A New Mathematical Metric for Inclusive Excellence in Teaching Applied Before and During the COVID-19 Era

Jeffrey Thomas Ludwig[1,*]

[1]Department of Mathematics, University of California, Irvine, United States

*Corresponding author: Department of Mathematics, University of California, Irvine, United States. E-mail: jtludwig@uci.edu

Abstract

In this paper a novel metric for evaluating inclusive excellence in teaching is introduced and applied to students' performance in classes before and during the COVID-19 era. The novel metric, named the Inclusive Excellence Ratio (IER), is designed to simultaneously reflect the two desirable characteristics embraced by inclusive excellence teaching: strong student performance and low variation in performance across all students. The computation of the IER given student test score data is simple and straightforward: it is the statistical sample mean divided by the sample standard deviation. Consequently, the IER is high when the students' test scores are high and variance is low, suggesting it may provide a useful quantitative measure for those educational innovators seeking to experiment with new, effective teaching methodologies that boost inclusive excellence. The IER is applied to evaluate a posteriori student performance taken from cumulative aggregate data from the University of California, Irvine (UCI) undergraduate math finance classes involving 378 students over two academic years (2018 to 2020), spanning five quarters before and one quarter during the COVID-19 era of remote teaching at UCI. Conclusions are drawn and discussed comparing the quality of in-person teaching environments to remote teaching environments.

Keywords: inclusive excellence, COVID-19, remote teaching

1. Introduction

My arrival as a teacher at the University of California, Irvine (UCI) in 2018 was accompanied by the opening of the Anteater Learning Pavilion, the first facility in California designed with active learning principles in mind with 55,000 square feet of learning space optimized for meaningful student engagement and productive collaboration. With high hopes of participating in UCI's commitment to transforming the student learning experience and supporting progressive educational research, I set to work on refining and enhancing the math finance curriculum at UCI with ardent vigor.

When the COVID-19 pandemic hit the United States of America in 2020 and social distancing swept in as essential, UCI closed its doors to in-person teaching at the very end of the winter quarter in 2020, as did many other institutions based on guidelines from the Center for Disease Control and Prevention at that time (Centers for Disease Control and Prevention, 2020).

Pandemonium instantaneously set in as professors were urged to transition from in-person to remote teaching almost overnight. Marshall et al. (2020) explains that a large majority (92.4%) of polled American teachers self-reported that they had never taught online before the COVID-19 pandemic. As college campuses across the United States of America scrambled to initiate emergency remote teaching, Zoom became a popular modality of teaching. UCI pivoted to a remote learning environment almost exclusively across the campus in Irvine, California.

In the spring quarter I taught the introduction to math finance class in the Department of Mathematics remotely, using a live videoconference format with the Zoom Video Communications, Inc. communications technology for lectures and office hours, while quizzes and exams were administered online using the Canvas learning management system. As the quarter progressed, like many of my colleagues at UCI, I developed a curiosity about the quality of teaching and overall student experiences during remote teaching compared to the traditional in-person engagement. Was the remote teaching better, or worse? And how might one determine an answer to this question?

I set about surveying the pedagogical scientific literature for a meaningful way to measure the quality of teaching. Recently, data including test scores have been used to help identify and compare effective teaching methods (Hacking, 1983; Skedsmo, 2019). Data analysis can be pivotal in holding individuals and organizations accountable for their performance (Proitz, 2017). Other metrics such as student evaluations of teaching (SETs) are used to compare professors but may be riddled with selection biases (Goos, 2017) and the scores may not accurately reflect the professors' quality of instructional skill (Knol, 2016). Third-party observational tools subjectively rate teachers on specific modalities related to instruction, classroom culture, socioemotional skills (Molina, 2020), but are highly variable given the subjective nature of these observational tools. After scouring through the literature, I found no metric that focuses specifically on using test scores and data to bolster inclusive excellence.

In the end, I was unable to find a meaningful, robust, accurate method for quantifying the quality of teaching. This provided motivation to innovate and create a novel metric that would assist me in answering the question: how does remote teaching affect the quality of teaching? This led to the creation of a novel metric, named the Inclusive Excellence Ratio (IER). The IER is designed to simultaneously reflect the two desirable characteristics embraced by inclusive excellence teaching: strong student performance and low variation in performance across all students. The computation of the IER given student test score data is simple and straightforward: it is the statistical sample mean divided by the sample standard deviation. Typically, the final score that students earn during a course will be a percentage-weighted combination of their homework assignments, quizzes, and exams. For example, the homework total may contribute 30%, the quizzes 30%, and midterm/final exams 40%. In this case, for a student who averaged 90% on homework, 80% on quizzes, and 75% on both the midterm and final exam, the final score for the kth student would be f_k calculated as follows:

$$f_k = 0.3 * 90 + 0.3 * 80 + 0.4 * 75 = 81 \tag{1}$$

If the total final student scores are f_k, calculated for k = 1 to N, with N students in the class, then the Inclusive Excellence Ratio IER may be computed as follows:

$$IER = \bar{x} / s \tag{2}$$

Where the sample mean $\bar{x}$ is

$$\bar{x} = (f_1 + f_2 + f_3 + \ldots + f_N) / N \tag{3}$$

And the sample variance s^2 is

$$s^2 = ((f_1 - \bar{x})^2 + (f_2 - \bar{x})^2 + \ldots + (f_N - \bar{x})^2) / (N - 1) \tag{4}$$

Consequently, the IER is high when the sample mean $\bar{x}$ of the test scores f_k is high and variance s^2 of the test scores f_k is low, suggesting it may provide a simple and useful quantitative measure for those educational innovators seeking to experiment with new, effective teaching methodologies that boost inclusive excellence. This descriptive metric will be higher when fewer students are left behind in the classical DFW (final letter grade of D or F or withdraw from the class) category. In sum, teachers who are able to innovate to achieve high IER's will produce educational environments that include as many students as possible in the objective of educational excellence.

In this paper we will proceed first with a detailed description of the study involving a posteriori student performance taken from cumulative aggregate data from the University of California, Irvine (UCI) undergraduate math finance classes involving 378 students over two academic years (2018 to 2020), spanning five quarters before and one quarter during the COVID-19 era of remote teaching at UCI.

Numerical and graphical results are then presented. The paper concludes with a discussion comparing the quality of inclusive excellence with in-person to remote teaching environments, and potential contextual explanations for the quantitative results.

2. Method

The IER was applied to evaluate a posteriori student performance taken from cumulative aggregate data from the University of California, Irvine (UCI) undergraduate math finance classes involving 378 students over two academic years (2018 to 2020), spanning five quarters before and one quarter during the COVID-19 era of remote teaching at UCI. The study test periods are described in Table 1.

In order to investigate the quality of remote teaching compared to the quality of in-person teaching, the IER was computed for all the students that participated in the in-person study periods (total of 288 students from Fall 2018 to Winter 2020) and compared to the IER for the remote study test period (90 students in the Spring of 2020).

3. Results

The statistical results are summarized in Table 2 while the histograms of the raw cumulative scores for the in-person and remote groups are plotted in Figures 1 and 2, respectively.
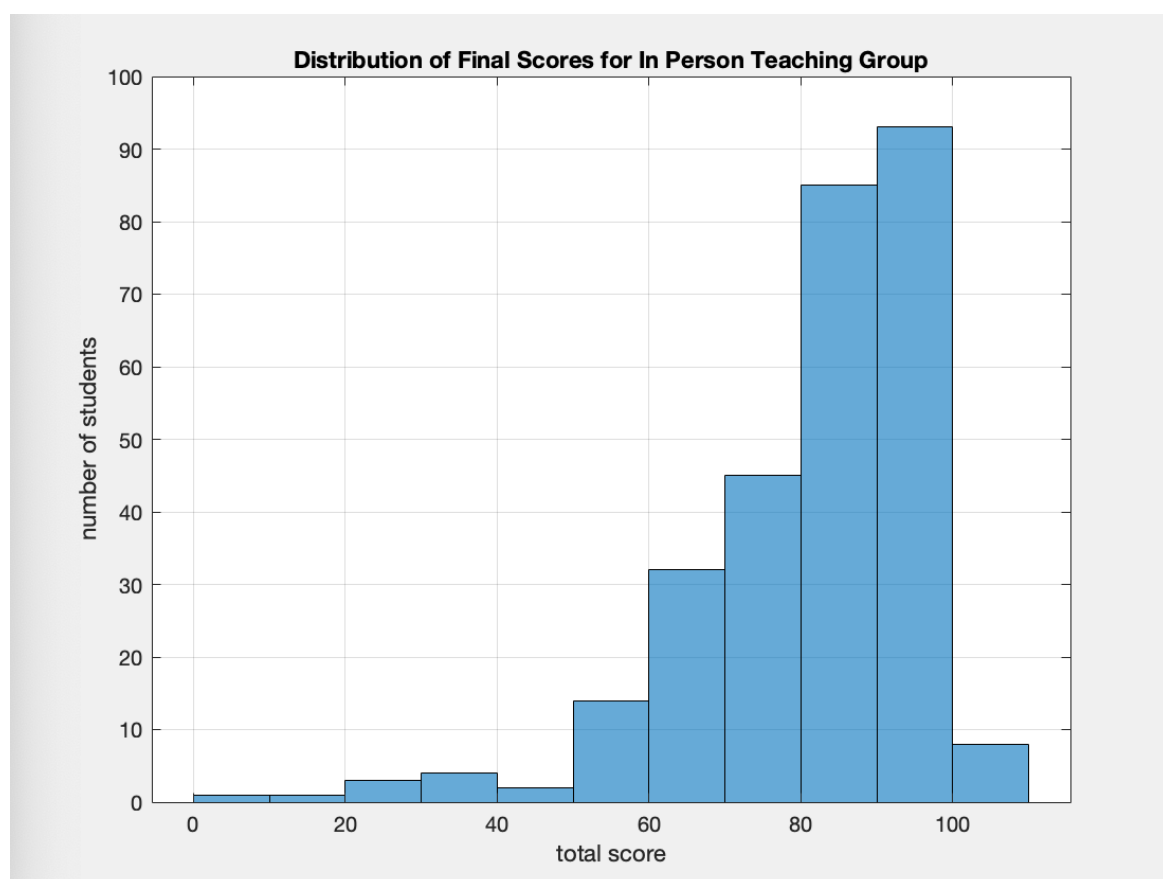

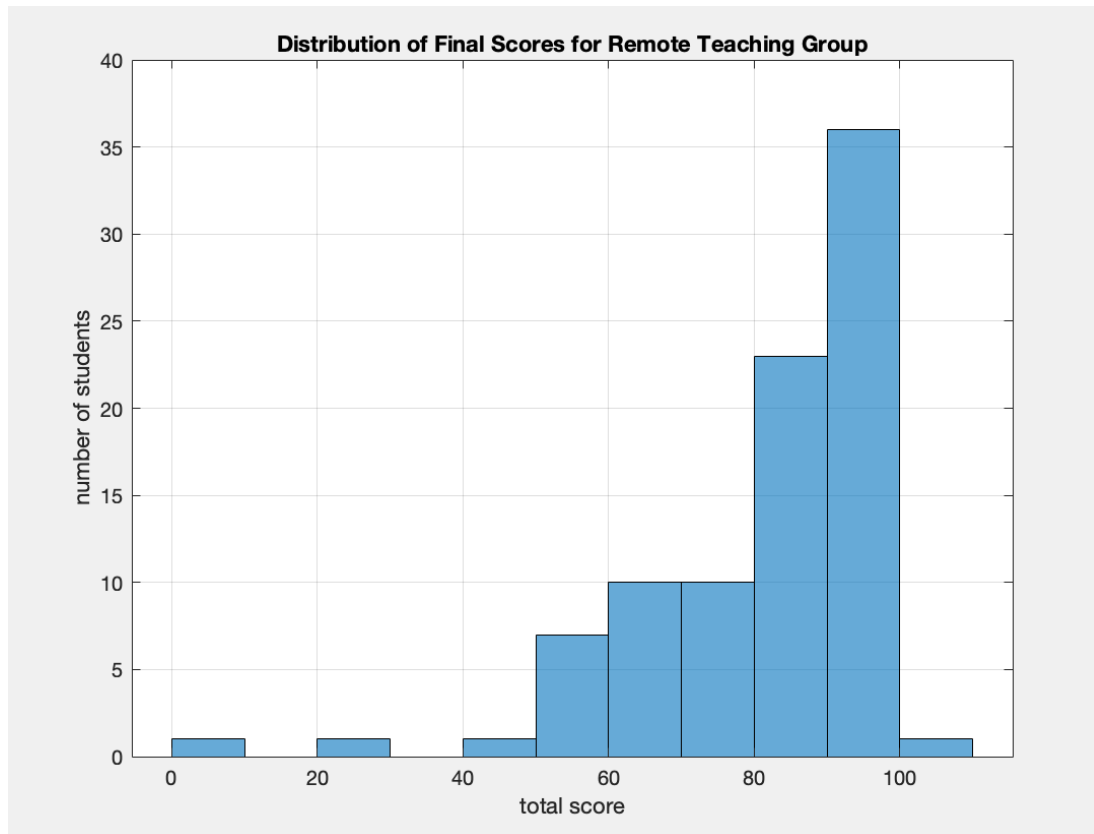**Figure 1.** Histogram of In-Person Period Student Cumulative Scores

Figure 2. Histogram of Remote Period Student Cumulative Scores

Table 1. Test Periods for Study

Quarter	Year	Total Students	Format
Fall	2018	29	In-Person
Winter	2019	46	In-Person
Spring	2019	78	In-Person
Fall	2019	88	In-Person
Winter	2020	47	In-Person
Spring	2020	90	Remote

Table 1 provides an overview of the study test periods. The study involved 378 undergraduate students in math finance classes over two academic years (2018 to 2020), spanning five quarters before and one quarter during the COVID-19 era of remote teaching at UCI.

In order to investigate the quality of remote teaching compared to the quality of in-person teaching, the IER was computed for all the students that participated in the in-person test periods (total of 288 students from Fall 2018 to Winter 2020) and compared to the IER for the remote study test period (90 students in the Spring of 2020). The statistical results are summarized in Table 2.

Table 2. Summary of Inclusive Excellence Ratio Test Results

Total Students	Format	Sample Mean	Sample Variance	IER
288	In-person	81.2	259.8	5.0
90	Remote	81.2	275.8	4.9

4. Limitations

Clearly there are many limitations to the conclusions that may or may not be drawn from this study and data analysis. First and foremost, the number of students in the remote learning data set was roughly a third of the sample size of the in-person data set.

Secondly, as the world was shocked with the emergence of the pandemic, international students were suddenly asked to relocate off-campus which clearly caused added stress and potentially internal emotional conflict that may have impaired their focus on studies. In addition, many of the students in this particular class relocated to international locations in time zones separated by more than 10 hours from the United States of America.

Furthermore, of course the students' cumulative test scores reflect their level of learning, but this is entangled with the professor's level of teaching. The students' learning may be impeded in a remote learning environment while also the professor's ability to teach in an emergent pandemic-related shift, lacking proper preparation, may also be impeded.

Also, student test scores may not accurately reflect the quality of teaching (e.g., high test scores may be attributable to grade inflation). Future directions for this study may include qualitative research methodologies to supplement the quantitative measures. Additionally, an experiment could be formulated using random assignments and a larger sample size, with adjustments for possible confounding seasonal effects. This would permit more rigorous statistical conclusions to be drawn, whereas in this study only descriptive and exploratory observations are made.

Lastly, the in-person students had ready access to in-person interactions with the professor and teaching assistant during lectures, weekly office hours, and private appointments. Live office hours and lectures via zoom may impede learning due to the restrictions of teleconferencing. For example, many students prefer to participate with audio and/or video muted, for privacy or other reasons; they may be more comfortable engaging in person than in teleconference. Rapid multiple student participation is difficult on zoom, and the audio and video quality depend on a robust, reliable and fast Wi-Fi connection.

Unfortunately, the math finance curriculum at UCI was modified such that the introduction to math finance class featured in this paper is no longer being taught. This means that further expansion of this study is impossible.

5. Discussion

In summary, the novel Inclusive Excellence Ratio (IER) is designed to simultaneously reflect the two desirable characteristics embraced by inclusive excellence teaching: strong student performance and low variation in performance across all students. The computation of the IER given student test score data is simple and straightforward: it is the statistical sample mean divided by the sample standard deviation. Consequently, the IER is high when the students' test scores are high and variance is low, suggesting it may provide a useful quantitative measure for those educational innovators seeking to experiment with new, effective teaching methodologies that boost inclusive excellence.

The statistical results from this initial study applying the Inclusive Excellence Ratio to measure teaching quality illuminates the fact that while both in-person and remote learning environments produced the same *average* performance (as measured by the sample mean of the cumulative student scores), there is a greater statistical variation of scores across all students in the remote teaching setting. Therefore, the in-person instruction may be associated with more desirable educational outcomes than remote instruction, as measured by the inclusive excellence ratio.

References

Centers for Disease Control and Prevention. (2020). *Considerations for Institutions of Higher Education.* Retrieved from https://www.cdc.gov/coronavirus/2019-ncov/community/colleges-universities/index.html

Goos, M., & Salomons, A. (2017). Measuring teaching quality in higher education: Assessing selection bias in course evaluations. *Research in Higher Education, 58*(4), 341-364. https://doi.org/10.1007/s11162-016-9429-8

Hacking, I. (1983). *Representation and intervening: Introductory topics in the philosophy of natural science.* Cambridge, Cambridge: University Press.

Knol, M., Dolan, C., Mellenbergh, G. J., & van der Maas, H. L. J. (2016). Measuring the Quality of University Lectures: Development and Validation of the Instructional Skills Questionnaire (ISQ). *PLOS ONE*. https://doi.org/10.1371/journal.pone.0149163

Marshall, D., Shannon, D., & Love, S. (2020). How teachers experienced the COVID-19 transition to remote instruction. *Phi Delta Kappan, The professional journal for educators*. September, 2020.

Molina, E., Fatima, S., Ho, A., Melo, C., Wilichowski, M., & Pushparatnam, A. (2020). Measuring the quality of teaching practices in primary schools: Assessing the validity of the Teach observation tool in Punjab, Pakistan. *Teaching and Teacher Education, 96*. https://doi.org/10.1016/j.tate.2020.103171

Proitz, T. S., Mausethagen, S., & Skedsmo, G. (2017). Investigative modes in research on data

use in education. *Nordic Journal of Studies in Educational Policy, 3*(1), 42-55. https://doi.org/10.1080/20020317.2017.1326280

Skedsmo, G., & Huber, S. G. (2019). Measuring teaching quality: some key issues. *Educational Assessment, Evaluation and Accountability, 31*, 151-153.

Copyrights