Students' Mathematics Achievements: A Comparison between Pre-and Post-COVID-19 Pandemic

Article · December 2021

DOI: 10.26907/esd.16.4.04

1 author:

Daniel Doz
University of Primorska
32 PUBLICATIONS **23** CITATIONS

SEE PROFILE

Students' Mathematics Achievements: A Comparison between Pre- and Post-COVID-19 Pandemic

Daniel Doz

University of Primorska, Koper, Slovenia
E-mail: doz_daniel@yahoo.it
ORCID: https://orcid.org/0000-0002-6942-6937

DOI: 10.26907/esd.16.4.04
Submitted: 22 March 2021; Accepted: 18 October 2021

Abstract
The COVID-19 pandemic represented a shift from face-to-face to online and distance learning. Teaching methods and assessing strategies changed, as well as grading standards. The focus of this paper is to address the differences between pre- and post-COVID-19 pandemic mathematics grades and, in particular, to investigate the possible differences between mid-term and end-term grades in Italy. To this end, 231 Italian middle and high school students' grades were analysed. Using the Wilcoxon rank test (a non-parametric statistical test) the results showed a statistically significant difference in pre- and post-COVID-19 quarantine grades. End-of-year grades were higher than those before the COVID-19 confinement. Furthermore, the results indicated that more than half of the students in the sample achieved a higher grade at the end of the school year. Gender differences in mathematics grades were examined, since the literature about gender gap in mathematics achievement is not coherent about whether boys outperform girls or vice versa. Statistically significant differences at the end of the first semester were reported, in favour of female students although gender differences were not detected at the end of the school year. The findings suggest that greater caution should be paid in interpreting students' grades pre- and post-COVID-19 confinement, since it cannot be excluded that such students' achievements are inflated. Excessively high students' grades that do not represent their actual knowledge and competencies could give educators and legislators misleading and even false information about the quality of distance learning and students' knowledge.
Keywords: mathematics, COVID-19, evaluation and assessment, distance education.

Достижения учеников в области математики: сравнительный анализ показателей в период до и после пандемии COVID-19

Даниэль Доз

Приморский университет, Копер, Словения
E-mail: doz_daniel@yahoo.it
ORCID: https://orcid.org/0000-0002-6942-6937

DOI: 10.26907/esd.16.4.04
Дата поступления: 22 марта 2021; Дата принятия в печать: 18 октября 2021

Аннотация
Пандемия COVID-19 способствовала переходу от очного образования к онлайн и дистанционному образованию. Изменения произошли не только в методах преподавания, но и в стратегиях и стандартах оценивания. Исследование нацелено на сравнение оценок по математике, полученных учащимися в период до и после пандемии COVID-19, и, в частности, на изучение различий в промежуточных и итоговых оценках учащихся Италии. Для достижения постав-

ленной цели был проведен анализ успеваемости 231 ученика средних и старших классов школ Италии. Результаты, полученные с помощью Т-критерия Вилкоксона (непараметрического статистического теста), свидетельствуют о статистически значимой разнице в оценках, полученных в период до и после локдауна, вызванного COVID-19. Полученные в конце года итоговые оценки оказались выше, чем оценки в период до локдауна. Кроме того, согласно полученным результатам, более половины учащихся из выборки по окончании учебного года также получили более высокий балл. В ходе исследования изучались гендерные отличия в успеваемости по математике. В научной литературе, посвященной данной тематике, до сих пор не существует единого мнения о том, существует ли гендерное преимущество мальчиков над девочками в изучении математики или наоборот. По итогам первого семестра были получены статистически значимые различия, согласно которым школьницы превосходили школьников. При этом в конце учебного года гендерных различий обнаружено не было. Полученные результаты свидетельствуют о том, что при анализе оценок учащихся, полученных ими до и во время локдауна во время пандемии COVID-19, следует проявлять особую осторожность, поскольку нельзя исключать вероятность завышения успеваемости учащихся. Завышенные оценки, не отражающие реальные знания и компетенции студентов, могут ввести преподавателей и работодателей в заблуждение, так как могут содержать ложную информацию о качестве знаний выпускников.
Ключевые слова: математика, COVID-19, анализ и оценивание, дистанционное обучение.

Introduction

The COVID-19 pandemic has had an important impact on both teaching and learning. The emergence of COVID-19 meant a transition from face-to-face to online teaching (Brown, 2020; Daniel, 2020; Gilles & Britton, 2020; Zhao et al., 2020; Sahu, 2020; Istenič, 2021a, 2021b). As far as mathematics education is concerned, current studies in the field highlighted that mathematics lessons are less formal and, hence, could improve students' learning by creating a more relaxed learning environment (Mulenga & Marbán, 2020).

The transition from face-to-face to online learning has also changed the way teachers have to assess their students. Tests and examinations have been suspended by several institutions (Daniel, 2020) and new assessment methods have been used (see Rahim, 2020; Tuah & Naing, 2021). Shifting from face-to-face lessons to online or distance learning led teachers to assess students through open-book exams, assignments, and portfolios (Khan & Jawaid, 2020; Tuah & Naing, 2021). Consequently, differences in students' grades are expected (Schramm et al., 2021). For instance, Gonzalez and colleagues (2020) found that the COVID-19 confinement had some positive effects on students' performance. They compared students' grades in the year 2019/2020 with two control groups in the previous two years, finding that students after the pandemic had better results than those control group students had in the years before the confinement. The authors argued that higher scores are expected due to the COVID-19 confinement. However, they attributed these differences both to the new learning methodology and the new assessment processes.

Similarly, García Peñalvo and colleagues (2021) conducted a longitudinal study in which one area investigated the trend in college students' final grades. They considered the generation from the 2019/2020 academic year and the previous three academic years as control group. The researchers found a statistically significant increase in students' achievements after the pandemic, thus indicating that after the pandemic students tend to have higher grades than before it. Similar conclusions were drawn by Pócsová and colleagues (2021), who compared students' achievements from the academic year 2019/2020 and 2020/2021 with those from previous years. The results show that distance learning had a significant impact on the improvement of grades in the students' evaluations. They found that while more students passed the final exams in the academic year 2020/2021, distance education did not affect excellent students.

Moreover, greater caution needs to be taken while interpreting students' achievements, since cheating represents a threat to the validity and reliability of students' grades. Some authors highlighted the problem of cheating as one of the issues of online assessment (Khan & Jawaid, 2020; Nguyen et al., 2020; Bilen & Matros, 2021; Gamage et al., 2020) and specific programs are used to detect cheating during online tests (Lee et al., 2020). Hence, the problem should be examined by comparing students' grades before and after the COVID-19 quarantine (Basilaia & Kvavadze, 2020), since unreliable students' grades harm the whole educational system. In particular, excessively high and inflated grades are believed to give educators, students, families, and policymakers false and unreliable information (Felda, 2018). Whether is this the case with students' grades after the COVID-19 isolation, greater caution in interpreting students' grades should be taken. Students might get a false feeling that they mastered the covered material, despite lacking formal knowledge (Chowdhury, 2018; Butcher et al., 2014).

At present, the effects of distance learning on students' grades in Italy have not yet been investigated. Hence, addressing such question is important for two reasons: firstly, that it could provide a clearer understanding of the real achievement/understanding of the subject, and secondly, the comparison of students' grades before and after the COVID-19 isolation (Basilaia & Kvavadze, 2020) could provide an additional method to detect grade inflation and raise questions about the reliability of teachers' grades at the end of the quarantine. This, in turn, could shed some light on student's false expectations about their own knowledge and achievements.

Another important issue which relates to academic achievements pre- and post-quarantine periods, is the gender gap (Istenič, 2021a). For instance, Elzainy and colleagues (2020) found an increase in the mean of students' grades in female students in both face-to-face and online learning environments. However, the aspect of gender gap, especially in mathematics, has not been thoroughly investigated by international literature which has focused primarily on gender gap in mathematics under normal circumstances.

Various research studies have shown that there is a significant gender difference in grades. The literature, however, is not homogeneous, since some studies highlighted that boys have higher grades than girls (Ross et al., 2012; Weis et al., 2013), while others state the opposite (O'Dea et al., 2018; Devine et al., 2012; Hyde et al., 1990). Some researchers claim that gender differences in mathematics achievements might be the consequence of gender stereotypes in mathematics, which state that mathematics is more a male domain (Alan et al., 2018; Tsui, 2007; Cvencek et al., 2011). In addition, some researchers did not find any gender difference in math achievement (Else-Quest et al., 2010; Hyde et al., 2008; Penner & Paret, 2008). Hence the question about whether the isolation period helped reduce gender differences in mathematics achievement has no empirical or theoretical answer, and so this research attempts to provide a preliminary answer to this question.

In this paper, we address the question whether students' mathematics grades changed significantly before and after distance learning. We conducted a study involving 231 students with the aim of investigating whether there are any differences in students' mathematics grades before and after the pandemic. We argue that significantly higher students' grades at the end of the quarantine period could provide evidence of grade inflation and cheating, hence making students' mathematics grades unreliable. The question about whether there are gender differences in mathematics grades before and after distance learning is also considered, since distance learning could represent either a possible solution to normalize grades among genders, or an issue for students of a specific gender.

The Italian context and objectives of the study

Due to the spread of the COVID-19 disease in Italy, the Government decided to close all schools at the end of February 2020 (DPCM, 2020; MIUR, 2020a; MIUR, 2020b). The Government suggested that educators adopt distance learning methods in order to guarantee an almost regular and "normal" continuation of the learning activities for students. Teachers were invited to use different online platforms and apps for synchronous and asynchronous teaching, and a virtual contact with students was suggested: teachers were asked not to send only homework and material to students, but also to interact with them, providing explanations and important feedback information (MIUR, 2020b).

Moreover, the Italian Ministry of Education invited teachers to adopt new grading standards, which needed to be approved by the Class councils and based on new criteria (MIUR, 2020a). In particular, for secondary education (that is, middle and high schools), the Ministry stressed that teachers had to continue their regular assessment and grading procedures, using, as usual, grades from 1 to 10, where 6 is the first passing grade, 10 represents excellence and grades lower than 6 are failing grades (MIUR, 2018). Furthermore, students who did not have the possibility of being assessed, due to technical or other verified issues, such as unstable internet connection or lack of technological tools, were nevertheless evaluated by the council of all teachers teaching in that class, to determine whether or not they should pass (MIUR, 2020a).

The Italian Ministry of Education stressed the importance of developing new grading standards and assessment methods (DPCM, 2020; MIUR, 2020a). However, little is known about the distribution of grades in Italy before and after the COVID-19 health emergency. In particular, it has been shown that in other countries students' grades increased after the coronavirus confinement (Gonzalez et al., 2020). However, there were several questions about the validity and the reliability of students' grades after the COVID-19 pandemic (Khan & Jawaid, 2020; Nguyen et al., 2020). Basilaia and Kvavadze (2020) suggested that students' achievements should be compared before and after the isolation due to the spread of the coronavirus in order to get a clearer picture about the quality of students' learning. Following this idea, this paper addresses a similar question, namely whether students' mathematics grades increased during the period of the COVID-19 confinement from February 2020 to June 2020.

Since the problem has not yet been studied extensively, gender differences in mathematics grades before and after the coronavirus confinement are also included. The research also aimed to verify, whether students' achievements at the end of the first semester correlated to those at the end of the second semester. The aim of the research was to investigate students' mathematics grades before and after the COVID-19 quarantine in Italian schools. In particular, we wanted to answer the following research questions:

RQ1: Is there any difference in mathematics grades at the end of the first semester and at the end of the school year, i.e., before and after the coronavirus health emergency?

RQ2: Is there any difference in grades between males and females?

RQ3: Are students' grade at the end of the first and second semester correlated?

RQ4: Which grades did students get at the end of the school year?

Methods

Methodology

A non-experimental quantitative cross-sectional research method was used. The nature of the research is descriptive. Descriptive and inferential statistical methods were used.

Sample

The simple random sample consisted in 231 Italian students who attend a middle or high school with Slovene as language of instruction in North-Eastern Italy. There were 121 (52.4%) male and 110 (47.6%) female students. The average age of the participants was M=15.9 (SD=1.49; min=13; max=23). There were 16 (6.9%) middle school students, 51 (22.1%) 1st year high school students, 73 (31.6%) 2nd year high school students, 55 (23.8%) 3rd year high school students, 24 (10.4%) 4th year high school students and 12 (5.2%) last year of high school students.

Procedure and data analysis

In order to answer our research question, we examined both the students' mathematics grades in their January's report card, and in their June's report card. After obtaining students' (or their parents', if the students were minors) signed informed consent and the school principals' approval, we collected the students' grades in mathematics at the end of the first and second semester directly from the schools' official registers, avoiding self-reported data, which permitted us to obtain valid and reliable data. Despite all students' grades in Italy being public data (MIUR, 2021), we nevertheless asked students (or their parents) for a written informed consent, in order to follow ethical standards (cf. Cooper & Coetzee, 2020). We collected data about students' gender and mathematics grades at the end of the first and second semester. Data was anonymized using a coding scheme, such that anonymity and objectiveness were assured in every step of the research. The collected data were accessible only to the researcher (OECD, 2021).

The gathered data was analysed using the *Jamovi* statistical software. Firstly, demographical data was analysed using descriptive statistical methods. Secondly, inferential statistical methods were used. In particular, the assumptions of normality and equality of variances were checked using the Kolmogorov-Smirnov test and the Levene's test respectively. A non-parametric Mann-Whitney U test and the Wilcoxon signed rank tests were used to answer the RQ1 and RQ2. In order to answer RQ3, the Spearman's ρ correlation coefficient was used. Furthermore, descriptive statistical methods were used to answer RQ4.

Results

Firstly, we investigated possible differences in mathematics grades at the end of the first semester and at the end of the second semester. The mean students' grade in mathematics at the end of the first semester was M=7.61 (SD=1.47; min=3; max=10; Mdn=8). The absolute and relative frequencies are presented in Table 1. The Kolmogorov-Smirnov test showed a violation of the assumption of normality (KS=.120; p=.003), while the Levene's test for the equality of variances showed that there is no violation of such assumption (F(1,229)=1.39; p=.239).

The average students' mathematics grade at the end of the school year, i.e., in June, was M=8.14 (SD=1.37; min=4; max=10; Mdn=8). Final mathematics grades were not normally distributed (KS=.162; p<.001), with equal variances (F(1,229)=3.86; p=.051).

Table 1. Frequencies of students' mathematics grades.

	First semester		*Second semester*	
Grade	*f*	*%*	*f*	*%*
3	1	0.4	0	0
4	2	0.9	1	0.4
5	16	6.9	5	2.2
6	38	16.5	26	11.3
7	44	19.0	44	19.0
8	47	20.3	60	26.0
9	49	21.2	69	29.9
10	21	9.1	39	16.9

Comparing the students' mean grades at the end of the first semester (i.e., in January) with those at the end of the second semester (i.e., in June), a difference of .528 (*SE*=.068) might be observed and the statistical significance of this difference was tested. Due to the violation of the assumption of normality (*KS*=.191; $p<.001$), the differences between first semester's and second semester's grades were checked using the non-parametric Wilcoxon signed rank. The test results indicate a statistically significant difference between the grades at the end of the first semester and those at the end of the school year (*W*=9156; $p<.001$). The results showed that grades at the end of the school year are statistically higher than those at the end of the first semester.

Our second aim was to investigate possible grade differences between genders. At the end of the first semester, males had the mean grade *M*=7.98 (*SD*=1.42; *min*=5; *max*=10), while females had an average of *M*=8.31 (*SD*=1.29; *min*=4; *max*=10). At the end of the second semester, i.e., at the end of the school year, boys' mathematics grade was on average *M*=7.45 (*SD*=1.49; *min*=4; *max*=10), while for girl it was *M*=7.79 (*SD*=1.43; *min*=3; *max*=10). Due to the violation of the assumption of normality, the non-parametric Mann-Whitney U test was used to check the differences between men and women for the grade at the end of the first semester and second semester. Firstly, the test showed a statistical significance (*U*=5698; $p=.054$) at the end of the semester, indicating that girls tend to have higher mathematics grades than boys. Secondly, the non-parametric Mann-Whitney U test showed no statistically significant differences in mathematics grades between boys and girls (*U*=5753; $p=.069$) at the end of the second semester, which indicates that boys and girls had similar final grades for mathematics.

Furthermore, we found that the differences in mathematics grades before and after quarantine are statistically significant both for males (*W*=2679; $p<.001$) and females (*W*=1957; $p<.001$). In particular, both boys and girls had final grades that are statistically higher than grades in their first report cards.

Our third aim was to verify whether there was a correlation between the students' grades at the end of the first and second semester. Using the Spearman's ρ correlation coefficient, we found a positive and statistically significant correlation ($\rho=.750$; $p<.001$), indicating that students with higher grades at the end of the first semester also had higher grades at the end of the second semester.

Our last aim was to understand the students' grades at the end of the second semester. In order to verify the distribution of grades at the end of the first semester and at the end of the school year, a contingency table was constructed (see Table 2). At the end of the first semester, there were 19 students (8.2%) with a grade lower than 6, i.e., with a failing grade. At the end of the school year, 4 students (1.7%) had an insufficient grade. In particular, the only student with the grade "3" at the end of the first semester got a "7"

at the end of the second semester. Furthermore, the number of excellent students rose from 21 (9.1%) to 39 (16.9%). In particular, among the 49 students with a "9" at the end of the first semester, 19 (38.8%) got an excellent grade at the end of the school year. Overall, 31 students (13.4%) got in their final report card a grade lower than at the end of the first semester, 81 students (35.1%) got the same grade, while 119 (51.5%) students got a higher grade at the end of the second semester.

Table 2. Comparing grades at the end of first semester with those at the end of the second semester.

January	*June*							
	4	*5*	*6*	*7*	*8*	*9*	*10*	*Total*
3	0	0	0	1	0	0	0	1
4	1	0	0	1	0	0	0	2
5	0	3	7	5	1	0	0	16
6	0	1	10	14	10	3	0	38
7	0	0	6	12	16	10	0	44
8	0	1	3	9	15	29	3	60
9	0	0	0	2	5	23	19	49
10	0	0	0	0	0	4	17	21
Total	1	5	26	44	47	69	39	231

Discussion

The COVID-19 pandemic had a significant impact on education worldwide (Brown, 2020; Daniel, 2020; Gilles & Britton, 2020). In respect to students' grades, higher grades after the coronavirus confinement are expected (Gonzalez et al., 2020), not only because students might feel more comfortable at their homes (Mulenga & Marbán, 2020), but also because some of them might cheat during online tests and assessments (Khan & Jawaid, 2020; Nguyen et al., 2020; Bilen & Matros, 2021; Gamage et al., 2020). In this paper, we addressed the question as to whether there are significant differences in mathematics grades before and after the confinement in the school year 2019/2020. Other research questions investigated in depth the phenomenon of grade distribution in Italy before and after the COVID-19 confinement, including if there are differences in mathematics grades between genders, whether there was a correlation between grades at the end of the first and second semester, and the grades students obtained.

With our first research question, we aimed to verify whether there are statistically significant differences between students' mathematics grades before and after the quarantine. In order to verify the reliability of students' math grades at the end of the school year, we considered the suggestions made by Basilaia and Kvavadze (2020). The findings indicate that there is a statistically significant differences between grades pre- and post-quarantine. In particular, students at the end of the second semester had higher math grades.

We found that our results are in agreement with the claims of several researchers (Gonzalez et al., 2020; García Peñalvo et al., 2021; Pócsová et al., 2021), who found statistically significant improvements in students' grades before and after the COVID-19 confinement. Such improvement could be the result of students' being more relaxed and at the comfort of their homes (Mulenga & Marbán, 2020), an overall improvement of the teaching and learning methods (Gonzalez et al., 2020), but also because of cheating during online assessments (Khan & Jawaid, 2020; Nguyen et al., 2020; Lee et al., 2020;

Bilen & Matros, 2021; Gamage et al., 2020). Dealing with the question, whether there is evidence of grade inflation before and after the COVID-19 distance learning isolation, is important for different reasons.

Firstly, mathematics grades should express the students' real competencies and abilities, and are used by educators to improve their teaching methods and students to modify their learning strategies (Felda, 2018). Unreal and distorted information about students' knowledge could give legislators and teachers false information about the quality of the instruction (Chowdhury, 2018; Felda, 2018; Butcher et al., 2014). Moreover, inflated grades could consequently lead educators and legislators to provide insufficient help and resources for those students who fail to master the national programs, since their grades do not represent the real knowledge and competencies.

Secondly, statistically significant differences between grades in the first and second report cards might indicate that students cheated. Gonzalez and colleagues (2020) showed that students' grades increased in the last year. However, they do not exclude the possibility that students cheated. Other researchers hypothesized (Fuller et al., 2020) and found (Bilen & Matros, 2021; Balderas & Caballero-Hernández, 2020) that some students have cheated during online assignments and tests. Hence, higher grades are expected (Gonzalez et al., 2020), but they are not completely reliable (Tuah & Naing, 2021; Elsalem et al., 2021).

Thirdly, unreliable and excessively high (inflated) grades could create an unfair environment, which would give college entrance staff false information about students' knowledge. Furthermore, it could lead universities to invest more time, energy and resources to adequately prepare their future students. McManus and colleagues (2020) report the example of the United Kingdom abolishing all A-levels, GCSE and SQA assessments for summer 2020. Such examinations were replaced by calculated grades: teachers were asked to estimate the grades students would attain. Universities in the UK relied on the students' achievements on such exams in order to select their future students. In such case, inflated grades could lead to unfair opportunities for students, since they could not be assessed objectively.

This last scenario is only partially relevant to Italy, as university admission does not rely solely on student achievement. On the other hand, there are some faculties that consider students' grades as part of their entrance requirements. Thus, excessively high and inflated students' grades could lead to an unfair selection. Similarly, where employers rely on students' final achievements, high school grades that are based on unrealistic achievements might lead to unfair job opportunities.

Our second research question regarded possible differences in mathematics grades between boys and girls. In the international literature there is no consistent answer to this question. Some researchers found that girls tend to have higher school grades than boys (O'Dea et al., 2018; Devine et al., 2012; Hyde et al., 1990); other researchers found the contrary, i.e., that boys tend to have higher school grades in mathematics (Ross et al., 2012; Weis et al., 2013). Furthermore, some studies found little or no gender differences in mathematics achievements (Else-Quest et al., 2010; Hyde et al., 2008; Penner & Paret, 2008).

Our research considered students' mathematics grades as they appear in their report cards. At the end of the first semester, girls tended to have higher grades in mathematics than boys, however at the end of the school year, differences between genders were not statistically significant. Thus, our findings partially support results from previous research in respect of the students' grades at the end of the first semester (e.g., O'Dea et al., 2018; Devine et al., 2012; Hyde et al., 1990), but not at the end of the school year (e.g., Else-Quest et al., 2010; Hyde et al., 2008; Penner & Paret, 2008). Hence, the presented phenomenon

cannot be fully understood by solely considering previous research studies. Additional studies are needed. In particular, the question as to whether distance learning could help to equalize students' achievements in mathematics and reduce gender gap in mathematics (see also Armah et al., 2020), especially due to gender stereotypes in mathematics (Stoet & Geary, 2012), needs to be considered in future research studies.

With our third research question, we aimed to verify whether there was a correlation between students' grades at the end of the first semester and at the end of the semester. We found a strong positive and statistically significant correlation between the two grades. This means that students with higher grades at the end of the first semester also had higher grades at the end of the second semester.

We analysed the grades obtained by comparing the grade distribution at the end of the first semester and those at the end of the school year. Our findings suggest that more than half of the students had a higher grade in mathematics at the end of the school year, while little more than 13% of the students received a lower grade in their final report card. Also, the number of excellent students rose and those of underachieving students diminished. If, however, grades at the end of the second semester were inflated or obtained through cheating, educators would not only struggle with unobjective grades that do not represent students' real knowledge and competencies, but students would also be deprived of the extra help they would have receive during the summer through support lessons, which were provided by the schools (MIUR, 2020c).

Limitations and future research

The present work is not without limitations. The aims of the research were to understand the phenomenon of grade distribution before and after quarantine, and hence no other variables that could influence students' grades were considered. In particular, the questions about whether students cheated during the COVID-19 distance learning assessment, or whether teachers inflated students' grades, remain unanswered by the present research. Furthermore, students' motivation, their satisfaction with the new teaching methods and effort during the COVID-19 could be further investigated. Further analysis is required, for a clearer picture about the phenomenon described.

Nevertheless, this paper provides some preliminary insight about the distribution of math grades before and after the pandemic in Italy. In particular the students' grades being higher in the second semester, requires greater caution in interpreting students' achievements. Finding that students' mathematics grades are higher during the pandemic that they were before it, we argue that it represents a possible evidence of grade inflation, where students receive incomplete or even false information about their actual knowledge and competencies. Legislators and educators should thus interpret students' achievements with additional attention. Moreover, we found that there are no statistically significant differences in mathematics grades between genders after the pandemic. Additional research is required to fully understand, whether distance learning could help to reduce the gender gap in mathematics, thus leading to a decrease in gender stereotypes.

Future research is needed, in order to fully understand the scenario depicted and to evaluate the influence that variables, such as motivation, students' satisfaction with the teachers' methods, anxiety and other factors, may have on students' achievements during distance learning.

Statement on conflict of interest

The Author declares that there is no conflict of interest.

References

Alan, S., Ertac, S., & Mumcu, I. (2018). Gender stereotypes in the classroom and effects on achievement. *Review of Economics and Statistics, 100*(5), 876-890. https://doi. org/10.26300/3 w06-kw30.

Armah, S. E., Akayuure, P., & Armah, R. B. (2020). A Comparative Study of Male and Female Distance Learners ' Mathematics Achievement . *Contemporary Mathematics and Science Education, 2*(1), ep21001. https://doi.org/10.30935/conmaths/9288.

Balderas , A., & Caballero -Hernández , J. A. (2020 , October). Analysis of Learning Records to Detect Student Cheating on Online Exams : Case Study during COVID -19 Pandemic . In *Eighth International Conference on Technological Ecosystems for Enhancing Multiculturality* (pp. 752-757). https://doi.org/10.1145/3434780.3436662.

Basilaia, G., & Kvavadze, D. (2020). Transition to online education in schools during a SARS-CoV-2 coronavirus (COVID-19) pandemic in Georgia. *Pedagogical Research, 5*(4), 1-9. https:// doi.org/10.29333/pr/7937.

Bilen, E., & Matros, A. (2021). Online cheating amid COVID-19. *Journal of Economic Behavior & Organization, 182*, 196-211. https://doi.org/10.1016/j.jebo.2020.12.004.

Brown, S. (2020). Teaching Science Methods Online During COVID-19. *The Electronic Journal for Research in Science & Mathematics Education, 24*(3), 14-18.

Butcher, K. F., McEwan, P. J., & Weerapana, A. (2014). The effects of an anti-grade- inflation policy at Wellesley College. *Journal of Economic Perspectives, 28*(3), 189-204. http://dx.doi.org/ 10.1257/jep.28.3.189.

Chowdhury, F. (2018). Grade Inflation: Causes, Consequences and Cure. *Journal of Education and Learning, 7*(6), 86-92. http://dx.doi.org/10.5539/jel.v7n6p86.

Cooper, A. K. & Coetzee, S. (2020). On the Ethics of Using Publicly-Available Data. In: Hattings M., Matthee M., Smuts H., Pappas , I., Dwivedi Y. K., Mäntymäki M. (eds.) *Responsible Design, Implementation and Use of Information and Communication Technology*. I3 E 2020. Lecture Notes in Computer Science, vol. 12067. Springer: Cham. https://doi.org/10.1007 /978-3-030-45002-1_14.

Cvencek , D., Meltzoff , A. N., & Greenwald , A. G. (2011). Math –gender stereotypes in elementary school children. *Child development , 82*(3), 766-779. https://doi.org/10.1111/ j.1467-8624.2010.01529.x.

Daniel, S. J. (2020). Education and the COVID-19 pandemic. *Prospects*, 1-6. https:// doi.org/ 10.1007/s11125-020-09464-3.

Devine , A., Fawcett , K., Szűcs , D., & Dowker , A. (2012). Gender differences in mathematics anxiety and the relation to mathematics performance while controlling for test anxiety. *Behavioral and brain functions, 8*(1), 33. https://doi.org/10.1186/1744-9081- 8-33.

DPCM (2020). Acccessed December 11, 2020 at: https://www.gazzettaufficiale.it/eli/ id/2020/ 02/23/20A01228/sg

Elsalem, L., Al-Azzam, N., Jum'ah, A. A., & Obeidat, N. (2021). Remote E-exams during Covid-19 pandemic: A cross-sectional study of students' preferences and academic dishonesty in faculties of medical sciences. *Annals of Medicine and Surgery, 62*, 326–333. https://doi.org/10.1016 /j.amsu.2021.01.054.

Else-Quest, N. M., Hyde, J. S., & Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: a meta-analysis. *Psychological bulletin, 136*(1), 103- 127. https://doi. org/10.1037/a0018053.

Elzainy, A., El Sadik, A., & Al Abdulmonem , W. (2020). Experience of e-learning and online assessment during the COVID -19 pandemic at the College of Medicine , Qassim University . *Journal of Taibah University Medical Sciences , 15*(6), 456 -462 . https ://doi . org /10.1016 /j. jtumed.2020.09.005.

Felda, D. (2018). Preverjanje matematičnega znanja. *Journal of Elementary Education, 11*(2), 175-188. http://dx.doi.org/10.18690/rei.11.2.175-188.2018.

Fuller, R., Joynes, V., Cooper, J., Boursicot, K., & Roberts, T. (2020). Could COVID-19 be our ' There is no alternative'(TINA) opportunity to enhance assessment?. *Medical teacher, 42*(7), 781-786. https://doi.org/10.1080/0142159X.2020.1779206.

Gamage, K. A., Silva, E. K. D., & Gunawardhana, N. (2020). Online delivery and assessment during COVID-19: Safeguarding academic integrity. *Education Sciences, 10*(11), 301. https://doi.org/10.3390/educsci10110301.

García Peñalvo, F. J., García Holgado, A., Vázquez Ingelmo, A., & Sánchez-Prieto, J. C. (2021). Planning, Communication and Active Methodologies: Online Assessment of the Software Engineering Subject during the COVID-19 Crisis. *RIED. Revista Iberoamericana de Educación a Distancia, 24*(2), 41-66. https://doi.org/10.5944/ried.24.2.27689.

Gilles, B., & Britton, S. (2020). Moving Online: Creating a Relevant Learning Experience for Preservice Teachers in the Time of COVID-19. *Electronic Journal of Science & Mathematics Education, 24*(3), 19-28.

Gonzalez, T., De la Rubia, M. A., Hincz, K. P., Comas-Lopez, M., Subirats, L., Fort, S., & Sacha, G. M. (2020). Influence of COVID-19 confinement on students' performance in higher education. *PloS one, 15*(10), e0239490. https://doi.org/10.1371/journal.pone.0239490.

Hyde, J. S., Fennema, E., & Lamon, S. J. (1990). Gender differences in mathematics performance: A meta-analysis. *Psychological Bulletin, 107*(2), 139–155. https://doi.org/10.1037/0033-2909.107.2.139.

Hyde, J. S., Lindberg, S. M., Linn, M. C., Ellis, A. B., & Williams, C. C. (2008). Gender similarities characterize math performance. *Science, 321*(5888), 494-495. http://dx.doi.org/10.1126/science.1160364.

Istenič, A. (2021a). Shifting to digital during COVID-19: are teachers empowered to give voice to students? *Educational Technology Research and Development*, 69, 43–46. https://doi.org/10.1007/s11423-021-09956-9.

Istenič, A. (2021b). Online learning under COVID-19: re-examining the prominence of video-based and text-based feedback. *Educational Technology Research and Development*, 1-5. https://dx.doi.org/10.1007%2Fs11423-021-09955-w.

Khan, R. A., & Jawaid, M. (2020). Technology enhanced assessment (TEA) in COVID 19 Pandemic. *Pakistan journal of medical sciences, 36*(COVID 19-S4), S108. https://dx.doi.org/10.12669%2Fpjms.36.COVID19-S4.2795.

Lee, J., Kim, R. J., Park, S. Y., & Henning, M. A. (2020). Using technologies to prevent cheating in remote assessments during the COVID-19 pandemic. *Journal of dental education*, 1-3. https://doi.org/10.1002/jdd.12350.

McManus, I. C., Woolf, K., Harrison, D., Tiffin, P. A., Paton, L. W., Cheung, K. Y. F., & Smith, D. T. (2020). Calculated grades, predicted grades, forecasted grades and actual A-level grades: Reliability, correlations and predictive validity in medical school applicants, undergraduates, and postgraduates in a time of COVID-19. *medRxiv*. http://dx.doi.org/10.1101/2020.06.02.20116830.

MIUR (2018). Accessed March 3, 2021 at: https://www.miur.gov.it/valutazione MIUR (2020 a). Accessed December 11, 2020 at :documents/20182/2467413/m_pi.AOOGABMI.Registro+Decreti.0000011.16-05-2020.pdf/5bb159fa-1a35-fd30-02e4-6726901979ad?t=1589631914392

MIUR (2020b). Accessed February 17, 2021 at: https://www.miur.gov.it/documents/20182/0/Nota+prot.+388+del+17+marzo+2020.pdf/d6acc6a2-1505-9439-a9b4-735942369994?version=1.0&t=1584474278499

MIUR (2020 c). Accessed February 27, 2021 at: https ://www .miur .gov .it/ documents /20182 /2432359 /OM +VALUTAZIONE +FINALE +ALUNNI +A.S.+19 - 20+ RECUPERO +APPRENDIMENTI +.0000011.16-05-2020.pdf/c665ee9e-1752-c808 - ce67-9f3e3c 02ef7e?version=1.0&t=1589784478152

MIUR (2021). Accessed August 5, 2021 at: https://www.miur.gov.it/privacy-tra-i- banchi-di-scuola

Mulenga , E. M., & Marbán , J. M. (2020). Is COVID -19 the Gateway for Digital Learning in Mathematics Education?. *Contemporary Educational Technology, 12*(2), ep269. https://doi.org/10.30935/cedtech/7949.

Nguyen , J. G., Keuseman , K. J., & Humston , J. J. (2020). Minimize Online Cheating for Online Assessments During COVID-19 Pandemic. *Journal of Chemical Education, 97*(9), 3429-3435. https://doi.org/10.1021/acs.jchemed.0c00790.

O'Dea , R. E., Lagisz , M., Jennions , M. D., & Nakagawa , S. (2018). Gender differences in individual variation in academic grades fail to fit expected patterns for STEM . *Nature communications, 9*(1), 1-8. https://doi.org/10.1038/s41467-018-06292-0.

OECD (2021). Retrieved from: https://www.oecd.org/digital/digital-government/ goot-practice-principles-for-data-ethics-in-the-public-sector.htm (5th August 2021).

Penner , A. M., & Paret , M. (2008). Gender differences in mathematics achievement : Exploring the early grades and the extremes. *Social Science Research, 37*(1), 239-253. https:// doi.org/10.1016/j.ssresearch.2007.06.012.

Pócsová , J., Mojžišová , A., Takáč , M., & Klein , D. (2021). The Impact of the COVID-19 Pandemic on Teaching Mathematics and Students' Knowledge, Skills, and Grades. *Education Sciences, 11*(5), 225. https://doi.org/10.3390/educsci11050225.

Rahim , A. F. A. (2020). Guidelines for online assessment in emergency remote teaching during the COVID-19 pandemic. *Education in Medicine Journal, 12*(3), 59–68. https:// doi.org/10.21315/eimj2020.12.2.6.

Ross , J. A., Scott , G., & Bruce , C. D. (2012). The gender confidence gap in fractions knowledge : Gender differences in student belief –achievement relationships . *School Science and Mathematics, 112*(5), 278-288. https://doi.org/10.1111/j.1949-8594.2012.00144.x.

Sahu, P. (2020). Closure of universities due to coronavirus disease 2019 (COVID -19): impact on education and mental health of students and academic staff. *Cureus, 12*(4), e7541. https:// dx.doi.org/10.7759%2Fcureus.7541.

Schramm, H., Rubin, I., & Schramm, N. (2021). Covid-19 and high school grades: An early case study. *Significance, 18*(2), 6. https://doi.org/10.1111/1740-9713.01500.

Stoet , G., & Geary , D. C. (2012). Can stereotype threat explain the gender gap in mathematics performance and achievement ?. *Review of General psychology , 16*(1), 93 - 102. https://doi.org/10.1037%2Fa0026617.

Tsui, M. (2007). Gender and mathematics achievement in China and the United States. *Gender Issues, 24*(3), 1-11. https://doi.org/10.1007/s12147-007-9044-2.

Tuah , N. A. A., & Naing , L. (2021). Is online assessment in higher education institutions during COVID-19 pandemic reliable?. *Siriraj Medical Journal, 73*(1), 61-68. https:// doi.org/10.33192/Smj.2021.09.

Weis , M., Heikamp , T., & Trommsdorff , G. (2013). Gender differences in school achievement: The role of self-regulation. *Frontiers in Psychology, 4*, 442. https://doi. org/10. 3389/fpsyg.2013.00442.

Zhao, Y., Guo, Y., Xiao, Y., Zhu, R., Sun, W., Huang, W., ... & Wu, J. L. (2020). The Effects of Online Homeschooling on Children, Parents, and Teachers of Grades 1–9 During the COVID -19 Pandemic . *Medical Science Monitor : International Med - ical Journal of Experimental and Clinical Research, 26*, e925591 -1. https://dx.doi. org/10.12659%2FMSM. 925591.

