



Article

E-Assessment in E-Learning Degrees: Comparison vs. Face-to-Face Assessment through Perceived Stress and Academic Performance in a Longitudinal Study

Roberto Sánchez-Cabrero * , Javier Casado-Pérez, Amaya Arigita-García, Elena Zubiaurre-Ibáñez ,
David Gil-Pareja and Ana Sánchez-Rico

Department of Education, Faculty of Social Sciences and Applied Languages, Villanueva de la Cañada, Alfonso X the Wise University, 28691 Madrid, Spain; jcasaper@uax.es (J.C.-P.); aarigita@uax.es (A.A.-G.); ezubiiba@uax.es (E.Z.-I.); dgilpar@uax.es (D.G.-P.); asancric@uax.es (A.S.-R.)

* Correspondence: rcabrero@uax.es

Featured Application: This study aimed to compare online exams with their classroom counterparts by comparing academic achievement and student perception of difficulty, stress and fairness. This assessment is prerequisite if university education is to be less dependent on the face-to-face experience throughout the educational process.



Citation: Sánchez-Cabrero, R.; Casado-Pérez, J.; Arigita-García, A.; Zubiaurre-Ibáñez, E.; Gil-Pareja, D.; Sánchez-Rico, A. E-Assessment in E-Learning Degrees: Comparison vs. Face-to-Face Assessment through Perceived Stress and Academic Performance in a Longitudinal Study. *Appl. Sci.* **2021**, *11*, 7664. <https://doi.org/10.3390/app11167664>

Academic Editor:
Paz Prendes-Espinosa

Received: 21 July 2021

Accepted: 19 August 2021

Published: 20 August 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: The COVID-19 pandemic has become both a challenge and an opportunity to implement certain changes in the world of education. One of the most important differences has been online evaluation, which had, until now, been marginal in most prestigious universities. This study compared the academic achievement of the last cohort that performed classroom assessment and the first group that was graded for an official degree using synchronous online evaluation. Other variables measured were the self-assessment of students in this second group, in order to understand how it affected their perception of the process using three different indicators: stress, difficulty, and fairness. Nine hundred and nineteen students participated in the study. The results indicate that online assessment resulted in grades that were 10% higher while enjoying the same degree of validity and reliability. In addition, stress and difficulty levels were also in line with the on-site experience, as was the perception that the results were fair. The results allow us to conclude that online evaluation, when proctored, provides the same guarantees as desktop exams, with the added bonus of certain advantages which strongly support their continued use, especially in degrees with many students who may come from many different locations.

Keywords: evaluation; e-assessment; e-learning; academic performance; COVID-19; exams

1. Introduction

In recent decades, there has been a great debate on the best methods of student assessment [1–3]. Studies have indicated that the improvement of student learning is linked to an optimal evaluation of their academic progress, however, the most appropriate way to achieve a quality assessment remains controversial. [4,5]. Nowadays, the consensus is that it should, when planned correctly, be a key stepping stone in the teaching-learning process and in improving the quality of education systems [6–9]. This has led to the central role that evaluation has in educational policy. Regulations consider evaluation to be an element of the curriculum, consistent with the other aspects included. Educational reforms in recent decades in Spain have never failed to underwrite its importance, progressively substituting previous forms with new systems which reflect the changes pervading the new laws. In international reports prepared by PISA, promoted by the OECD (Organisation for Economic Co-operation and Development), the reports coordinated by the IEA (International Association for the Evaluation of Educational Achievement) or those

created by the OREALC/UNESCO (Oficina regional de Educación para América Latina in Spanish/United Nations Educational, Scientific and Cultural Organisation) have provided countless indicators aimed at comparing and improving evaluation systems [10,11].

This substantial role is due to the fact that exam-based assessment is the way in which teachers can gather information on the process of teaching-learning, determining whether goals are being met and competencies acquired by students [12–14]. Evaluations are not merely grading tools, but also formative in nature [15–17]. Determining content acquisition is not the only use they have, for they can also be diagnostic regarding the needs students have. Evaluation is a data-gathering mechanism that can help ascertain student progress and effect positive change on their education while serving as a monitoring tool of the complete educational process [18–20].

This entails that, apart from an individual assessment, the assessment also allows the identification of broader problems that can ultimately lead to the improvement of education systems. Evaluation has, thus, extended beyond learning to be used for teaching and the operation of education centers [21]. All the dimensions of the educational process are now subject to assessment [22–24]. The ensuing results may result in significant changes to systems. Evaluation, far from being consigned to classrooms, has become fundamentally strategic in determining the nature of any potential education reforms.

This key role of evaluation and the impact it can play in educational transformation is not without controversy. It has been questioned whether it does, in fact, measure learning effectively [11,25]. This is a two-pronged question because it sheds light on student achievement but must remain consistent with the methodology used. In addition, the evaluation method must be formally coherent with the pedagogical objectives, warranting assessments that are designed to elucidate certain selected aspects [9,22,24,26]. A second debate revolves around the presumed objectivity of standardized tests such as PISA [25]. Finally, doubt has been cast on the capacity that assessment results have to actually improve the overall process, prompting many proposals that could streamline the connection between objectives and results, and hence to educational policy [13,27,28]. As Biggs points out in his seminal work on constructive alignment [29], there is overall a sharp tension between learning objectives and evaluation methods, two poles with distinct priorities and philosophies regarding how results should affect education reform.

The ever-increasing presence of digital technology in education has muddled these debates even further complicated [30–33]. Its growing importance in recent years has compelled educators to adapt to a medium that is increasingly present at all levels. ICTs have permeated different spheres of education, including assessment, which, in turn, have had to forego in many circumstances various traditional classroom techniques and instruments [24,34].

Being able to do exams online has opened many possibilities, allowing for greater flexibility and de-centering the whole educational process [32,35,36]. Physical constraints disappear when evaluation can take place remotely from home [37,38]. The new changes, however, have generated new problems related to the capacity that testing has of reflecting student learning, and thus of being a diagnostic tool of the teaching-learning process [38,39].

Some studies have been tentatively exploring this issue [35,40–42] and the results have shown that online assessment is possible according to current educational quality standards. However, more data needs to emerge on the characteristics of online assessment and its effectiveness in providing information on student learning. The widespread application of online assessments in most universities came with the outbreak of the COVID-19 pandemic. The online evaluation had been very isolated until that moment, but the crisis forced important steps to be taken and online assessment became the only method for students and degrees [43]. These advances must take place in combination with further consideration on the role of evaluation and how it shapes education reform, in order to ascertain whether online testing can also play this role [44–46].

In the Spanish university system, few institutions had even considered online exam-based evaluation prior to the COVID-19 pandemic, a position that closely reflected that of all major higher education organizations across the world [47]. Spring 2020, however, forced everyone to embrace remote learning almost immediately, accelerating a process that was perhaps foreseen, but not in the near future. This prompted a deluge of scientific research on this situation [48–50], but most of it focused on other educational aspects, or used provisional assessment data, without providing any comparative data from analogous evaluation conditions. According to González-González et al. [51], most countries are migrating education to an online environment: around 89% of the World student population. The greatest caveat is the quality of evaluation and controlling for fraud, warranting electronic supervision tools that can offset these dangers, at least to a classroom standard [52].

The exceptional situation generated by the COVID-19 pandemic led the Alfonso X the Wise University to focus strongly on online exam-based assessment, making use of the *Respondus Monitor* and *LockDown Browser*, both applications owned by *Respondus Inc.*, Redmond, WA, USA [53]. This decision enabled the university to keep the evaluation calendar in all university degrees which had previously required on-site assessment.

According to the authors, the combination of both *Respondus Monitor* and *LockDown Browser* enables students to take exams remotely while guaranteeing the integrity of the process [53]. The remote proctoring system has generated great controversy since its use has been extended in the university sector during the COVID-19 pandemic. For example, Silverman et al. [54] show in their study a summary of the arguments against institutional adoption of remote proctoring services with a focus on equity, an account of the decision to avoid remote proctoring on the University of Michigan. This software, once installed, activates the student's webcam and executes certain prior steps: a short recording of the user, a picture of the student and his ID, and a recording of the student's environment [53]. After this, the *Browser* blocks the window in full screen, impeding toggling windows, copying, printing, or accessing other programs during the exam. Exam supervision is carried out automatically by *Review Priority*, a third program that *Respondus monitor* delegates this task to, and which is accessible using the Dashboard on *Browser* [53]. *Priority*, using the feed from the student's cameras during the exam, flags any anomalies in the video, and an overall proctoring result is generated, which the teacher can review later using the recorded video. However, reviewing the video is not at all necessary, except for notice of irregular behavior on the part of the software.

Instructors can review other meta-information of the exam live, such as the time elapsed, student name and grade, proctoring result, as well as how many students have finished or are still doing it. *Respondus Monitor*, hence, does not determine whether students cheat, but rather leaves this decision to the teacher. However, according to recent studies, it is a very effective deterrent [55–57].

This study had two research objectives which will use the experience at the *Alfonso X the Wise University*. First of all, to compare the academic achievement in the three required courses of the *Master's Degree in Secondary Teacher Training*, professional teaching qualification—for the last cohort that had an on-site assessment, as well as the first one to do so remotely. This analysis will factor in other variables such as gender, age, and academic discipline of students. A second objective is to understand the significance of correlations between online evaluation and student perception of stress, difficulty, and satisfaction, again factoring in the academic discipline. The analysis aims to reveal whether grades were different using the two methods, as well as the student well-being during, and opinion of, remote assessment.

2. Materials and Methods

2.1. Participants

The study sample includes 919 participants (521 women and 398 men) who were enrolled in the mentioned *Master's Degree in Secondary Teacher Training* of the *Alfonso X*

the Wise University. This degree is a Bologna master's program and it is necessary to become a secondary school teacher in the Spanish educational system. The degree is the second option for many students, who decide to become teachers at a later age. Many students choose this degree years after practicing another profession, so they are usually students who are mostly in their thirties. The mean age of the sample was 34.91 years (34.44 women, 35.53 men) with a standard deviation of 7.68 (women 7.5, men 7.88). The sample was grouped into academic disciplines which are determined by the student's prior degree: Biology and Geology, Economy, Technology, Physics and Chemistry, Geography and History, English, Spanish Language and Literature, and Mathematics. These groups complied with the criteria of being evaluated either in the school year 2019–2020 and 2020–2021 and providing other factor data specified. As it is an online degree, where only the final assessment was face-to-face, all participants had sufficient ICT skills and access to broadband internet from the beginning of the course. Therefore, their digital skills and the digital divide are not variables that can influence the results of the study. Table 1 provides numbers for the variables used in the study.

Table 1. Sample description.

Type of Assessment Used	Frequency	Percentage	Mean Age	Standard Deviation
On-site (year 2019–2020)	402	43.7	34.74	7.83
Online (year 2020–2021)	517	56.3	35.04	7.57
Academic Discipline	Frequency	Percentage	Mean Age	Standard Deviation
Biology and Geology	184	20.0	33.01	7.24
Economy	178	19.4	36.08	7.54
Technology	265	28.8	37.98	7.14
Physics and Chemistry	69	7.5	32.90	6.49
Geography and History	77	8.4	32.70	7.91
English	51	5.5	31.96	8.06
Spanish Language and Literature	29	3.2	28.97	5.63
Mathematics	66	7.2	34.26	7.59
Gender	Frequency	Percentage	Mean Age	Standard Deviation
Female	521	56.7	34.44	7.50
Male	398	43.3	35.53	7.88
Sample Total	919	100.0	34.91	7.68

Table 1 clearly shows that the population of students of the Secondary Education Master at the *Alfonso X the Wise University*, according to the sample obtained, are mostly from the 2020–2021 cohort (56.3%), and their most frequent previous academic area is Technology (28.8%).

2.2. Assessment Methods and Variables

The data in this study was obtained using different sources. Exam results were obtained from the standard February exams of the three mandatory courses within the Master's program: *Education and Social and Family Environment*, *Learning and development of personality*, and *Educational Processes and Contexts* for the 2019–2020 and 2020–2021 editions, which enjoyed high reliability based on a Cronbach's alpha coefficient of $\alpha = 0.801$. Further information was collected using a simple questionnaire that included basic data and a consent form. Finally, student surveys included three self-assessment questions which were appended to the three exams of the 2020–2021 cohort, which used a single-response Likert scale.

All assessment instruments were validated by an external committee that oversaw the scientific and ethical issues. Their role was key in approving and monitoring the experiment. In order to be able to be part of the sample, participants needed to provide informed consent in writing, in compliance with the Helsinki Declaration on the ethical principles of human experimentation [58].

The variables used were attributes, academic performance, and self-assessment. The four attributes were: gender (with two options allowed: female and male), age (a discrete quantitative variable), cohort (two options, 2019–2020 and 2020–2021, i.e., on-site vs remote), and academic discipline, based on the degree obtained prior to the Master's (with eight nominal possibilities: Biology and Geology, Economy, Technology, Physics and Chemistry, Geography and History, English, Spanish Language and Literature and Mathematics).

The academic performance variables were three: grades on a 0–10 number scale obtained in the final exam for the three mandatory courses within the Master's program: *Education and Social and Family Environment*, *Learning and development of personality*, and *Educational Processes and Contexts*. The exams were structurally the same in both calls and the duration was similar. In the years 2019–2020 and 2020–2021 the same contents were evaluated. The exams were carried out by the same teachers in both calls as well. Therefore, it is possible to affirm that the main differences in the results are due to the change in the evaluation format.

Finally, self-assessment variables were nine items on a Likert scale. The first three were related to the perceived stress levels, the following three to the perceived difficulty of the exam, and the final three were aimed at understanding overall student satisfaction with the online grading experience. Table 2 details the three items described.

Table 2. Self-assessment items regarding the online evaluation experience.

Items	Number of Options
How would you define your stress level right before taking the online exam?	5
How would you define your stress level while taking the online exam?	5
Comparing this remote exam to previous on-site experiences, what statement would better reflect your opinion?	3
Regarding the grade a student gets in an exam, overall:	3
Do you consider that remote exams enable the student to prove his learning?	3
Do you consider that the competences acquired are reflected in an online exam?	5
How would you rate your experience taking, for the first time, an online exam in the university?	5
In the second semester, would you like to repeat the experience of taking exams remotely?	5
Do you think that online evaluation has a future, or is it a temporary situation?	5

2.3. Experiment Design

A descriptive analysis has been carried out, using sequential correlation and comparison, of the student sample. The first step was a descriptive statistical analysis, using frequency distribution of nominal and ordinal variables, as well as statistical indicators such as the average and the standard deviation of the age and grades quantitative datasets. A correlational analysis was carried out using Pearson's correlation coefficient (r) for the quantitative variables, such as grades, and Spearman's correlation (ρ) when ordinal vari-

ables are used, allowing for the combination of quantitative and qualitative data, such as the self-assessment variables. Inferential analyses used Student's t-distribution on independent (or unpaired) samples in order to compare the results of the online vs on-site cohorts. Finally, the significance of confidence intervals—99% (α : 0.01) and 95% (α : 0.05)—were taken into account.

3. Results

The assessments took place in a timely manner, with no issues to report. All students were successfully examined, so there is no possibility of skewed results. The descriptive results of the study can be seen in Table 3. It shows the averages and standard deviation of the three courses for both cohorts, as well as breaking down the results by gender and academic discipline.

Table 3. Student grade average by course and standard deviation according to different variables.

Type of Assessment Used	AyDP (M/SD)	EyESyF (M/SD)	PyCE (M/SD)
On-site (year 2019–2020)	7.03 (1.76)	7.54 (1.75)	6.60 (1.31)
Online (year 2020–2021)	8.27 (1.52)	9.02 (1.14)	7.95 (1.22)
Gender	AyDP (M/SD)	EyESyF (M/SD)	PyCE (M/SD)
Female	7.82 (1.69)	8.44 (1.62)	7.44 (1.41)
Male	7.61 (1.80)	8.29 (1.62)	7.26 (1.44)
Academic Discipline	AyDP (M/SD)	EyESyF (M/SD)	PyCE (M/SD)
Biology and Geology	8.24 (1.59)	8.77 (1.40)	7.67 (1.35)
Economy	7.42 (1.76)	7.93 (1.83)	7.17 (1.45)
Technology	7.66 (1.72)	8.56 (1.40)	7.40 (1.37)
Physics and Chemistry	7.91 (1.64)	8.42 (1.58)	7.37 (1.34)
Geography and History	7.27 (1.79)	7.83 (1.70)	7.03 (1.51)
English	7.95 (1.69)	8.23 (1.83)	7.35 (1.51)
Spanish Language and Literature	7.59 (1.99)	7.80 (1.92)	6.76 (1.67)
Mathematics	7.60 (1.88)	8.64 (1.55)	7.46 (1.47)
Sample Total	7.73 (1.74)	8.37 (1.62)	7.36 (1.43)

AyDP = Learning and development of personality; EyESyF = Education and Social and Family Environment; PyCE = Educational Processes and Contexts.

Table 3 reveals significant differences in the grades of both cohorts, which are systematically better for the online evaluation. There are also slight differences between gender. Finally, the students of the Biology and Geology degrees systematically fared better than the rest.

Table 4 below shows the frequency distribution of the different answers possible in the self-assessment questions designed as a Likert scale, which were filled by the 517 participants of the online exam cohort of 2020–2021.

The results observed in Table 4 indicate that student stress perception declines rapidly once an exam has begun. Ultimately, stress levels follow a similar pattern to that of on-site evaluation; 54.2% of respondents indicate it is the same. Regarding grades, students do not perceive much difference between online and on-site assessment, considering both to be fair. Most students are satisfied with the remote exam experience and would like to repeat it in the second semester (63.1%). Finally, most students believe that online evaluation will become common in the future (54.4%). In order to respond to the research objectives, these perceptions revealed in the descriptive analysis of results need to be contrasted using correlational and inferential statistical analysis.

Table 4. Frequency table of the self-assessment answers.

How would you define your stress level right before taking the online exam?	Frequency	Percentage
No stress	13	2.5
A little stress	95	18.4
A moderate level of stress	142	27.5
Pretty stressed	198	38.3
Very stressed	69	13.3
How would you define your stress level while taking the online exam?	Frequency	Percentage
No stress	66	12.8
A little stress	205	39.7
A moderate level of stress	179	34.6
Pretty stressed	51	9.9
Very stressed	16	3.1
Comparing this remote exam to previous on-site experiences, what statement would better reflect your opinion?	Frequency	Percentage
On-site evaluation is normally more stressful	183	35.4
Both types are similarly stressful	280	54.2
A remote exam will make me more stressed than an on-site one next time	54	10.4
Regarding the grade a student gets in an exam, overall:	Frequency	Percentage
On-site exams are easier for students	25	4.8
There are no differences regarding the grade a student may obtain	416	80.5
Remote exams are easier for students	76	14.7
Do you consider that remote exams enable the student to prove his learning?	Frequency	Percentage
To a greater extent than on-site exams	27	5.2
Both types provide the same opportunity	455	88.0
To a lesser extent than on-site exams	35	6.8
Do you consider that the competences acquired are reflected in an online exam?	Frequency	Percentage
Not at all	15	2.9
Less than on-site exams	36	7.0
The same as on-site exams	394	76.2
More than on-site exams	10	1.9
Totally	62	12.0
How would you rate your experience taking, for the first time, an online exam in the university?	Frequency	Percentage
I was very disappointed with the experience	2	0.4
I was slightly disappointed with the experience	20	3.9
Neither disappointed nor satisfied	103	19.9
I was moderately satisfied with the experience	172	33.3
I was very satisfied with the experience	220	42.6

Table 4. Cont.

In the second semester, would you like to repeat the experience of taking exams remotely?	Frequency	Percentage
I much prefer on-site evaluation	3	0.6
I much prefer on-site evaluation, but I would not mind taking them online again	27	5.2
I don't care, both types are the same	47	9.1
I prefer remote exams, but I would not mind taking them on-site	114	22.1
I much prefer remote evaluation	326	63.1
Do you think that online evaluation has a future, or is it a temporary situation?	Frequency	Percentage
It is an extravagance which will be abandoned in the future	1	0.2
It will evolve into new and more complex forms of evaluation	25	4.8
It will adopt a supporting role to the onsite method with time	99	19.1
It will become more frequent, even common, as a type of evaluation	281	54.4
It will become unavoidable in the future	111	21.5
Sample total	517	100.0

The first one of these goals, comparing the academic performance between the on-site and the online cohort, and how factors such as age, gender, and academic discipline may influence them is the reason for Table 5, which shows the results of Student's t-distribution for independent samples, comparing both cohorts.

Table 5. Comparison of the online vs on-site evaluation using Student's t-distribution for two unpaired samples.

Total Sample	<i>t</i>	Significance
Grade in <i>Learning and development of personality</i>	−11.202	0.000
Grade in <i>Education and Social and Family Environment</i>	−14.742	0.000
Grade in <i>Educational Processes and Contexts</i>	−15.970	0.000
Age	−0.578	0.563
Gender: Female	<i>t</i>	Significance
Grade in <i>Learning and development of personality</i>	−8.304	0.000
Grade in <i>Education and Social and Family Environment</i>	−10.444	0.000
Grade in <i>Educational Processes and Contexts</i>	−11.719	0.000
Age	0.462	0.645
Gender: Male	<i>t</i>	Significance
Grade in <i>Learning and development of personality</i>	−7.418	0.000
Grade in <i>Education and Social and Family Environment</i>	−10.369	0.000
Grade in <i>Educational Processes and Contexts</i>	−10.775	0.000
Age	−1.490	0.137
Academic Discipline: Biology and Geology	<i>t</i>	Significance
Grade in <i>Learning and development of personality</i>	−5.940	0.000
Grade in <i>Education and Social and Family Environment</i>	−6.766	0.000
Grade in <i>Educational Processes and Contexts</i>	−9.764	0.000
Age	2.009	0.046

Table 5. Cont.

Academic Discipline: Economy	<i>t</i>	Significance
Grade in <i>Learning and development of personality</i>	−3.812	0.000
Grade in <i>Education and Social and Family Environment</i>	−7.917	0.000
Grade in <i>Educational Processes and Contexts</i>	−6.648	0.000
Age	−0.096	0.924
Academic Discipline: Technology	<i>t</i>	Significance
Grade in <i>Learning and development of personality</i>	−6.145	0.000
Grade in <i>Education and Social and Family Environment</i>	−6.459	0.000
Grade in <i>Educational Processes and Contexts</i>	−4.659	0.000
Age	−0.883	0.378
Academic Discipline: Physics and Chemistry	<i>t</i>	Significance
Grade in <i>Learning and development of personality</i>	−4.604	0.000
Grade in <i>Education and Social and Family Environment</i>	−2.668	0.010
Grade in <i>Educational Processes and Contexts</i>	−5.330	0.000
Age	0.065	0.949
Academic Discipline: Geography and History	<i>t</i>	Significance
Grade in <i>Learning and development of personality</i>	−2.527	0.014
Grade in <i>Education and Social and Family Environment</i>	−3.812	0.000
Grade in <i>Educational Processes and Contexts</i>	−5.468	0.000
Age	−1.011	0.315
Academic Discipline: English	<i>t</i>	Significance
Grade in <i>Learning and development of personality</i>	−2.053	0.045
Grade in <i>Education and Social and Family Environment</i>	−4.432	0.000
Grade in <i>Educational Processes and Contexts</i>	−5.463	0.000
Age	0.134	0.894
Academic Discipline: Spanish Language and Literature	<i>t</i>	Significance
Grade in <i>Learning and development of personality</i>	−3.084	0.007
Grade in <i>Education and Social and Family Environment</i>	−4.059	0.001
Grade in <i>Educational Processes and Contexts</i>	−4.937	0.000
Age	−0.487	0.630
Academic Discipline: Mathematics	<i>t</i>	Significance
Grade in <i>Learning and development of personality</i>	−3.298	0.002
Grade in <i>Education and Social and Family Environment</i>	−4.676	0.000
Grade in <i>Educational Processes and Contexts</i>	−5.698	0.000
Age	0.592	0.556

The results in Table 5 indicate that the variable cohort, which distinguishes students that were evaluated on-site and remotely, is significant throughout, both independently and in relation to all the gender and academic discipline variables. This confirms the descriptive analysis of Table 3. Age, however, appears to bear no effect in relation to any other variable, save for Biology and Geology, which only reinforces that this variable did not affect cohort results (i.e., age had no impact on remote learning).

The second research objective, tackling the correlation between the online cohort and their perception of stress, difficulty, and satisfaction, as well as academic achievement, is the topic for Table 6. This table shows below the correlations between the different grades in the three courses using Pearson's correlation coefficient (r) as a form of statistical contrast, given that all three variables were quantitative.

Table 6. Correlation tables between grades using Pearson’s correlation coefficient (r).

Course Title	AyDP	EyESyF	PyCE
<i>Learning and development of personality</i>	1	0.564 **	0.603 **
<i>Education and Social and Family Environment</i>		1	0.571 **
<i>Educational Processes and Contexts</i>			1

** Correlation is significant at the 0.01 level.

The results in Table 6 reveal clear correlations which are highly significant ($p = 0.000$) between all three courses graded, indicating a strong correlation between them. This enables the inclusion of a new variable, academic performance, which is the average of all three grades since independently the three grades would be completely redundant. Table 7 shows the resulting correlation between the self-assessment ordinal variables and academic performance, using Spearman’s correlation coefficient (rho)

Table 7. Correlations using Spearman’s correlation coefficient as a contrasting statistic.

	St 1	St 2	St 3	D 1	D 2	D 3	Sa 1	Sa 2	Sa 3
Academic performance	−0.082	−0.192 **	−0.049	−0.031	−0.016	−0.034	−0.195 **	0.082	0.009
Stress 1	1	0.561 **	0.270 **	−0.022	−0.011	−0.045	−0.224 **	−0.087 *	0.010
Stress 2		1	0.246 **	0.017	−0.076	−0.011	−0.239 **	−0.083	0.027
Stress 3			1	−0.242 **	0.232 **	−0.323 **	−0.328 **	−0.389 **	−0.181 **
Difficulty 1				1	−0.198 **	0.158 **	0.131 **	0.167 **	0.094 *
Difficulty 2					1	−0.324 **	−0.228 **	−0.180 **	−0.212 **
Difficulty 3						1	0.251 **	0.231 **	0.206 **
Satisfaction 1							1	0.494 **	0.234 **
Satisfaction 2								1	0.395 **

* Correlation is significant at the 0.05 level/** Correlation is significant at the 0.01 level.

In Table 7, it is remarkable how only the items “How would you define your stress level while taking the online exam?” (Stress 2) and “How would you rate your experience taking, for the first time, an online exam in the university?” (Difficulty 3) reveal correlation—inverse in the first and direct in the second—with academic performance. In addition, the items of each area—Stress, Difficulty, Satisfaction—are highly correlated within each area, while Difficulty and Satisfaction are also highly correlated amongst each other. Finally, Difficulty 3 is highly correlated with all other items, whereas “Comparing this remote exam to previous on-site experiences, what statement would better reflect your opinion?” (Stress 3) also does this, except for, interestingly, academic performance.

4. Discussion

An initial exploration of the results outlined above indicates that remote evaluation generally works well. Apart from excellent instrumental reliability ($\alpha = 0.801$), the fact that there is a strong direct correlation between courses is a powerful argument in favor of it thanks to the stability of the individuals partaking in it. Despite the change in the evaluation method, student performance still constitutes the major determinant of their grade. Even though in general grades were higher after remote evaluation, differences in student performance were maintained.

Notwithstanding the expert committee that approved of the self-assessment items, other aspects appear to support the robustness of their design: strong correlation within each category—stress, difficulty, satisfaction—as well as between the overall opinion of remote exams and the probability of desiring to repeat, or that those that were more satisfied with the experience felt less stress doing it. These correlations indicate that there is a coherence between the results and the responses.

The first research objective was focused on comparing the academic performance in the three compulsory courses of the Secondary Education Master's in both the last on-site evaluation and the first remote one, while also factoring in certain attributes such as age, gender, and academic discipline. It is safe to say that performance improved significantly in the latter cohort, with an increase of more than 10% in average grades. Age, gender, or discipline did apparently affect this improvement in any way. Despite these variables being usually blamed for academic performance [59–61], they did not seem to change at all from one cohort to the next.

These results could be due to the various circumstances which may have affected remote evaluation momentarily, and not because of the type of evaluation per se [41,42,49]. The change from on-site to online was sudden and unexpected due to the arrival of the COVID-19 pandemic, barring any premeditation or preparation [48,50]. Instructors preparing the evaluation items had no experience in designing online exams [22], and there was strong student pressure for fair assessment methods [62,63]. This may have prompted teachers to design exams that were substantially easier, thus attempting to compensate for any possible detriment caused by the sudden shift to a methodology which the instructors were ill-prepared for [50]. In addition, a new and unknown assessment environment was naturally mistrusted [51].

The second research objective sought to compare the correlations in the second cohort with their self-assessment of perceived stress, difficulty, and satisfaction, as well as academic performance. Surprisingly, perceived stress was significantly lower once the online exam had begun. Digital assessment environments are unknown for students, which would warrant a high degree of uncertainty, coupled with potential technical issues which were a looming threat [64]. This could easily have spiraled into greater student insecurity and, hence, stress [42,65]. Our results, however, indicate that once the exam has begun, stress levels are even lower than with on-site exams, and the security the student feels with the environment increases quickly. Once the exam is over, most students consider that there is little difference between on-site and online evaluation, and both prompt the same amount of stress, if not less for remote exams. Both are, in the student's perception, valid and fair assessment tools.

These overly positive indicators explain that most students would like to repeat remote assessment, despite the fact that the remote proctoring method initially generated enormous mistrust among students, as Silverman et al. [54] have shown in their study. Once the initial uncertainty is overcome, the digital environment is reliable and safe for the student. Most students believe that in the future remote exams will be the most common type, yet another dimension in which we become increasingly accustomed to using digital technologies. This is a very optimistic outlook, given that it was the dramatic change that COVID-19 forced upon universities by pushing them into the age of remote learning [48,50]. Student attitude has been welcoming and they have adapted quickly to the new environment, which is daily becoming more usual and satisfactory [66]. This is good news for universities that wish to expand this dimension in the near future.

Perceived difficulty indicators clarify that online exams are significantly easier, which may have facilitated this broad acceptance of the evaluation system described here. Nonetheless, only 14.7% of students were aware of this. Even though a possible influence in the reduction of real difficulty in the exam is not ruled out, perceived difficulty remained unaltered for most students, deeming this influence insufficient to explain the overall results [11,19,48].

Another interesting result is that students who were more stressed during the exam obtained worse results, which is a relation that would appear to be expected [67]. Nonetheless, the stress suffered can be an element that caused the inferior grade, or a consequence of it. Those with higher grades were also most satisfied with the experience, perhaps showing that they were the ones that adapted best to the new methodology. Students who are savvy in digital technologies will tend to be more comfortable in situations like this, which could result in a less stressful experience, and hence greater overall satisfaction [32,49].

5. Conclusions

During 2020 and 2021, the health emergency caused by the global COVID-19 pandemic and academic performance constituted one of the greatest educational challenges in history. Universities and schools could not do on-site exams, which were the only modalities foreseen throughout the world. Even many online degrees in Spain tended to have classroom evaluations, forcing remote learning institutions such as the *Alfonso X the Wise University* to invest heavily in evaluation rooms; the on-site exam was the only reliable way to go. Given the dramatic situation in spring 2020 in Spain, universities opted for a wide range of measures with inconsistent results. Some swapped exams for papers submitted asynchronously, others preferred practical exams conducted synchronously, and yet others merely did the on-site exam remotely. Online exam-based evaluation, however, is qualitatively distinct from on-site exam-based evaluation, and it requires technologies and software which can substitute direct human proctoring while guaranteeing fairness and equality for all students. As a result, many institutions obtained adulterated or disappointing results. The *Alfonso X the Wise University* made use of *Respondus Monitor* to meet this challenge, based on their experience and quality, and this study bears witness to the adequacy of this decision.

The main conclusion of this study is that online assessment, if done in conditions that avoid fraud and that are accessible for students, is as legitimate as on-site assessment. In fact, in given circumstances, it is even preferable—in the case of the *Online Master's Degree in Secondary Teacher Training* in the *Alfonso X the Wise University*, which has many students from all over the country. Online assessment allows many students to take the exams regardless of their location. This conclusion is borne out by the fact that a 10% increase in grades is still consistent with a very stable and reliable evaluation process, both for instructors and students. Grade distribution is based on performance and individual grade differences are in line with the on-site assessment of the different courses.

It can thus be inferred, based on the interindividual consistency of the results and between the different subjects, that the grade improvement is not due to the methodology itself, but rather to the circumstances surrounding the process which led to the creation of exams that were significantly easier. The contents evaluated were the same and the structure of the exam and its duration were similar, as well. For most of the students, the perception of justice and difficulty of the assessment was exactly the same as the past experiences with the face-to-face assessment format. The lack of experience of teachers in creating online exams could explain this, as well as an initial position of mistrust for both students and instructors. These particulars, however, will disappear as this system is repeated in different calls, probably resulting in a diminishing difference that some minor monitoring of exam designs by teachers in upcoming calls would solve.

This pioneering experience of online synchronous evaluation of students results in initial anticipatory stress which sharply descends as the exam begins. The perceived stress is, at least, the same as that of an on-site exam, despite the video recording in their personal space and the intrusion into the privacy of their home, since 54.2% of those surveyed believed it so. Of the other respondents, most believed classroom evaluation was more stressful than remote evaluation, three times more than those who believe the opposite. In addition, student perception of difficulty and fairness indicates that 75% of the sample believe there is no difference between on-site and online.

Finally, the data provided leads to the conclusion that online exams have facilitated the task of grading large quantities of students in a short time. It does this while ensuring the reliability and validity of the evaluation, at least on par with classroom exams, and, in addition, reducing the costs and difficulties associated with student travel. It is therefore foreseeable that this system will easily expand in the future, especially in large online degrees, such as the *Online Master's Degree in Secondary Teacher Training* in the *Alfonso X the Wise University*, where this evaluation has undoubtedly arrived to remain. However, current trends in educational evaluation tend towards the evaluation of significant knowledge and

the development of competencies. This challenge for the university environment indicates that the evaluation processes must still be improved.

Author Contributions: Conceptualization, R.S.-C., A.A.-G., E.Z.-I. and J.C.-P.; methodology, R.S.-C. and J.C.-P.; software, A.A.-G. and J.C.-P.; validation, R.S.-C. and E.Z.-I.; formal analysis, R.S.-C. and E.Z.-I.; investigation, R.S.-C. and E.Z.-I.; resources, D.G.-P., A.S.-R. and A.A.-G.; data curation, R.S.-C.; writing—original draft preparation, R.S.-C. and E.Z.-I.; writing—review and editing, R.S.-C., A.A.-G., D.G.-P., A.S.-R., J.C.-P. and E.Z.-I.; visualization, R.S.-C., D.G.-P. and A.S.-R.; supervision, R.S.-C.; project administration, R.S.-C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Institutional Review Board of Alfonso X The Wise (protocol code 01 in 01/02/21).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Restrictions apply to the availability of these data. Data was obtained from [third party] and are available [from the authors at URL] with the permission of [third party].

Acknowledgments: The authors would like to express their gratitude to the participants of the research and to the managing team of Encuestafacil.com, which lent out their application free of charge so that we could create a questionnaire, and their servers in order to collect some of the data.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Wagner, D.A.; Lockheed, M.; Mullis, I.; Martin, M.O.; Kanjee, A.; Gove, A.; Dowd, A.J. The Debate on Learning Assessments in Developing Countries. *Comp. A J. Comp. Int. Educ.* **2012**, *42*, 509–545. [\[CrossRef\]](#)
2. Mulongo, G.; Amod, Z. Participation in Cross-National Learning Assessments and Impact on Capacity Development: Programmes, Practice, Structures and Teacher Competency. Case Study of Kenya, Tanzania and South Africa. *Eval. Program Plan.* **2017**, *65*, 94–105. [\[CrossRef\]](#)
3. Pisani, L.; Borisova, I.; Dowd, A.J. *International Development and Early Learning Assessment Technical Working Paper*; Save the Children: London, UK, 2015.
4. Wagner, D.A. *Learning as Development: Rethinking International Education in a Changing World*; Routledge: London, UK, 2017; ISBN 978-1-136-29451-8.
5. Lockheed, M. Measuring Progress with Tests of Learning: Pros and Cons for “Cash on Delivery Aid” in Education. *SSRN J.* **2008**, *147*. [\[CrossRef\]](#)
6. Dochy, F.J.R.C.; McDowell, L. Introduction: Assessment as a Tool for Learning. *Stud. Educ. Eval.* **1997**, *23*, 279–298. [\[CrossRef\]](#)
7. Rust, C. The Impact of Assessment on Student Learning How Can the Research Literature Practically Help to Inform the Development of Departmental Assessment Strategies and Learner-Centred Assessment Practices? *Act. Learn. High. Educ.* **2002**, *3*, 145–158. [\[CrossRef\]](#)
8. Cadez, S.; Dimovski, V.; Groff, M.Z. Research, Teaching and Performance Evaluation in Academia: The Salience of Quality. *Stud. High. Educ.* **2017**, *42*, 1455–1473. [\[CrossRef\]](#)
9. Villarroel, V.; Bloxham, S.; Bruna, D.; Bruna, C.; Herrera-Seda, C. Authentic Assessment: Creating a Blueprint for Course Design. *Assess. Eval. High. Educ.* **2018**, *43*, 840–854. [\[CrossRef\]](#)
10. Baroutsis, A.; Lingard, B. Counting and Comparing School Performance: An Analysis of Media Coverage of PISA in Australia, 2000–2014. *J. Educ. Policy* **2017**, *32*, 432–449. [\[CrossRef\]](#)
11. Fischbach, A.; Keller, U.; Preckel, F.; Brunner, M. PISA Proficiency Scores Predict Educational Outcomes. *Learn. Individ. Differ.* **2013**, *24*, 63–72. [\[CrossRef\]](#)
12. Bergsmann, E.; Schultes, M.-T.; Winter, P.; Schober, B.; Spiel, C. Evaluation of Competence-Based Teaching in Higher Education: From Theory to Practice. *Eval. Program Plan.* **2015**, *52*, 1–9. [\[CrossRef\]](#)
13. Keinänen, M.; Ursin, J.; Nissinen, K. How to Measure Students’ Innovation Competences in Higher Education: Evaluation of an Assessment Tool in Authentic Learning Environments. *Stud. Educ. Eval.* **2018**, *58*, 30–36. [\[CrossRef\]](#)
14. Gruppen, L.D.; ten Cate, O.; Lingard, L.A.; Teunissen, P.W.; Kogan, J.R. Enhanced Requirements for Assessment in a Competency-Based, Time-Variable Medical Education System. *Acad. Med.* **2018**, *93*, S17–S21. [\[CrossRef\]](#)
15. Boston, C. The Concept of Formative Assessment. *Pract. Assess. Res. Eval.* **2002**, *8*, 9. [\[CrossRef\]](#)
16. Underwood, S.M.; Posey, L.A.; Herrington, D.G.; Carmel, J.H.; Cooper, M.M. Adapting Assessment Tasks to Support Three-Dimensional Learning. *J. Chem. Educ.* **2018**, *95*, 207–217. [\[CrossRef\]](#)

17. Schut, S.; Driessen, E.; van Tartwijk, J.; van der Vleuten, C.; Heeneman, S. Stakes in the Eye of the Beholder: An International Study of Learners' Perceptions within Programmatic Assessment. *Med Educ.* **2018**, *52*, 654–663. [[CrossRef](#)] [[PubMed](#)]
18. Sevillano, M.L. *Didáctica en el Siglo XXI: Ejes en el Aprendizaje y Enseñanza de Calidad*; McGraw-Hill Interamericana de España: Madrid, Spain, 2004; ISBN 978-84-481-9820-6.
19. Kizilcec, R.F.; Perez-Sanagustin, M.; Maldonado, J.J. Self-Regulated Learning Strategies Predict Learner Behavior and Goal Attainment in Massive Open Online Courses. *Comput. Educ.* **2017**, *104*, 18–33. [[CrossRef](#)]
20. Pitt, E.; Norton, L. "Now That's the Feedback I Want!" Students' Reactions to Feedback on Graded Work and What They Do with It. *Assess. Eval. High. Educ.* **2017**, *42*, 499–516. [[CrossRef](#)]
21. Sánchez-Cabrero, R.; Estrada-Chichón, J.L.; Abad-Mancheño, A.; Mañoso-Pacheco, L. Models on Teaching Effectiveness in Current Scientific Literature. *Educ. Sci.* **2021**, *11*, 409. [[CrossRef](#)]
22. Bennett, S.; Dawson, P.; Bearman, M.; Molloy, E.; Boud, D. How Technology Shapes Assessment Design: Findings from a Study of University Teachers. *Br. J. Educ. Technol.* **2017**, *48*, 672–682. [[CrossRef](#)]
23. Romero-Martin, M.R.; Castejon-Oliva, F.-J.; Lopez-Pastor, V.-M.; Fraile-Aranda, A. Formative Assessment, Communication Skills and ICT in Initial Teacher Training. *Comunicar* **2017**, *25*, 73–82. [[CrossRef](#)]
24. Walker, D.-M. Classroom Assessment Techniques: An Assessment and Student Evaluation Method. *Creat. Educ.* **2012**, *3*, 903–907. [[CrossRef](#)]
25. Zhao, Y. Two Decades of Havoc: A Synthesis of Criticism against PISA. *J. Educ. Chang.* **2020**, *21*, 245–266. [[CrossRef](#)]
26. Bennett, R.E. Formative Assessment: A Critical Review. *Assess. Educ. Princ. Policy Pract.* **2011**, *18*, 5–25. [[CrossRef](#)]
27. Arigita-García, A.; Sánchez-Cabrero, R.; Barrientos-Fernández, A.; Mañoso-Pacheco, L.; Pericacho-Gómez, F.J. Pre-Eminence of Determining Factors in Second Language Learning: An Educator's Perspective from Spain. *Heliyon* **2021**, *7*, e06282. [[CrossRef](#)]
28. Barrientos-Fernández, A.; Pericacho-Gómez, F.-J.; Sánchez-Cabrero, R. Competencias sociales y emocionales del profesorado de Educación Infantil y su relación con la gestión del clima de aula. *Estud. Sobre Educ.* **2020**, *38*, 59–78. [[CrossRef](#)]
29. Biggs, J. Enhancing Teaching through Constructive Alignment. *High Educ.* **1996**, *32*, 347–364. [[CrossRef](#)]
30. Gaytan, J.; McEwen, B.C. Effective Online Instructional and Assessment Strategies. *Am. J. Distance Educ.* **2007**, *21*, 117–132. [[CrossRef](#)]
31. Pardo, A.; Jovanovic, J.; Dawson, S.; Gasevic, D.; Mirriahi, N. Using Learning Analytics to Scale the Provision of Personalised Feedback. *Br. J. Educ. Technol.* **2019**, *50*, 128–138. [[CrossRef](#)]
32. Siddiq, F.; Gochyyev, P.; Wilson, M. Learning in Digital Networks-ICT Literacy: A Novel Assessment of Students' 21st Century Skills. *Comput. Educ.* **2017**, *109*, 11–37. [[CrossRef](#)]
33. Sánchez-Cabrero, R.; Arigita-García, A.; Barrientos-Fernández, A.; León-Mejía, A.C. Online Explorative Study on the Learning Uses of Virtual Reality among Early Adopters. *J. Vis. Exp.* **2019**, e60188. [[CrossRef](#)]
34. Chetwynd, F.; Dobbyn, C. Assessment, Feedback and Marking Guides in Distance Education. *Open Learn. J. Open Distance E Learn.* **2011**, *26*, 67–78. [[CrossRef](#)]
35. Moss, J.; Hendry, G. Use of Electronic Surveys in Course Evaluation. *Br. J. Educ. Technol.* **2002**, *33*, 583–592. [[CrossRef](#)]
36. Cidral, W.A.; Oliveira, T.; Di Felice, M.; Aparicio, M. E-Learning Success Determinants: Brazilian Empirical Study. *Comput. Educ.* **2018**, *122*, 273–290. [[CrossRef](#)]
37. Hentea, M.; Shea, M.J.; Pennington, L. A Perspective on Fulfilling the Expectations of Distance Education. In Proceedings of the 4th Conference on Information Technology Curriculum, Lafayette, IN, USA, 16–18 October 2003; Association for Computing Machinery: New York, NY, USA, 2003; pp. 160–167.
38. Tucker, S.Y. *Assessing the Effectiveness of Distance Education versus Traditional On-Campus Education*; AERA: New Orleans, LA, USA, 2000.
39. Xenos, M. Prediction and Assessment of Student Behaviour in Open and Distance Education in Computers Using Bayesian Networks. *Comput. Educ.* **2004**, *43*, 345–359. [[CrossRef](#)]
40. Vonderwell, S.; Liang, X.; Alderman, K. Asynchronous Discussions and Assessment in Online Learning. *J. Res. Technol. Educ.* **2007**, *39*, 309–328. [[CrossRef](#)]
41. Gikandi, J.W.; Morrow, D.; Davis, N.E. Online Formative Assessment in Higher Education: A Review of the Literature. *Comput. Educ.* **2011**, *57*, 2333–2351. [[CrossRef](#)]
42. Khorsandi, M.; Kobra, A.; Ghobadzadeh, M.; Kalantari, M.; Seifei, M. Online vs. Traditional Teaching Evaluation: A Cross-Sectional Study. *Procedia Soc. Behav. Sci.* **2012**, *46*, 481–483. [[CrossRef](#)]
43. Gutiérrez, E.J.D.; Espinoza, K.G. Educar y evaluar en tiempos de Coronavirus: La situación en España. *Multidiscip. J. Educ. Res.* **2020**, *10*, 102–134. [[CrossRef](#)]
44. Giusti, A.D. Policy Brief: Education during COVID-19 and beyond. *Rev. Iberoam. Technol. Educ. Educ. Technol.* **2020**, *26*, e12. [[CrossRef](#)]
45. García, E.; Weiss, E. *COVID-19 and Student Performance, Equity, and U.S. Education Policy: Lessons from Pre-Pandemic Research to Inform Relief, Recovery, and Rebuilding*; Economic Policy Institute: Washington, DC, USA, 2020.
46. Thomas, M.S.C.; Rogers, C. Education, the Science of Learning, and the COVID-19 Crisis. *Prospects* **2020**, *49*, 87–90. [[CrossRef](#)]
47. Padayachee, P.; Wagner-Welsh, S.; Johannes, H. Online Assessment in Moodle: A Framework for Supporting Our Students. *S. Afr. J. High. Educ.* **2018**, *32*, 211–235. [[CrossRef](#)]

48. Watermeyer, R.; Crick, T.; Knight, C.; Goodall, J. COVID-19 and Digital Disruption in UK Universities: Afflictions and Affordances of Emergency Online Migration. *High. Educ.* **2021**, *81*, 623–641. [[CrossRef](#)] [[PubMed](#)]
49. Khalaf, K.; El-Kishawi, M.; Moufti, M.A.; Al Kawas, S. Introducing a Comprehensive High-Stake Online Exam to Final-Year Dental Students during the COVID-19 Pandemic and Evaluation of Its Effectiveness. *Med. Educ. Online* **2020**, *25*. [[CrossRef](#)] [[PubMed](#)]
50. Garcia-Penalvo, F.J.; Corell, A.; Abella-Garcia, V.; Grande, M. Online Assessment in Higher Education in the Time of COVID-19. *Educ. Knowl. Soc.* **2020**, *21*. [[CrossRef](#)]
51. González-González, C.S.; Infante-Moro, A.; Infante-Moro, J.C. Implementation of E-Proctoring in Online Teaching: A Study about Motivational Factors. *Sustainability* **2020**, *12*, 3488. [[CrossRef](#)]
52. Bretag, T.; Harper, R.; Burton, M.; Ellis, C.; Newton, P.; Rozenberg, P.; Saddiqui, S.; van Haeringen, K. Contract Cheating: A Survey of Australian University Students. *Stud. High. Educ.* **2019**, *44*, 1837–1856. [[CrossRef](#)]
53. Respondus Inc. LockDown Browser + Respondus Monitor. Available online: <https://web.respondus.com/he/monitor/> (accessed on 28 April 2021).
54. Silverman, S.C. What Happens When You Close the Door on Remote Proctoring? Moving Toward Authentic Assessments with a People-Centered Approach. *Improv. Acad. A J. Educ. Dev.* **2021**, *39*, 115–131. [[CrossRef](#)]
55. Gilbert, L.; Nunez-Argote, L.; Hunter, K. Prevention of Digital Cheating with Respondus Monitor. *Clin. Lab. Sci.* **2017**, *30*, 88.
56. Alessio, H.M.; Malay, N.; Maurer, K.; Bailer, A.J.; Rubin, B. Examining the Effect of Proctoring on Online Test Scores. *Online Learn.* **2017**, *21*, 146–161. [[CrossRef](#)]
57. Castaño, M.; Noeller, C.; Sharma, R. Implementing Remotely Proctored Testing in Nursing Education. *Teach. Learn. Nurs.* **2021**, *16*, 156–161. [[CrossRef](#)]
58. World Medical Association. World Medical Association Declaration of Helsinki—Ethical Principles for Medical Research Involving Human Subjects. *JAMA* **2013**, *310*, 2191–2194. [[CrossRef](#)]
59. Sanchez-Cabrero, R. La adaptación al ámbito educativo de los futuros docentes de secundaria según su área académica. *Rev. San Gregor.* **2021**, *1*, 114–129. [[CrossRef](#)]
60. Falch, T.; Naper, L.R. Educational Evaluation Schemes and Gender Gaps in Student Achievement. *Econ. Educ. Rev.* **2013**, *36*, 12–25. [[CrossRef](#)]
61. Astin, A.W.; Antonio, A.L. *Assessment for Excellence: The Philosophy and Practice of Assessment and Evaluation in Higher Education*; Rowman & Littlefield Publishers: Lanham, MD, USA, 2012; ISBN 978-1-4422-1362-3.
62. La Verdad, Exámenes “Online” a Trompicones y Entre Quejas de los Estudiantes. 2020. Available online: <https://www.laverdad.es/murcia/examenes-online-trompicones-20200603001115-ntvo.html> (accessed on 21 July 2021).
63. Levante, Profesores de la UNED Denuncian un Aprobado General en Varias Carreras. 2020. Available online: <https://www.levante-emv.com/comunitat-valenciana/2020/12/31/profesores-uned-denuncian-aprobado-general-26904289.html> (accessed on 21 July 2021).
64. Paechter, M.; Maier, B. Online or Face-to-Face? Students’ Experiences and Preferences in e-Learning. *Internet High. Educ.* **2010**, *13*, 292–297. [[CrossRef](#)]
65. Ilgaz, H.; Afacan Adanir, G. Providing Online Exams for Online Learners: Does It Really Matter for Them? *Educ. Inf. Technol.* **2020**, *25*, 1255–1269. [[CrossRef](#)]
66. Palmer, S.; Holt, D. Students’ Perceptions of the Value of the Elements of an Online Learning Environment: Looking Back in Moving Forward. *Interact. Learn. Environ.* **2010**, *18*, 135–151. [[CrossRef](#)]
67. Ramli, N.H.; Alavi, M.; Mehrihezahad, S.A.; Ahmadi, A. Academic Stress and Self-Regulation among University Students in Malaysia: Mediator Role of Mindfulness. *Behav. Sci.* **2018**, *8*, 12. [[CrossRef](#)] [[PubMed](#)]