by proctor mode. This study was designed to fill this void by evaluating:

1) the comparability of English entrance exam scores by proctor mode, and

2) the predictive validity of an English entrance exam by comparing the academic performance of students placed in first-year English courses based on their scores by proctor mode.

Literature review

Since there are numerous terms to describe the various proctoring modalities, for consistency in this article, the following four terms will be used to describe the most common proctor modes. *Onsite Live* proctoring (OL) refers to the traditional onsite, in-person proctoring that takes place in a testing center or classroom in real time. *Remote Live* proctoring (RL) refers to situations in which the proctor is not in the same location as the candidate, but virtually present to authenticate examinees and monitor the testing session in real time. *Remote Recorded* proctoring (RR) refers to testing sessions that are digitally recorded and reviewed by a proctor after the session is completed. *Remote Automated* proctoring (RA) refers to the use of artificial intelligence to monitor exams. Usually this is used in combination with *Remote Live* proctoring or *Remote Recorded* proctoring as AI flags concerns for the proctor to address immediately for RL or after the fact for RR.

Numerous studies exploring the effects of proctor modality on test scores have been conducted. In higher education, many of these studies focus on proctored versus unproctored exams for courses offered on campus or through distance education. Most of these studies tend to report significantly higher test scores for unproctored exams as compared to onsite live proctored exams (Speth & Burr, 2025; Vazquez et al., 2021; Zhang et al., 2024); or remote live proctored exams (Alessio et al., 2018; Daffin Jr. & Jones, 2018; Goedl & Malla, 2020; Hylton et al., 2016); or remote recorded proctored exams (Akaaboune et al., 2022; Alessio et al., 2017; Dendir & Maxwell, 2020; Reisenwitz, 2020; Shaushenova et al., 2021; Speth & Burr, 2025). In these studies, a variety of explanations are posited to explain the higher scores on unproctored exams versus proctored exams, with the most common assertion being that cheating had occurred. However, a few studies found no differences between test scores of unproctored exams versus OL proctored exams (Hollister & Berenson, 2009; Howard, 2020; Ladyshewsky, 2015); RL

proctored exams (Gulati et al., 2024); or RR proctored exams (Kharbat & Daabes, 2021; Oeding et al., 2024).

Given the tendency of higher scores on unproctored exams as compared to proctored exams and the plausible reason for it, the findings from most of the aforementioned studies suggest that exams should be proctored, but what type of proctoring is best? Numerous studies have been conducted to answer that question, and in most cases, these studies have revealed little to no significant differences in test scores by proctor modality. For instance, Hope et al. (2021) found no significant gaps in the scores of an online medical exam administered at testing centers with live proctoring to 6048 candidates and remotely via live proctoring to 2044 candidates. Similarly, a study by Emeka et al. (2024) involving 428 computer science students found no significant difference between test scores for students proctored onsite live and proctored remote live via Zoom. Based on their results, the authors concluded computer science exams could be delivered effectively using different proctoring methods (Emeka et al., 2024). Woldeab and Brothen (2019) also reported no significant differences on a final exam taken by 631 psychology students. The majority of the students (n = 587) were proctored in a testing center while the rest (n = 44) were remotely proctored in real time (Woldeab & Brothen, 2019).

In some cases, findings have varied diachronically between cohorts, especially in health and medicine educational fields. Using a retrospective approach, Hall et al. (2021) compared the performance on composite exams for pharmacy students. In this study, scores from 387 exams proctored remote live during the pandemic in 2020 were compared to 368 historical scores from a previous cohort proctored onsite live in 2019. The results revealed for the first sitting, exam scores were lower for the remote live proctored cohort but for the next two sittings there were no significant differences in exam scores by proctor modality (Hall et al., 2021). Another retrospective study involving 898 undergraduate medical students conducted by Jaap et al. (2021) reported similar findings. For this study, students in their fourth year performed better on RL proctored exams during the pandemic compared to OL proctored exams prior to the pandemic, but performances on exams for fifth year students were equivalent by proctor mode (Jaap et al., 2021). Even with this difference, Jaap et al. (2021) concluded that proctor mode had

no "negative effects on candidate performance." During the same year, another study involving high-stakes testing in the medical education field conducted by Andreou et al. (2021) offered complementary findings: there were no significant score differences between 121 candidates proctored OL as compared to 472 candidates proctored RL while completing a proficiency test for admissions to the Advanced Master of General Practice. Based on these results, the authors concluded, "Remote proctoring could allow for diverse and more flexible ways of administering exams without sacrificing academic integrity and exam quality" (Andreou et al., 2021, p.7).

Larger-scale studies grounded in medical testing in higher education have also yielded comparable results. Ong et al. (2024) examined scores from 27,115 candidates completing the Internal Medicine In Training Examination to reveal minor differences between 42% of residents proctored remote live versus 58% of residents proctored onsite live. Although these differences were statistically significant, given the differences were under one percentage point, they were deemed educationally nonsignificant (Ong et al., 2024).

Other retrospective studies have offered conflicting patterns, such as that conducted by Arnold (2022), which compared thousands of exams in economic courses proctored onsite live between 2015 and 2018 versus remote live between 2019 and 2020. Overall, no general pattern emerged but mean grades on RL proctored exams compared to OL proctored exams were significantly higher in accounting, microeconomics, and organizational and strategy courses, and significantly lower for mathematics, statistics, and econometrics courses (Arnold, 2022). In contrast, a study conducted prior to the pandemic involving microeconomics revealed 29 students proctored remote live scored lower on a final exam than 28 students proctored onsite live (Harmon et al., 2017). Since remote proctoring was relatively new at the time of this study, students' unfamiliarity with this type of proctoring was suggested as an explanation for the results.

In terms of studies involving English assessments, although there are many articles focusing on remote proctoring (e.g., Isbell et al., 2023; Isbell & Kremmel, 2020; Ockey, 2021), very few studies involve comparisons of language test scores by proctor mode. One exception was a study conducted by Zumbo (2021), in which the Canadian Academic

English Language (CAEL) test scores from sessions delivered at test centers to 765 candidates were compared to those delivered remotely online to 690 candidates. The CAEL test evaluates academic English speaking, reading, listening, and writing skills for candidates seeking admission to Canadian higher education and professional organizations. Using matching concordance analyses, test takers earned statistically similar band scores whether they took the CAEL at a test center or online (Zumbo, 2021). As such, the conclusion from this study was "the test scores from the online and test center modes of administration are fully interchangeable" (p. 141).

Similar but not as conclusive results were reported by Valkova (2024) who conducted a study comparing test performance on Aptis General, an online British Council language proficiency test, by proctor modality. The 70 participants in this study were English faculty at a Mexican university who took the language assessment in a test center with a live proctor and remotely using a record and review option with AI flagging possible issues. The findings revealed no significant differences between scores from OL and RR proctoring for the Listening, Grammar and Vocabulary components, but for the Reading component, the mean scores for the OL proctoring were significantly lower than the mean scores for the RR proctoring (Valkova, 2024). Test order and environment were presented as a possible explanation for the difference on the Reading component. Specifically, since the Reading test was the final section of the Aptis test, the authors suggested that perhaps students testing remotely had become more comfortable with the platform, the remote proctor system and the testing environment as compared to the start of the exam and thus performed better (Valkova 2024).

Another study involving English and mathematics entry assessments was conducted by James (2022). For this study, entry assessment scores for 918 students either proctored onsite live or remote live were compared. No significant differences were found between proctor modality and performance on three English tests and two mathematics tests, but there was a significant difference between proctored groups for the advanced mathematics tests. For that test, the RL proctored students scored significantly higher than the OL proctored students (James, 2022). The author could not explain this result,

and hence, concluded more research, especially for high-stakes math assessments, by proctor modality was warranted.

Meanwhile, Castellano and others (2024) conducted a comparison study by proctor mode for a large-scale admissions test with three subject areas: Quantitative, Verbal and Writing. For this study, 101 candidates were proctored live at a testing center and 85 were tested remotely in their homes. The findings revealed no significant difference in mean scores across all three tests by proctor mode (Castellano et al., 2024) though missing from this study was a description of the type of remote proctoring employed.

Although the results from these studies evaluating the comparability of test scores via proctor mode are encouraging, they alone do not fully validate the scores. To do that, it is necessary to conduct more in-depth studies exploring other types of validity. Predictive validity studies that measure the extent to which test scores accurately predict future performance would be ideal for research involving entry assessment scores and performance in higher education. Assuming the entrance scores are valid for all types of proctoring modes, the students should be successful in courses or programs based on their placement test scores, but is that the case? According to one of the rare predictive validity studies involving proctor type in higher education, it is not (Hample, 2022). In this study, Hample (2022) explored the differences of college placement scores by proctor mode, onsite live versus no proctoring, and performance in subsequent courses. The findings revealed mean scores for the mathematics placement test for the group of students who were proctored were significantly lower than for the group of students who were not proctored (Hample, 2022). Based on these scores, students were placed in four different levels of college mathematics courses. For all four courses, the mean course GPAs were greater for the 627 students placed via proctored exams than for 1664 students who were placed via unproctored exams. These differences were significant for the college algebra and pre-calculus courses but not pre-college math or calculus courses (Hample, 2022). The results appear to indicate the unproctored assessments were not as valid as the proctored assessments and thereby impacted the accuracy of the placement.

Even though there is ample research into the comparability of test scores by proctor modality, research into the predictive validity of test scores by proctor modality is scarce. This study was designed to address this scarcity by comparing students' performances on an English entrance exam *and* achievement in a subsequent first-year English course by proctor mode.

Method

This study was conducted at a public higher education institution offering undergraduate and master's degrees as well as certificates and diplomas. In 2023-24, this institution served over 29,000 students (Thompson Rivers University, 2024). For undergraduate programs, the primary admission requirement is completion of high school and evidence of English language proficiency, as demonstrated through coursework (high school or remedial) or English assessment test scores. For students with English as a second or alternate language, the assessment tests typically utilized are TOEFL iBT, IELTS, and Duolingo. For students whose primary language is English the typical assessment tool utilized is ACCUPLACER®.

Entrance assessment tool - ACCUPLACER®

ACCUPLACER® is a web-based testing platform that offers a series of tests to measure students' key skills in English and mathematics needed for academic success. For this study, three English tests were utilized: Reading, Writing and WritePlacer®. The Reading and Writing tests consist of multiple-choice questions administered using computer-adaptive technology. This process, which tailors the sequence of questions to the examinee's level of ability, requires fewer test items, thereby reducing testing time while enhancing security (College Board, 2018a). The WritePlacer® requires candidates to complete an on-demand essay to a randomly assigned prompt. The candidate's response is scored utilizing automated assessment technology (College Board, 2018b). At this institution, the two multiple choice tests, Reading and Writing, are untimed, while the on-demand writing test, WritePlacer®, has a one-hour time limit.

The scoring for the two multiple-choice tests, Reading and Writing, ranges from 200 to 300, while the scoring for the WritePlacer® test ranges from 0 to 8 (College Board, 2017). The cutscores for first-year English at TRU, established based on a variety of validity studies, are 250 for Reading and Writing, and a five on WritePlacer®. With these scores or higher, applicants can enroll in first-year English courses that focus on university or professional writing.

The ACCUPLACER® tests are administered onsite live at the TRU Assessment Centre, regional centers, and other external testing sites or remote live by one of the online proctoring services approved by the College Board. For the former, the Assessment Centre is a Certified Test Center by the National College Testing Association (NCTA), indicating that the Centre and its staff operate and adhere to the professional standards and best practices for all types of testing (National College Testing Association, 2025). Although not necessarily certified, the regional centers, and other external testing sites such as other higher education testing centers or high schools are vetted by the TRU Assessment Centre to ensure they also adhere to similar standards. As for the remote proctors, the College Board contracts vendors to provide remote proctoring services. These services are provided 24 hours a day year-round, and up until 2024 included Remote Live and Remote Recorded options. After 2024, only Remote Live proctoring services were allowed for ACCUPLACER® testing. The TRU Assessment Centre always employed the Remote Live option to ensure a consistent proctor experience whether students were testing onsite or remotely.

Participants

From 2018 to 2023, 778 applicants wrote the full English assessment, completing all three English ACCUPLACER® tests. Since these students completed the English assessment, the assumption is that they did not meet the entrance requirements at TRU, so it is likely the applicants were recent high school graduates who did not achieve 73% in grade 12 English, mature students who have been out of school for more than two years, or home schooled students who did not have an official grade for English 12. The average age of

these applicants was 25.5 (SD = 7.9). Of these, 537 were proctored in person and 241 were proctored online.

Out of the 778 applicants, 292 achieved the necessary scores to enroll in first-year English, and of these, 138 applicants enrolled and completed a first-year English course offered onsite at the university. The other 154 students who achieved the necessary scores appear to have entered programs that did not require students to complete a first-year English course or the students had yet to complete their first-year English course. The pathways for the students who did not achieve the necessary scores will be the focus of a future study. The average age for the subset of students who completed a first-year English course was 24.3 (SD = 6.9). For this subset of students, 113 were proctored in person and 25 were proctored online.

Data collection and analysis

The ACCUPLACER® English test scores for the 778 applicants and final grade in a first-year English course for the subset of 138 students served as the primary data sources for this study.

For the 778 applicant pool, descriptive statistics for the tests scores by proctor mode were tabulated and distributions of test scores by proctor mode were presented through histograms. *T*-tests were utilized to compare the mean test scores between the onsite live proctored and remote live proctored group for the Reading and Writing tests. Due to the ordinal nature of the WritePlacer® scores, a Mann-Whitney *U* test was employed to compare test score distributions between OL and RL proctored groups. The same descriptive and analytical statistics were calculated for the subgroup of students who completed an on-campus in-person first-year English course (n = 138).

To assess course performance, the English course grade each student received was collected and converted to a grade point, using the institutional numerical scale of 0 to 4.33 (Thompson Rivers University, 2017). On this scale, zero equates to a letter grade of F (failed) or DNC (did not complete) and 4.33 indicates an A+ letter grade. Descriptive statistics for the course grade points by proctor mode were tabulated. A *t*-test was utilized to compare course grade point averages (GPAs) between the onsite live and remote live

proctored students. Course performance was also classified into two categories: success for grades above a C- or nonsuccess for a grade of D, Failed (F) or Did Not Complete (DNC). Using cross-tabulation, the dichotomous outcome was reported and the percentage of successful students versus unsuccessful students by proctor modality was evaluated using the Fisher's Exact test.

For all calculations, a significance level of .05 was employed. Since there were multiple comparisons, the Bonferroni correction factor was applied, adjusting the level of significance to .017 for the comparison between all test scores by proctor mode (Table 2) and .0125 for the comparison between tests scores and GPA for the subgroup of students who completed a first-year English course (Table 4).

Results

All applicants

The descriptive statistics of the ACCUPLACER® English tests by proctor type for the applicant group were much the same (Table 1). The mean score for the ACCUPLACER® Reading test was slightly greater for the onsite live proctored group as compared to the remote live proctored group, but the medians were identical. For the ACCUPLACER® Writing test, the mean and median scores were marginally higher for the RL proctored group as compared to the OL proctored group. For the ACCUPLACER® WritePlacer® essay, mean scores and the median score were greater for the group that was proctored onsite live (Table 1).

Table 1. Descriptive statistics for ACCUPLACER® English test scores by proctored group

	Reading		Writing		WritePlacer®	
Statistics	Onsite Live	Remote Live	Onsite Live	Remote Live	Onsite Live	Remote Live
Mean	258.84	258.65	255.01	255.55	4.51	4.42
CI (95%) Upper	260.26	260.66	256.20	257.41	4.61	4.56
CI (95%) Lower	257.42	256.65	255.31	253.69	4.41	4.29
Median	259.00	259.00	255.00	256.00	5.00	4.00
SD	16.69	15.80	14.05	14.64	1.14	1.09

Maximum	300	300	300	300	8	7
Minimum	200	206	200	200	1	1
n	537	241	537	241	537	241

The distributions of the ACCUPLACER® Reading test scores (Figure 1), Writing test scores (Figure 2) and WritePlacer® test scores (Figure 3) by proctor type for the applicant group were comparable.

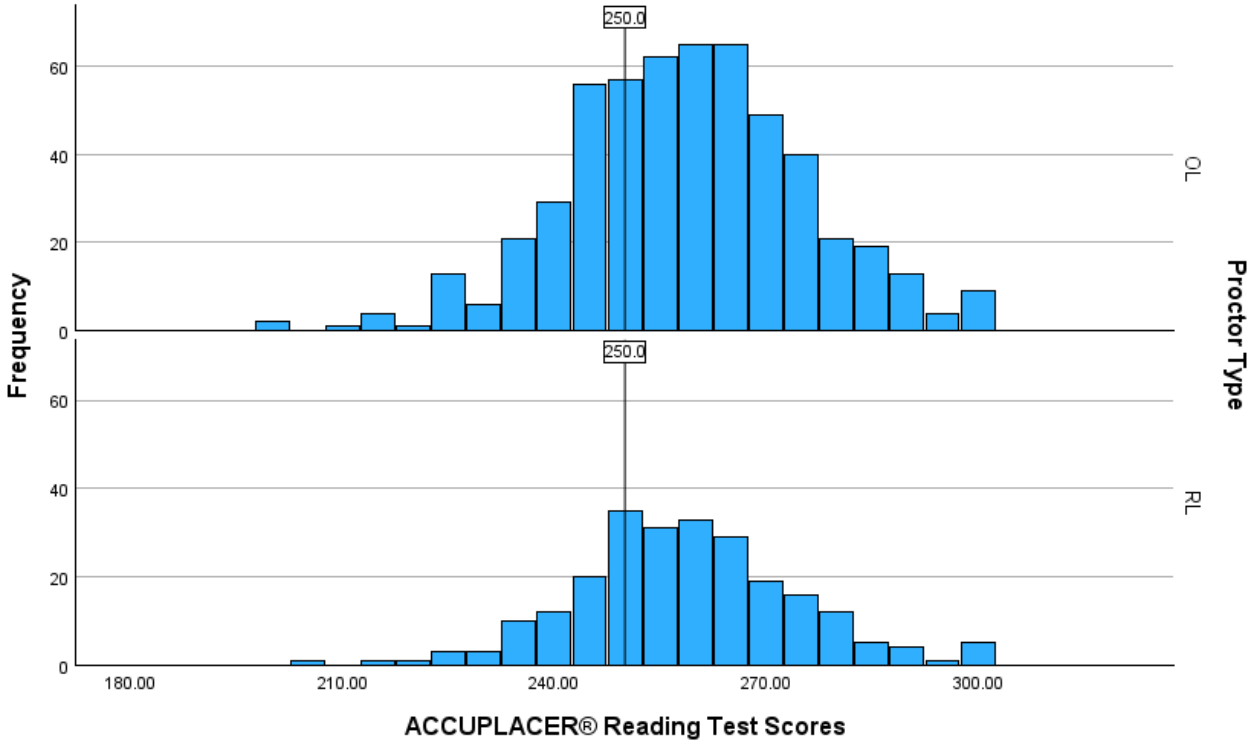

Figure 1. Histogram of the ACCUPLACER® reading test scores by proctor mode

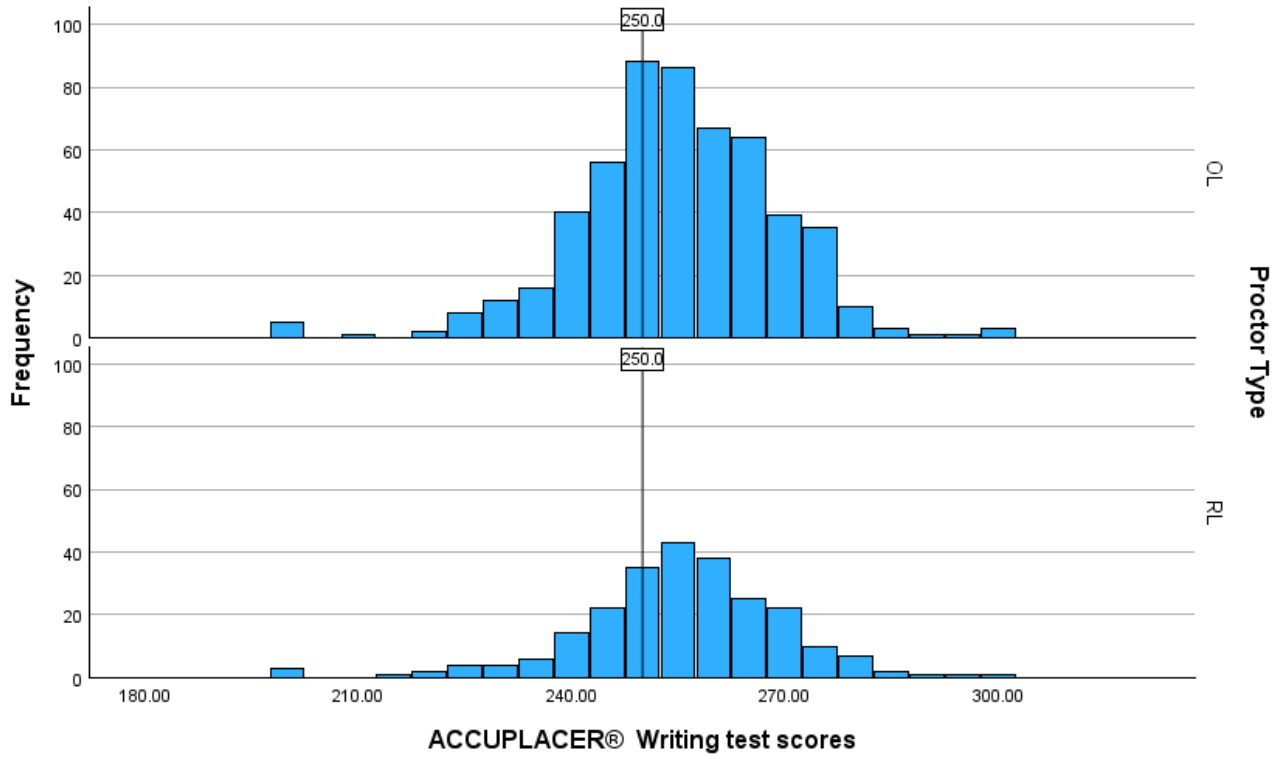

Figure 2. Histogram of the ACCUPLACER® writing test scores by proctor mode

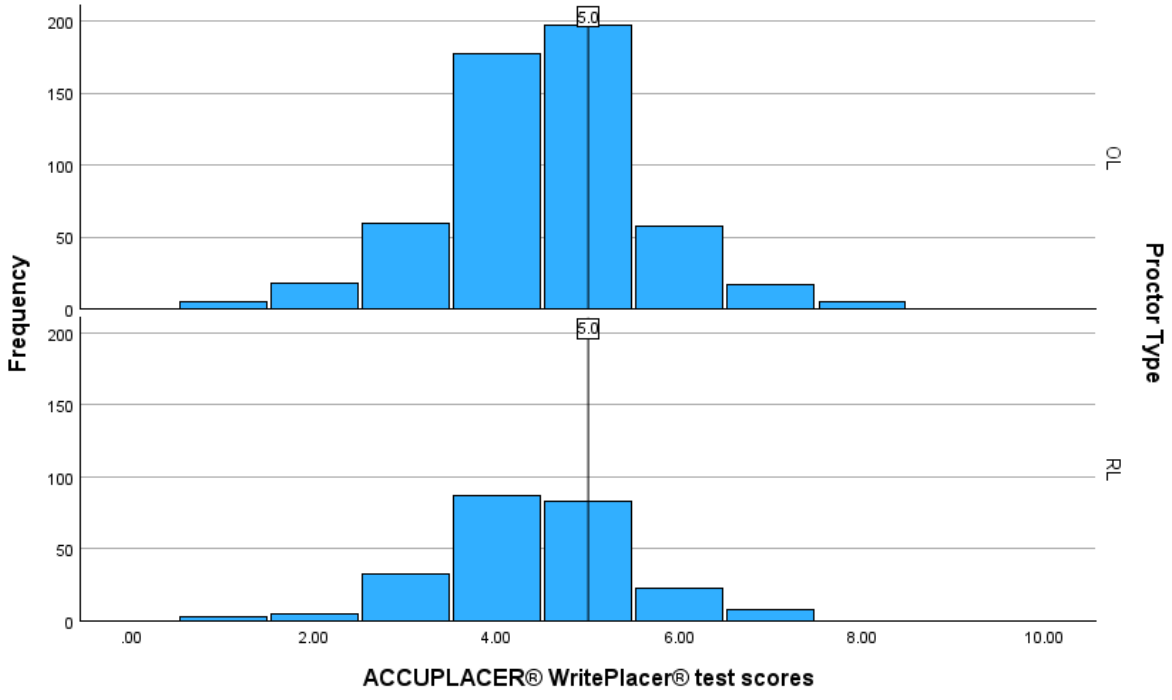

Figure 3. Histogram of the ACCUPLACER® WritePlacer® test scores by proctor mode

Based on *t*-tests, the Reading and Writing test mean scores were not significantly different between the onsite live and remote live proctored groups (Table 2). There also was no significant difference between the distribution of scores for the WritePlacer® test by proctored group based on the Mann-Whitney *U* test (Table 2).

Table 2. Comparison of test scores by proctored group

ACCUPLACER® test	Onsite Live		Remote Live		Difference		
	Mean	SD	*Mean*	SD	*Mean*	*T-Test*	*p**
Reading	258.84	16.69	258.65	15.80	0.19	t = .15	.882
Writing	255.01	14.05	255.55	14.64	0.54	t = .49	.625
	Mean Rank		Mean Rank		Mann-Whitney *U* Test		
WritePlacer®	394.89		377.50		z = 1.05		.296
n	537		241				

*two-tailed.

Subgroup of students who completed a first-year English course

The distributions of the ACCUPLACER® English tests by proctor type were commensurable for the subgroup of students who also completed a first-year English course (Table 3). The ACCUPLACER® Reading and Writing test mean and median scores were marginally greater for the remote live proctored group as compared to the onsite live proctored group; however, for the ACCUPLACER® WritePlacer® test, the mean and median were identical for both proctored groups (Table 3). As for the course GPA, again the means were nearly identical, but the median for students proctored remotely was slightly higher as compared to students proctored onsite (Table 3).

Table 3. Descriptive statistics for ACCUPLACER® English test scores and English course GPA by proctored group.

	Reading		Writing		WritePlacer®		Course GPA	
Statistics	Onsite	Remote	Onsite	Remote	Onsite	Remote	Onsite	Remote
Mean	266.84	268.88	262.59	263.96	5.40	5.40	3.14	3.19
95% CI Upper	269.17	273.41	264.58	267.83	5.54	5.64	3.33	3.68
95% CI Lower	264.51	264.35	260.60	260.09	5.26	5.16	2.95	2.69
Median	266.00	267.00	262.00	264.00	5.00	5.00	3.33	3.67
SD	12.51	10.98	10.68	9.37	.74	.58	1.02	1.21
Maximum	300	300	300	287	8	7	4.33	4.33
Minimum	250	255	239	248	5	5	.00	.00
n	113	25	113	25	113	25	113	25

Based on *t*-tests and the Mann-Whitney *U* test, there were no significant differences between the scores on the ACCUPLACER® tests and the performance in the first-year English courses by proctor modality (Table 4).

Table 4. Comparison of test scores and course GPA for participants by proctored group

Assessment	Onsite Live		Remote Live		Difference		
	Mean	SD	Mean	SD	Mean	*t*-test	*p**
ACCUPLACER® Reading	266.84	12.514	268.88	10.982	2.04	.753	.453
ACCUPLACER® Writing	262.59	10.675	263.96	9.374	1.37	.591	.555
First-year English course GPA	3.14	1.021	3.19	1.206	0.05	.207	.836
	Mean Rank		Mean Rank		Mann-Whitney *U* test		*p**
ACCUPLACER® WritePlacer®	68.74		72.92		.59		.553
n	113		25				

*two-tailed.

The success rate for the students proctored onsite live was slightly higher (93.8%) as compared to the students proctored remote live (92.0%) (Table 5). However, this difference was not significant based on the Fisher's exact test (two-tailed $p = .666$).

Table 5. Success rates in first-year English participants by proctored group

	First-year English Participants		
Proctor Group	Success	Total	Success Rate
Onsite Live	106	113	93.8%
Remote Live	23	25	92.0%
Both	129	138	93.5%

Discussion

Adhering to standardized procedures enhances the validity of entrance tests in higher education. Since proctoring is a key component of the procedure, it is imperative to ensure different forms of proctoring are valid. To contribute and expand on the existing proctor research, this study compared the English entrance scores of students who were proctored onsite live to that of students who were proctored remote live, and the performance of students who enrolled in a subsequent first-year English course by proctor mode. The results indicated test scores were not significantly impacted by the type of proctoring for all three English tests, nor was student performance in first-year English courses.

This study reaffirms the findings from previous studies that revealed little to no difference in test performance by candidates proctored onsite live as compared to candidates proctored remote live (Andreou et al., 2021; Castellano et al., 2024; Emeka et al., 2024; Hall et al., 2021; Hope et al., 2021; Jaap et al., 2021; Ong et al., 2024; Woldeab & Brothen, 2019). It also goes further to demonstrate that proctoring mode, specifically OL or RL, does not impact the predictive validity of the test scores. This differs from the findings by Hample (2022), but both studies appear to indicate that proctoring, especially onsite live or remote live, improves the predictive validity of test scores as compared to no proctoring.

One of the best explanations for finding no differences between test scores and course performance in this study is that onsite live and remote live proctoring both involve one of the most important components of proctoring, namely *active* proctoring. This is captured by the term "live" which refers to continuous proctoring in real-time throughout the test session. This constant monitoring helps to minimize academic integrity issues related to testing such as limiting opportunities for the candidates to access restricted materials or aid during the exam since the proctor can intervene whenever necessary to prevent or stop unacceptable behaviors. Active proctoring also may assist with preventing proxy testing and content theft.

Specific features associated with the testing tool also may have contributed to these findings. The computer-adaptive nature of the ACCUPLACER® testing platform results in questions on the multiple-choice tests being selected at random and depending upon the answer to the previous question. Consequently, it is difficult to predict question order, thereby limiting cheating opportunities as compared to a static multiple-choice test. Moreover, as with most computer-adaptive tests, ACCUPLACER® has a large item bank, further reducing predictability and preventing overexposure of any one item. As for the essay type question, the prompt is randomly assigned by the ACCUPLACER® testing platform, again, limiting opportunities to cheat as compared to a single prompt essay test. Test security is also enhanced by the lockout feature offered by the ACCUPLACER® system. While the test is in progress, if the student clicks outside the test administration window, the system will automatically lockout the student until the proctor intervenes

(College Board, 2018b). All these features are available for onsite or remote testing, providing a similar testing experience no matter the proctoring mode.

The findings from this study should reassure testing professionals, faculty and students involved with entrance assessments at post-secondary institutions. Institutional staff can be fairly confident about offering remote live proctoring services as an option for students requiring entrance assessments since test scores do not appear to differ between OL and RL proctored exams, nor does performance in English courses. The same is true for students, knowing test site does not appear to impact performance and having the flexibility to choose their preferred site should alleviate some of the stress related to entrance assessments. In addition, remote testing can be beneficial to some students in terms of eliminating timely travel and its associated costs, as well as providing a more familiar environment for testing.

Limitations and future research

This study has several limitations that hinder its generalizability and implications. First, the data were collected from a unique sample from one institution, so the findings may not apply to other post-secondary institutions since populations, programs and admission processes do differ. Second, the findings may only apply to entrance scores from the ACCUPLACER® testing system. Studies involving different assessment tools utilized by other institutions are needed to confirm or refute these findings. Third, this study focused on English entrance scores and placement into English courses, so the results may not be transferrable to other test content and placement into other subject areas, and hence, need to be investigated. Finally, this study compared onsite live to remote live proctoring, so the findings may not apply to other modes of remote proctoring. Future research will need to replicate this study to evaluate the comparability and predictive validity of test scores via other remote proctoring modalities.

Conclusion

Onsite live proctoring is the well-established mode of proctoring, but advances in technology, increased participation in online and distance learning plus a pandemic have led to the growth of remote proctoring. Since remote proctoring was designed to emulate onsite live proctoring, test performance should not be affected by proctor modality. To date, this supposition has been verified by many studies exploring the comparability of test scores by proctor mode, but there has been a dearth of studies exploring the predictive validity of test scores by proctor modality. This study confirms the findings from the comparability studies and provides the needed findings about the predictive validity of test scores by proctor mode. Specifically, this study reveals no significant difference between entrance test scores proctored onsite live versus remote live; nor were there any significant differences in course performance by students placed in first-year English based on the entrance test scores proctored onsite live or remote live. While these findings are promising, more research is needed, especially in terms of the predictive validity since it relates to the most germane and consequential outcome.

Author disclosures

There are no conflicts of interest.

The authors had the following roles in the research and writing:

Cindy James: conceptualization, data curation, formal analysis, investigation, methodology, validation, writing; Veronika Larsen: validation, writing.

ORCID iD

Cindy James https://orcid.org/0009-0004-7748-9286

References

Akaaboune, O., Blix, L. H., Carrington, L. G., & Henderson, C. D. (2022). Accountability in Distance Learning: The Effect of Remote Proctoring on Performance in Online

Accounting Courses. *Journal of Emerging Technologies in Accounting*, *19*(1), 121–131. https://doi.org/10.2308/JETA-2020-040

Alessio, H. M., Malay, N., Maurer, K., Bailer, A. J., & Rubin, B. (2017). Examining the Effect of Proctoring on Online Test Scores. *Online Learning*, *21*(1), 146–161. https://doi.org/10.24059/olj.v21i1.885

Alessio, H. M., Malay, N., Maurer, K., Bailer, A. J., & Rubin, B. (2018). Interaction of Proctoring and Student Major on Online Test Performance. *International Review of Research in Open and Distributed Learning*, *19*(5). https://doi.org/https://doi.org/10.19173/irrodl.v19i5.3698

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for Educational and Psychological Testing*. American Educational Research Association.

Andreou, V., Peters, S., Eggermont, J., Wens, J., & Schoenmakers, B. (2021). Remote versus on-site proctored exam: comparing student results in a cross-sectional study. *BMC Medical Education*, *21*(1), 624. https://doi.org/10.1186/s12909-021-03068-x

Arnold, I. J. M. (2022). Online proctored assessment during COVID-19: Has cheating increased? *Journal of Economic Education*, *53*(4), 277–295. https://doi.org/10.1080/00220485.2022.2111384

Castellano, K. E., Johnson, M. S., & Lawless, R. (2024). Are Large-Scale Test Scores Comparable for At-Home Versus Test Center Testing? *Applied Psychological Measurement*, *48*(6), 295–299. https://doi.org/10.1177/01466216241253795

College Board. (2017). *Next-Generation Test Specifications: Version 2.0*. https://accuplacer.collegeboard.org/accuplacer/pdf/next-generation-test-specifications-manual.pdf

College Board. (2018a). *ACCUPLACER Program Manual*. https://accuplacer.collegeboard.org/accuplacer/pdf/accuplacer-program-manual.pdf

College Board. (2018b). *ACCUPLACER User's Guide*. https://accuplacer.collegeboard.org/accuplacer/pdf/accuplacer-user-guide.pdf

Daffin Jr., L. W., & Jones, A. A. (2018). Comparing Student Performance on Proctored and Non-Proctored Exams in Online Psychology Courses. *Online Learning*, *22*(1), 131–145. https://doi.org/10.24059/olj.v22i1.1079

Dendir, S., & Maxwell, R. S. (2020). Cheating in online courses: Evidence from online proctoring. *Computers in Human Behavior Reports*, *2*, 100033. https://doi.org/https://doi.org/10.1016/j.chbr.2020.100033

Emeka, C., West, M., Zilles, C., & Silva, M. (2024). A Comparison of Proctoring Regimens for Computer-Based Computer Science Exams. *Proceedings of the 2024 on Innovation and Technology in Computer Science Education V. 1*, 185–191. https://doi.org/10.1145/3649217.3653536

Goedl, P. A., & Malla, G. B. (2020). A Study of Grade Equivalency between Proctored and Unproctored Exams in Distance Education. *American Journal of Distance Education*, *34*(4), 280–289. https://doi.org/10.1080/08923647.2020.1796376

Gulati, R., West, M., Zilles, C., & Silva, M. (2024). Comparing the Security of Three Proctoring Regimens for Bring-Your-Own-Device Exams. *Proceedings of the 55th ACM Technical Symposium on Computer Science Education V. 1*, 429–435. https://doi.org/10.1145/3626252.3630809

Hall, E. A., Spivey, C., Kendrex, H., & Havrda, D. E. (2021). Effects of Remote Proctoring on Composite Examination Performance Among Doctor of Pharmacy Students. *American Journal of Pharmaceutical Education*, *85*(8), 8410. https://doi.org/https://doi.org/10.5688/ajpe8410

Hample, R. (2022). Proctored Versus Unproctored Math Placement Tests: Does It Matter? In *ProQuest Dissertations and Theses*. http://dx.doi.org/10.34944/dspace/8307

Han, S., Nikou, S., & Yilma Ayele, W. (2024). Digital proctoring in higher education: a systematic literature review. *International Journal of Educational Management*, *38*(1), 265–285. https://doi.org/10.1108/IJEM-12-2022-0522

Harmon, O. R., Lambrinos, J., & Tomolonis, P. (2017, October 22). *Proctored in Person or via Webcam: Is There a Difference?* http://dx.doi.org/10.2139/ssrn.3057006

Hollister, K. K., & Berenson, M. L. (2009). Proctored Versus Unproctored Online Exams: Studying the Impact of Exam Environment on Student Performance. *Decision Sciences Journal of Innovative Education*, *7*(1), 271–294. https://doi.org/10.1111/j.1540-4609.2008.00220.x

Hope, D., Davids, V., Bollington, L., & Maxwell, S. (2021). Candidates undertaking (invigilated) assessment online show no differences in performance compared to those undertaking assessment offline. *Medical Teacher*, *43*(6), 646–650. https://doi.org/10.1080/0142159X.2021.1887467

Howard, D. (2020). Comparison of Exam Scores and Time Taken on Exams between Proctored On-Campus and Unproctored Online Students. *Online Learning*, *24*(4), 204228. https://doi.org/10.24059/olj.v24i4.2148

Hylton, K., Levy, Y., & Dringus, L. P. (2016). Utilizing webcam-based proctoring to deter misconduct in online exams. *Computers & Education*, *92–93*, 53–63. https://doi.org/10.1016/j.compedu.2015.10.002

Isbell, D. R., & Kremmel, B. (2020). Test Review: Current options in at-home language proficiency tests for making high-stakes decisions. *Language Testing*, *37*(4), 600–619. https://doi.org/10.1177/0265532220943483

Isbell, D. R., Kremmel, B., & and Kim, J. (2023). Remote Proctoring in Language Testing: Implications for Fairness and Justice. *Language Assessment Quarterly*, *20*(4–5), 469–487. https://doi.org/10.1080/15434303.2023.2288251

Jaap, A., Dewar, A., Duncan, C., Fairhurst, K., Hope, D., & Kluth, D. (2021). Effect of remote online exam delivery on student experience and performance in applied knowledge tests. *BMC Medical Education*, *21*(1), 86. https://doi.org/10.1186/s12909-021-02521-1

James, C. L. (2022). Comparing Performance on Entry Assessments by Proctoring Modality during the COVID-19 Pandemic. *Journal of the National College Testing Association*, *5*(1), 23–39.

https://ncta.memberclicks.net/assets/docs/JNCTA/2022%20Vol%205%20Issue%201.pdf

Kharbat, F. F., & Daabes, A. S. A. (2021). E-Proctored Exams during the COVID-19 Pandemic: A Close Understanding. *Education and Information Technologies*, *26*(6), 6589–6605. https://doi.org/10.1007/s10639-021-10458-7

Kuleva, M., & Miladinov, O. (2024). Exploring the Efficacy of Online Proctoring in Online Examinations: A Comprehensive Review. *Proceedings of the International Scientific and Practical Conference*, 192–196. https://doi.org/10.17770/etr2024vol2.8058

Ladyshewsky, R. K. (2015). Post-graduate student performance in 'supervised in-class' vs. 'unsupervised online' multiple choice tests: implications for cheating and test security. *Assessment & Evaluation in Higher Education*, *40*(7), 883–897. https://doi.org/10.1080/02602938.2014.956683

National College Testing Association. (2025). *Test Centre Certification*. https://www.ncta-testing.org/test-center-certification

Ockey, G. J. (2021). An Overview of COVID-19's Impact on English Language University Admissions and Placement Tests. *Language Assessment Quarterly*, *18*(1), 1–5. https://doi.org/10.1080/15434303.2020.1866576

Oeding, J., Gunn, T., & Seitz, J. (2024). The Mixed-Bag Impact of Online Proctoring Software in Undergraduate Courses. *Open Praxis*. https://doi.org/10.55982/openpraxis.16.1.585

Öncül, B. (2021). Dealing with Cheating in Online Exams: A Systematic Review of Proctored and Non-Proctored Exams. *International Technology and Education Journal*, *5*(2), 45–54. https://dergipark.org.tr/en/pub/itej/issue/75188/1233307

Ong, T. Q., Krumm, B., Wells, M., Read, S., Harris, L., Altomare, A., & Paniagua, M. (2024). A Comparison of Remote vs In-Person Proctored In-Training Examination Administration for Internal Medicine. *Academic Medicine*, *99*(7), 778–783. https://doi.org/10.1097/ACM.0000000000005651

Reisenwitz, T. H. (2020). Examining the Necessity of Proctoring Online Exams. *Journal of Higher Education Theory and Practice*, *20*(1). https://doi.org/10.33423/jhetp.v20i1.2782

Shaushenova, A., Zulpykhar, Z., Zhumasseitova, S., Ongarbayeva, M., Akhmetzhanova, S., Mutalova, Z., Niyazbekova, S., & Zueva, A. (2021). The Influence of the proctoring systems on the results of online tests in the conditions of distance learning. *Ad Alta: Journal of Interdisciplinary Research*, *11*(2), 250–256. https://www.magnanimitas.cz/ADALTA/1102/papers/A_shaushenova.pdf

Speth, J., & Burr, J. (2025). Impact of proctored examinations on students' performance in an online undergraduate pharmacology course. *Pharmacy Education*, *25*(1), 129–135. https://doi.org/10.46542/pe.2025.251.129135

Steger, D., Schroeders, U., & Gnambs, T. (2020). A meta-analysis of test scores in proctored and unproctored ability assessments. *European Journal of Psychological Assessment*, *36*(1), 174–184. https://doi.org/10.1027/1015-5759/a000494

Thompson Rivers University. (2017). *Grading Systems*. https://www.tru.ca/__shared/assets/ED_03-5_Grading_Systems__2017-11-26_5647.pdf

Thompson Rivers University. (2024). *TRU Factbook Fiscal Year 2023-24*. https://www.tru.ca/__shared/assets/Factbook_2023-2459748.pdf

Valkova, R. (2024). Remote Proctoring in language assessment: Exploring the impact on test-takers' scores and perceptions. *Studies in Language Assessment*, 141–198. https://doi.org/10.58379/FCZM2855

Vazquez, J. J., Chiang, E. P., & Sarmiento-Barbieri, I. (2021). Can we stay one step ahead of cheaters? A field experiment in proctoring online open book exams. *Journal of Behavioral and Experimental Economics*, *90*, 101653. https://doi.org/https://doi.org/10.1016/j.socec.2020.101653

Woldeab, D., & Brothen, T. (2019). 21st Century Assessment: Online Proctoring, Test Anxiety, and Student Performance. *International Journal of E-Learning &*

Distance Education, *34*(1). https://www.ijede.ca/index.php/jde/article/view/1106

Zhang, N., Larose, J., & Franklin, M. (2024). Effect of unproctored versus proctored examinations on student performance and long-term retention of knowledge. *Journal of Chiropractic Education (Allen Press Publishing Services Inc.)*, *38*(2), 114–119. https://doi.org/10.7899/JCE-23-16

Zumbo, B. (2021). *A Novel Multimethod Approach to Investigate Whether Tests Delivered at a Test Centre are Concordant with those Delivered Remotely Online: An Investigation of the Concordance of the CAEL.* https://doi.org/10.14288/1.0400581