

# Inflated scores on the online exams during the COVID-19 pandemic school lockdown

Christina D. Hwang<sup>1</sup>, Soonnam S. Lee<sup>2</sup>

<sup>1</sup> Liberal Arts and Science Academy, Austin, Texas

<sup>2</sup> Department of Chemistry, St. Edward's University, Austin, Texas

## SUMMARY

Due to the COVID-19 pandemic, colleges and universities in the US had to provide students with online-only instruction and assessment in the middle of the Spring 2020 semester. Since this was neither adequately planned nor thoroughly prepared due to little time, its effects on student academic performance are unknown. This study tested a hypothesis that students' test scores were significantly higher on online exams during the COVID-19 school lockdown than those of the in-person exams before the lockdown. We compared the scores of the same students between one in-person exam and two online exams in a fundamental engineering class in the Spring 2020 semester. We further compared the scores between the students who took three in-person exams in the 2019 semesters and those in the Spring 2020 semester. The same instructor taught the class in both cases, and all exams had similar difficulties. We found that students achieved significantly inflated test scores when taking online exams, with the increase being in the range of 20% to 50%. Therefore, we have proven the hypothesis. Score inflation also depended on the level of prevention of student collaboration on online exams, with a loosely-controlled online exam's scores increasing by 50% compared to exams with strict measures to prevent cheating. In comparison, students showed a similar performance at three in-person exams in the 2019 semesters, with differences ranging 6% to 8%. The current study could help show how the COVID-19 pandemic influences long-term academic performance of institutions, instructors, and more importantly, students.

## INTRODUCTION

Due to the COVID-19 pandemic, many colleges and universities in the US closed their campuses in the middle of the Spring 2020 semester and implemented online-only instruction and assessment. Although online instructional platforms have been used in the past, the instructional change due to the COVID-19 pandemic was neither adequately planned nor thoroughly prepared due to little time. Interestingly, these unprecedented changes in learning environment have caused contrasting results. For example, college students had a hard time keeping track of assignments and dates in

online classes (1). On the contrary, only a moderate negative impact of the COVID-19 school lockdown on college students' academic accomplishment was reported (2). Nevertheless, it is generally accepted that the effects of the COVID-19 pandemic on students learning are multifaceted and complex. They encompass student academic accomplishment, admissions, retention, graduation, and student loans that would also be affected by socio-economic factors such as income and access to tutoring (3).

In addition, long-term interdisciplinary research is needed to advance the understanding of complex effects of the COVID-19 pandemic on the academic performance of students. As the COVID-19 is an on-going pandemic, there have been no reported results from any interdisciplinary long-term research. As such, preliminary and short-term studies may play a critical role in filling this gap and supporting the findings from larger-scale, long-term research in the future.

This study aimed to assess the immediate effect of sudden changes to learning environment on student academic performance. We speculated that there might be student collaboration and access to resources during the online exams or less test anxiety with better testing conditions at home. Thus, we hypothesized that students' test scores would be significantly higher on the online exams during the COVID-19 school lockdown compared to the in-person exams before the lockdown. To address this hypothesis, we compared student accomplishment on exams as a measure of academic performance before and during the COVID-19 school lockdown. Specifically, we compared the in-person and live-proctored exam scores of the same group of students in an engineering course from the Spring 2020 semester. In addition, we compared the exam scores between the normal semesters in 2019 when students took fully in-person exams and the Spring 2020 semester when students took an in-person exam before the COVID-19 school lockdown and online exams during the lockdown. We found that students achieved significantly inflated test scores on the online exams, and the score increase was more significant on the loosely-controlled online exam. These results likely reflect local effects of the COVID-19 school lockdown on the online exams' score inflation and they will play an important role in the advancement of understanding how the COVID-19 pandemic influences long-term academic performance of institutions, instructors, and more importantly, students.

RESULTS

Changes in exam scores in Spring 2020 semester

Students achieved higher scores, on average, on the online exams after the COVID-19 school lockdown than on the in-person exams before the COVID-19 school lockdown (Figure 1). In comparison to the in-person Exam 1, a greater increase of >50% was observed on the first online exam (Exam 2) compared to the second online exam (Exam 3), where the increase was by ~20%. A further comparison conducted with a paired t-test among the average scores at each exam revealed that the average score on the in-person Exam 1 was statistically lower ( $p < 0.05$ ) than those on the online Exams 2 and 3. Also, the average score on the online Exam 2 was significantly higher ( $p < 0.05$ ) than that on the online Exam 3 (Table 1).

Comparison between the normal 2019 semesters and the Spring 2020 semester

The trend of the exam scores in the Spring 2020 semester was dissimilar to that of the normal 2019 semesters (Figure 2). In the fully in-person 2019 semesters, there were slight changes in which the exam scores decreased by ~6% on the in-person Exam 2 and increased by ~8% on the in-person Exam 3 in comparison to the in-person Exam 1 score. However, a profound concave-down trend was observed in the Spring 2020 semester where the Exam 1 was conducted in-person, but both the Exams 2 and 3 were online. That is, in comparison to the score on the fully in-person Exam 1, students achieved higher scores by ~50% on the online Exam 2 and by ~20% on the online Exam 3 in the Spring 2020 semester.

DISCUSSION

The scores of the in-person Exam 1 that was given before the COVID-19 school lockdown in Spring 2020 were normally distributed (Figure 3). In comparison, the exam scores for the online Exams 2 and 3 during the lockdown in Spring 2020 did not follow a normal distribution but were left-skewed. As such, most of the students achieved higher scores on the online Exams 2 and 3 than on the in-person Exam 1 in Spring 2020. The Anderson-Darling (AD) test which was conducted to check a normality of data distribution also supported such results. The  $p$ -value of the in-person Exam 1 ( $AD=0.438$ ) was not statistically significant ( $p=0.290$ ), indicative of a normal distribution of the exam scores. On the contrary, the scores for the online Exams 2 ( $AD=9.332$ ;  $p < 0.005$ ) and 3 ( $AD=9.381$ ;  $p < 0.005$ ) were not in a normal distribution. Additionally, the AD values for the online Exams 2 and 3 were 9.332 and 9.381, respectively. In general, larger values of the AD statistic indicate that the data are not in a normal distribution.

A high-achieving academic performance of the online exams during the COVID-19 school lockdown was somewhat contradictory to the previous findings. For example, student performance, particularly for those who were already

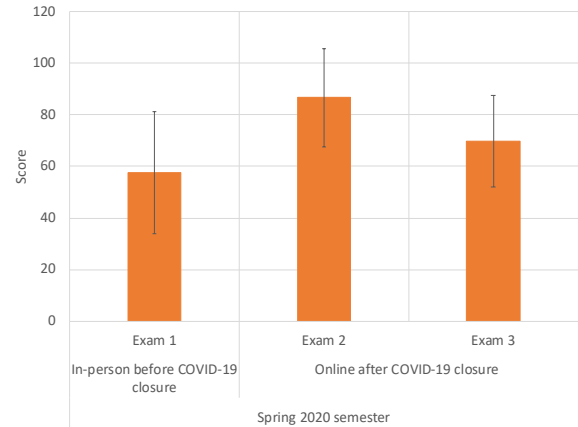


Figure 1: Comparison of the exam scores before and after COVID-19 school lockdown in the Spring 2020 semester. Data shown are the averages with standard deviations (n = 105). The average score on the in-person Exam 1 was statistically lower ( $p < 0.05$ ) than those on the online Exams 2 and 3.

		Exam 1	Exam 2	Exam 3
Exam 1	p-value	-	<0.0001	<0.0001
	$\mu_d^*$	-	-29.16	-12.19
	95% CI**	-	(-34.94, -23.38)	(-18.30, -6.07)
Exam 2	p-value	-	-	0.000
	$\mu_d$	-	-	19.97
	95% CI	-	-	(13.57, 20.37)

\* Difference between the averages.

\*\* 95% confidence interval of  $\mu_d$ .

Table 1: Results of the paired t-test for the exam scores in Spring 2020 semester. Exam 1 was in-person, whereas Exams 2 and 3 were online.

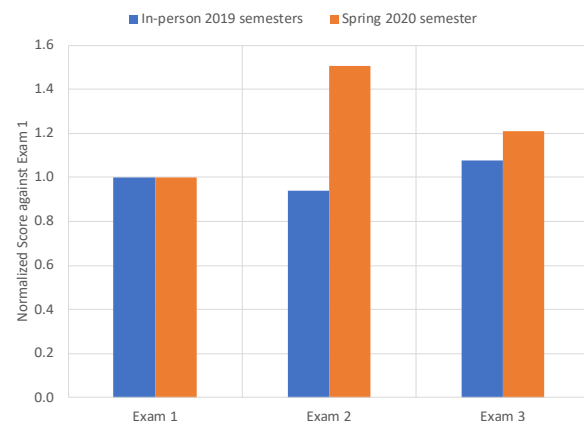
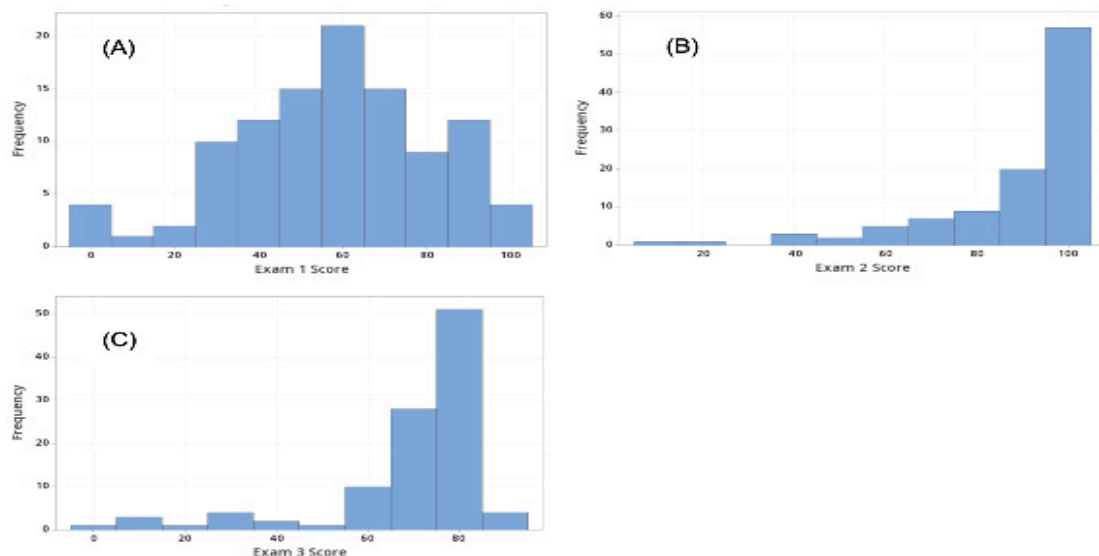


Figure 2: Comparison of the normalized exam scores between the 2019 semesters (fully in-person Exams 1, 2, and 3, n = 118) and the Spring 2020 semester (in person Exam 1 and online Exams 2 and 3, n = 105). Normalization was done by dividing the exam scores with the Exam 1 scores. For example, the scores of the Exams 1, 2, and 3 in in-person 2019 semesters were normalized by dividing them with the Exam 1 score in in-person 2019 semesters. In the fully in-person 2019 semesters, the exam scores decreased by ~6% on the in-person Exam 2 and increased by ~8% on the in-person Exam 3 in comparison to the in-person Exam 1. However, in the Spring 2020 semester, students achieved higher scores by ~50% on the online Exam 2 and by ~20% on the online Exam 3 in comparison to the in-person Exam 1.



**Figure 3: Histogram distribution of the exam scores in Spring 2020 semester.** Exam 1 was in-person before the COVID-19 school lockdown (Figure 3A), whereas Exams 2 and 3 (Figures 3B and 3C, respectively) were online during the lockdown. The scores of the in-person Exam 1 were normally distributed (Figure 3A), whereas the scores of the online Exam 2 (Figure 3B) and Exam 3 (Figure 3C) did not follow a normal distribution but were left-skewed.

academically struggling, can seriously suffer in online courses (4). In addition, some students may have had issues accessing effective technology for taking classes and exams. In fact, it was reported that roughly 20% of college students had difficulty maintaining access to working laptops and reliable high-speed internet (5). Despite that, this study showed that students performed statistically better on the online exam than on the in-person exam. This improved academic performance of students on the online Exams 2 and 3 during the COVID-19 school lockdown was not observed for the in-person Exams 2 and 3 taken in the normal 2019 semesters, suggesting that the increase in academic performance was a result of the online test format or other COVID mitigation measures. It is highly noted that the difficulties of the corresponding exams were similar to each other. As such, in the normal 2019 semesters, the average scores among the in-person Exams 1, 2, and 3 were not significantly different from each other: 6% difference between the Exam 1 and the Exam 2 and 8% between the Exam 1 and the Exam 3.

Based on the existing literature on the correlation between online exams and test scores, it is speculated that abnormal hikes of the exam scores observed on the online Exams 2 and 3 during the COVID-19 school lockdown might be attributed to student academic dishonesty (e.g., cheating) to some extent (6, 7). In general, students cheat more often online and when they are unmonitored (6, 7). Therefore, it would be surprising if no students cheated on the online Exams 2 and 3. Also, there could have been a technical difficulty for the professor to administer a webcam-based, live proctoring on more than 100 students on the online exams.

The aforementioned students' cheating on the online exams could have been reduced if the exam questions had been paraphrased rather than verbatim from the

previous exams (8). This is, in part, attributed to the fact that paraphrasing limits a student's ability to search for answers on the Internet. Exam performances using paraphrased questions helped students maintain academic integrity and better reflect students' knowledge (8). Another strategy to reduce cheating on online exams is delaying score availability to reduce the benefits of student collaboration. For example, students are not allowed to see their scores and feedback immediately after test completion. Instead, they can see them on a later date after the testing window ends so that a student who finishes early and knows the correct answers cannot advise other students who are still in the test (9). The results from our study are in agreement with this strategy. As a matter of fact, students were allowed to see the online Exam 2 scores and feedback immediately after submission, while students were permitted to see the online Exam 3 scores and feedback one day after the test. Exams 2 and 3 had a similar level of difficulties to each other. As shown in **Figure 1**, students achieved an averaged exam score of 86.7 on the online Exam 2 but their averaged exam score decreased to 69.7 on online Exam 3 most likely due to reduced benefits from the aforementioned student collaboration.

On the other hand, online exams can provide students with better testing conditions such as a more relaxing home environment that reduces test anxiety which could lead to better performance on the online exams (10). Therefore, this could also be contributing to inflated test scores on the online Exams 2 and 3. However, it should be noted that a mixed correlation between test anxiety and exam performance has also been reported. According to Stowell et al., students with low classroom test anxiety had significantly higher test anxiety levels when taking an online quiz (11). The authors also reported that students with high classroom test anxiety

had similarly high levels of test anxiety on the online quiz as well (11).

If the apparent increase of test scores on online exams observed in the current study does not reflect true improvement in student knowledge, significant pitfalls will follow in the near future with respect to long-term student academic performance. This speculation is supported by the findings of Still and Still who reported a grade inflation on online exams (12). They found that, on average, “A students” scored 7% higher, “B students” scored 11% higher, and “C students” scored 6% higher on the online exams than on the in-person exams. They were concerned about grade inflation because it might not be an accurate reflection of student knowledge and long-term comprehension of the class materials (12).

In conclusion, test score inflation in a fundamental engineering class was found on the online exams that were administered via a webcam live proctoring during the COVID-19 school lockdown. It is speculated that this score inflation was attributed to either student cheating (e.g., student collaboration, access to resources, etc.) on the online exams or less test anxiety due to better testing conditions at home. Also, the professor might have relaxed some of the strict grading requirements that had been in place before the COVID-19 pandemic. The data and findings from this study are believed to play a meaningful role in the advancement of understanding the long-term effects of the COVID-19 pandemic on student academic performance.

## MATERIALS AND METHODS

### Data source

Test scores analyzed in the current study were obtained from a Fluid Mechanics class taught at a higher-education institution. The particular class is a prerequisite for upper-level engineering courses. The same professor taught five sections of the class in-person in the Spring 2019 semester and two sections in-person in the Fall 2019 semester. Neither graduate student instructors nor other teaching assistants involved in monitoring or proctoring the exams. Typically, 20 to 25 students enrolled in each section in 2019. In the Spring 2020 semester, the same professor gave one large section of the class in-person before the COVID-19 school lockdown and online after the lockdown. Three Exams 1, 2, and 3 with similar difficulties and formats were given to students in the class. In the normal 2019 semesters, all three exams were fully in-person. However, in the Spring 2020 semester, only the first exam was fully in-person, and the two other exams were online due to the COVID-19 school lockdown. Similar numbers of students and similar ratios of female to male students were present in the class exams between the semesters (Table 2).

### Analysis

Test scores were averaged for the 2019 ( $n = 118$ ) and 2020 ( $n = 105$ ) semesters. Also, standard deviations were calculated for the averages. Changes in test scores due to

the COVID-19 school lockdown were statistically compared among the averages at an in-person exam (Exam 1) and two online exams (Exams 2 and 3) in the Spring 2020 semester. For this comparison among the same student group in the spring 2020 semester, a paired t-test was conducted with a significance level ( $\alpha$ ) of 0.05 as follows:

Null hypothesis ( $H_0$ ): There was no significant difference between two mean values. In other words, the mean difference ( $\mu_d$ ) = 0.

Alternative hypothesis ( $H_a$ ): There was a significant difference between two mean values. In other words,  $\mu_d \neq 0$ .

If the calculated *p-value* was less than or equal to the significance level ( $\alpha = 0.05$ ), the  $H_0$  was rejected, and it was concluded that there was a significant difference between two mean values. However, if the *p-value* was larger than  $\alpha$ , the  $H_0$  was accepted, and it was concluded that two means were not significantly different.

For the comparison between the 2019 semesters and the Spring 2020 semester, the average scores on the two exams (Exams 2 and 3) were normalized with those on the first exam (Exam 1) for each category (i.e., 2019 semesters vs. Spring 2020 semester). For example, the scores of the Exams 1, 2, and 3 in in-person 2019 semesters were normalized by dividing them with the Exam 1 score in in-person 2019 semesters. This normalization was done because student groups were different for each category.

To give more insight to the changes in test scores between the in-person and online exams in the Spring 2020 semester, the distribution of test scores was evaluated. Skewness and normality were compared with the histograms and by the Anderson-Darling goodness-of-fit statistic with a significance level of 0.05 as follows:

Null hypothesis ( $H_0$ ): The test scores are in a normal distribution.

Alternative hypothesis ( $H_a$ ): The test scores are not in a normal distribution.

If the calculated *p-value* was less than or equal to 0.05 ( $\alpha$ ), the  $H_0$  was rejected, and it was concluded that the test scores did not follow a normal distribution. However, if the *p-value* was larger than the  $\alpha$ , the  $H_0$  was accepted, and it was concluded that the test scores followed a normal distribution.

It should be noted that the t-test is commonly used for the data that are normally distributed and that the data obtained in the current study were not normally distributed. Non-parametric tests such as the Mann-Whitney test and the Kruskal-Wallis test are generally recommended for data that are not normally distributed (13) so that they could have been used in the current study. However, the t-test can work well for non-normally distributed data when there are large observations (e.g.,  $N > 25$ ) (14). Also, the use of t-test has an advantage over non-parametric tests. A non-parametric test only detects the difference between the observations but not compare the magnitude of the differences (15). That is, a non-parametric test will only produce a *p-value* whereas a t-test will also produce the observed mean difference between the



	Spring and Fall 2019 semesters	Spring 2020 semester
Total students	118	105
Female	28	21
Male	90	84

**Table 2. The numbers of students in the class in the 2019 and 2020 semesters.** In Spring and Fall 2019 semesters, there were five sections of the class. Each section with 20 to 25 students. Students took fully in-person three exams. In Spring 2020 semester, there was one large sections of class. Students took a fully in-person exam before the COVID-19 school lockdown and two online exams via a webcam live-proctoring after the COVID-19 school lockdown.

groups with a 95% confidence interval. In fact, the current study had more than 100 observations per group (Table 2). It should be also noted that a paired t-test was used for the data obtained in Spring 2020 semester when the same group of students took one in-person exam and two online exams (Table 2). There are other paired tests that make comparisons of the difference between two variables for the same subject, such as the Wilcoxon test (13).

**Received:** April 05, 2022

**Accepted:** May 30, 2022

**Published:** August 24, 2022

## REFERENCES

1. Wray, Maggie. "10 mistakes students make with online learning." *Creating Positive Futures*, 11 Apr. 2020, <https://creatingpositivefutures.com/10-mistakes-students-make-with-online-learning>. Accessed 10 Jul. 2022.
2. Garza, Victoria. "COVID-19's Impact on Spring Student Academic Success." ACC Newsroom, 11 Jun. 2020. [www.austincc.edu/news/2020/06/covid-19s-impact-spring-student-academic-success](http://www.austincc.edu/news/2020/06/covid-19s-impact-spring-student-academic-success). Accessed 27 Jun. 2020.
3. DePietro, Andrew. "Here's a Look at the Impact of Coronavirus (COVID-19) on Colleges and Universities in the U.S." *Forbes*, 30 Apr. 2020. [www.forbes.com/sites/andrewdepietro/2020/04/30/impact-coronavirus-covid-19-colleges-universities/#312c84ef61a6](http://www.forbes.com/sites/andrewdepietro/2020/04/30/impact-coronavirus-covid-19-colleges-universities/#312c84ef61a6). Accessed 27 Jun. 2020.
4. Xu, Di, and Shanna S. Jaggars. "Performance Gaps Between Online and Face-to-Face Courses: Differences Across Types of Students and Academic Subject Areas." *The Journal of Higher Education*, vol. 85, no. 5, Sep./Oct. 2014. doi: 10.1353/jhe.2014.0028.
5. Gonzales, Amy L., et al. "Technology Problems and Student Achievement Gaps: A Validation and Extension of the Technology Maintenance Construct." *Communication Research*, vol. 47, no. 5, Aug. 2018. doi:10.1177/0093650218796366.
6. Dyer, Jarret M., et al. "Academic Dishonesty and Testing: How Student Beliefs and Test Settings Impact Decisions to Cheat." *Journal of the National College Testing*

- Association, vol. 4, no. 1, 2020.
7. Downey, Maureen. "Colleges Struggle to Keep Students Honest after Classes and Exams Went Online due to Pandemic." *The Atlanta Journal-Constitution*, 30 Apr. 2020. [www.ajc.com/blog/get-schooled/georgia-tech-warns-physics-students-who-cheated-admit-risk-failing/3iQOCvBiwuqEXQivS63VqK](http://www.ajc.com/blog/get-schooled/georgia-tech-warns-physics-students-who-cheated-admit-risk-failing/3iQOCvBiwuqEXQivS63VqK). Accessed 28 Jun. 2020.
8. Golden, Joanna, and Mark Kohlbeck. "Addressing Cheating when Using Test Bank Questions in Online Classes." *Journal of Accounting Education*, vol. 52, Sep. 2020. doi: 10.1016/j.jaccedu.2020.100671
9. Smith Budhai, Stephanie. "Fourteen Simple Strategies to Reduce Cheating on Online Examinations." *Faculty Focus*, 11 May 2020. [www.facultyfocus.com/articles/educational-assessment/fourteen-simple-strategies-to-reduce-cheating-on-online-examinations](http://www.facultyfocus.com/articles/educational-assessment/fourteen-simple-strategies-to-reduce-cheating-on-online-examinations). Accessed 28 Jun. 2020.
10. Stowell, Jeffrey R., and Dan Bennett. "Effects of Online Testing on Student Exam Performance and Test Anxiety." *Journal of Educational Computing Research*, vol. 42, no. 2, 2010. doi: 10.2190/EC.42.2.b
11. Stowell, Jeffrey R., et al. "Emotions Experienced by Students Taking Online and Classroom Quizzes." *Journal of Educational Computing Research*, vol. 47, no. 1, 2012. doi: 10.2190/EC.47.1.e
12. Still, Mary Lynn, and Jeremiah D. Still. "Contrasting Traditional In-Class Exams with Frequent Online Testing." *Journal of Teaching and Learning with Technology*, vol. 4, no. 2, 2015. doi: 10.14434/jotlt.v4n2.13481
13. Montgomery, Douglas C. *Design and Analysis of Experiments*. John Wiley & Sons, 2021.
14. le Cessie, Saskia, et al. "Who is afraid of non-normal data? Choosing between parametric and non-parametric tests." *European Society of Endocrinology*, vol. 182, 2020. doi: 10.1530/EJE-19-0922
15. Kim, Tae Kyun, and Jae Hong Park. "More about the basic assumptions of t-test: normality and sample size." *Korean Journal of Anesthesiology*, vol. 72, no. 4, 2019. doi: 10.4097/kja.d.18.00292

**Copyright:** © 2022 Hwang and Lee. All JEI articles are distributed under the attribution non-commercial, no derivative license (<http://creativecommons.org/licenses/by-nc-nd/3.0/>). This means that anyone is free to share, copy and distribute an unaltered article for non-commercial purposes provided the original author and source is credited.