

Grade Inflation in Teacher Ratings: When Scores Rise Without Performance Gains

Edelaine M. Gabriel

*Institute of Graduate and Professional Studies, Lyceum-
Northwestern University, Dagupan City, Philippines*

ABSTRACT

This study investigated the phenomenon of grade inflation in teacher performance ratings, examining the extent to which rising evaluation scores reflect actual improvements in teaching effectiveness. Using longitudinal data over a five-year period, the research analyzed trends in teacher ratings, their relationship with student achievement, and the influence of evaluator characteristics and institutional pressures. Findings revealed a consistent upward trend in mean teacher performance ratings, increasing from 4.45 to 4.88, indicating significant score compression near the maximum scale. Despite this increase, changes in teacher ratings showed a non-significant negative correlation ($r = -0.15$) with student achievement gains, suggesting a disconnect between evaluation scores and actual performance outcomes. Further analysis demonstrated that veteran raters assigned significantly higher ratings than newer evaluators, indicating that leniency increases with experience. The most influential

predictor of grade inflation was identified as administrative pressure to meet performance targets, which accounted for a substantial proportion of variance in inflated scores. These findings suggest that teacher evaluation systems may be compromised by systemic and cultural factors that prioritize institutional image over accurate performance assessment. The study concludes that current evaluation practices risk losing their validity as tools for professional accountability and development. It highlights the need for structural reforms, including recalibration of rating scales, integration of objective performance indicators, and implementation of bias mitigation strategies, to ensure that evaluation systems accurately reflect teacher effectiveness and support meaningful educational improvement.

Keywords: *Grade Inflation, Teacher Evaluation, Performance Measurement*

INTRODUCTION

Grade inflation has emerged as a significant concern in educational systems worldwide, with implications for academic integrity, student motivation, and institutional credibility. Research indicates that rising grades do not always reflect genuine improvements in student performance, raising questions about the validity of grading practices and their impact on teacher

evaluations (Baglione et al., 2025; Bachan, 2017). While some argue that higher grades result from pedagogical improvements (Jephcote et al., 2021), others contend that external pressures-such as student expectations, institutional policies, and faculty workload-distort grading standards (Ternikov & Blyakher, 2024; Ullah et al., 2024). This literature review synthesizes key findings on the causes, consequences, and perceptions of grade inflation, particularly in relation to teacher evaluations.

One prominent factor contributing to grade inflation is the pressure on faculty to maintain high student satisfaction, often linked to teaching evaluations. Park and Cho (2023) found that relaxed grading policies during the COVID-19 pandemic altered students' grade expectations, subsequently influencing their evaluations of instructors. When students received higher grades than anticipated, they were less likely to penalize teachers in evaluations, suggesting that inflated grades may artificially inflate teaching ratings. Similarly, Baglione and Smith (2022) discovered that undergraduate students, while acknowledging some grade inflation, generally believed grades were an accurate reflection of achievement, though female students were more likely to suspect that faculty inflated grades to secure better evaluations. These findings highlight a potential feedback loop where lenient grading leads to favorable evaluations, reinforcing the cycle of inflation.

Faculty workload and institutional incentives further exacerbate grade inflation. Ternikov and Blyakher (2024) demonstrated that excessive teaching loads and resource scarcity encourage faculty to simplify assessments or award higher grades to reduce conflict and workload stress. In Pakistan, Ullah et al. (2024) found that competition between universities for student enrollment led to lenient grading, particularly when untrained visiting faculty—who sought permanent positions—felt pressured to satisfy students. These institutional dynamics suggest that grade inflation is not merely an individual faculty issue but a systemic problem driven by broader academic and economic pressures.

The consequences of grade inflation extend beyond distorted teacher evaluations. Kiziltas (2024) revealed that inflated grades in primary and secondary schools often mask student underperformance, creating a false sense of achievement. Griffin and Townsley (2021) found that traditional grading models incorporating non-academic factors (e.g., homework completion, employability skills) inflated or deflated nearly half of students' grades by at least 5%, undermining the accuracy of academic assessments. Such practices can disadvantage students in competitive environments where grades serve as key indicators of competence (Gershenson, 2020). Moreover, grade inflation erodes public trust in educational standards, as noted by Horowitz et al. (2023), who found that many tenured U.S.

faculty members perceived declining academic rigor and grade inflation as serious ethical concerns.

The debate over whether rising grades reflect genuine improvement or systemic inflation remains unresolved. Jephcote et al. (2021) argued that higher grades in UK universities could signal pedagogical enhancements rather than inflation, emphasizing the need for contextual analysis. However, Bachan (2017) countered that unexplained grade increases in UK institutions pointed to inflation, particularly after controlling for student ability and institutional efficiency. Similarly, Kostal et al. (2016) documented persistent grade inflation in U.S. colleges over two decades, suggesting that rising grades were not fully attributable to improved student performance. These conflicting perspectives underscore the complexity of disentangling genuine academic progress from systemic grading biases. Cultural and policy factors also shape grading practices. In Indonesia, Arrafii (2020) found that "hodgepodge grading"-incorporating non-cognitive factors like effort and attendance-led to grade inflation, exacerbated by top-down competency-based assessment policies. Meanwhile, Morrison (2021) applied a desert-based justice framework to critique England's degree classification system, arguing that utilitarian pressures (e.g., labor market demands) distorted the retrospective fairness of grades. Such studies highlight how national

educational policies and cultural norms influence grading leniency.

METHODS

Research Design

This chapter presents research design, the sources of data which includes the locale of the study and the research population, instrumentation and data collection, and the tools for data analysis.

Participants and Locale

This chapter presents research design, the sources of data which includes the locale of the study and the research population, instrumentation and data collection, and the tools for data analysis.

Sampling and Data Collection

This chapter presents research design, the sources of data which includes the locale of the study and the research population, instrumentation and data collection, and the tools for data analysis.

Data Analysis

This study utilized a non-experimental, quantitative research design, specifically a descriptive-correlational and predictive approach. This design was appropriate as it did not involve the manipulation of

variables but instead sought to describe phenomena, examine the relationships between them, and determine their predictive power based on existing data. It allowed for a systematic and statistical evaluation of teacher performance ratings in relation to other key variables over a specified period. The design was structured to provide quantitative answers to the core objectives of the investigation without influencing the educational environment.

Ethical Considerations

The research was conducted within the geographical and administrative confines of the Schools Division Office (SDO) of Tarlac City. Located in the province of Tarlac in the Central Luzon region of the Philippines, Tarlac City served as an urbanized educational center that hosted a significant number of public elementary and secondary schools. The urban setting of the SDO provided a context where performance metrics and school-level competition were particularly salient, making it a highly relevant locale for investigating phenomena such as grade inflation and performance pressure. Prior to data collection, formal approval and ethical clearance were secured from the Office of the Schools Division Superintendent to ensure full cooperation and access to the necessary anonymized school-level data.

RESULTS AND DISCUSSION

1. This study investigated the phenomenon of grade inflation in teacher performance ratings, examining the extent to which rising evaluation scores reflect actual improvements in teaching effectiveness.
2. Using longitudinal data over a five-year period, the research analyzed trends in teacher ratings, their relationship with student achievement, and the influence of evaluator characteristics and institutional pressures.
3. Findings revealed a consistent upward trend in mean teacher performance ratings, increasing from 4.45 to 4.88, indicating significant score compression near the maximum scale.
4. Despite this increase, changes in teacher ratings showed a non-significant negative correlation ($r = -0.15$) with student achievement gains, suggesting a disconnect between evaluation scores and actual performance outcomes.
5. Further analysis demonstrated that veteran raters assigned significantly higher ratings than newer evaluators, indicating that leniency increases with experience.
6. The most influential predictor of grade inflation was identified as administrative pressure to meet performance targets, which accounted for a substantial proportion of variance in inflated scores.

7. These findings suggest that teacher evaluation systems may be compromised by systemic and cultural factors that prioritize institutional image over accurate performance assessment.
8. It highlights the need for structural reforms, including recalibration of rating scales, integration of objective performance indicators, and implementation of bias mitigation strategies, to ensure that evaluation systems accurately reflect teacher effectiveness and support meaningful educational improvement.

PRESENTATION, ANALYSIS, AND INTERPRETATION OF DATA

This chapter presents the discussion of findings brought from the data gathering procedure.

Longitudinal Trend of Mean Teacher Performance Ratings Over the Past Five Academic Years Table 1
Longitudinal Trend of Mean Teacher Performance Ratings Over the Past Five Academic Years Academic Year | Mean Teacher Performance Rating (1-5 Scale) | Annual Change (Δ Rating) | Interpretation of Trend Year 1 | 4.45 | N/A | High initial rating.

CONCLUSION AND RECOMMENDATIONS

Conclusions

Based on the results, the following conclusions were deduced:

The system's relentless upward drift confirms the presence of systemic grade inflation, rendering the distinction between performance levels functionally irrelevant for human resource decisions.

The rising evaluation scores are unjustified by objective performance gains, confirming that the performance review system measures evaluator leniency rather than teacher efficacy.

Grade inflation is a learned cultural behavior, suggesting that long-term exposure to the school's administrative norms normalizes the awarding of near-perfect scores.

Administrative pressure transforms the evaluation process into a tool for organizational risk management, where ratings are inflated to report institutional success and avoid accountability for missed targets.

Recommendations

1. Based on the conclusions, the following are recommended:
2. Immediately reset the evaluation instrument's score distribution to utilize the full scale, enforcing a median score closer to the rubric's mid-point (e.g., 3.0)

3. Mandate that all performance evaluations include a weighted component tied directly to objective, validated student achievement data to restore accountability to the review process. Implement mandatory, recurrent rater bias training focused on recalibrating veteran evaluators to the rubric's true standards and minimizing the effect of professional habituation. Restructure the evaluation process to remove the School Head's direct reporting incentive by linking performance data directly to a neutral, third-party audit system.

REFERENCES

- Arrafii, M. A. (2020). Grades and grade inflation: Exploring teachers' grading practices in Indonesian EFL secondary school classrooms. *Pedagogy, Culture and Society*, 28(3), 477-499. <https://doi.org/10.1080/14681366.2019.1663246>
- Bachan, R. (2017). Grade inflation in UK higher education. *Studies in Higher Education*, 42(8), 1580-1600. <https://doi.org/10.1080/03075079.2015.1019450>
- Baglione, S. L., & Smith, Z. (2022). Grade inflation: Undergraduate students' perspective. *Quality Assurance in Education: An International Perspective*, 30(2), 251-267. <https://doi.org/10.1108/QAE-04-2021-0063>
- Baglione, S. L., Smith, Z., & Roach, O. (2025). Grade inflation: Graduate students' perspective. *Educational Research and Evaluation*, 30(1-2), 115-136. <https://doi.org/10.1080/13803611.2024.2422617>
- Chowdhury, F. (2018). Grade inflation: Causes, consequences and cure. *Journal of Education and Learning*, 7(6), 86-92. <https://files.eric.ed.gov/fulltext/EJ1191199.pdf>
- Gershenson, S. (2020). End the "Easy A": Tougher grading standards set more students up for success. *Education Next*, 20(2), 18-24. <https://www.educationnext.org/journal/spring-2020-vol-20-no-2/>

Griffin, R., & Townsley, M. (2021). Points, points, and more points: High school grade inflation and deflation when homework and employability scores are included. *Journal of School Administration Research and Development*, 6(1), 1-11. <https://files.eric.ed.gov/fulltext/EJ1301287.pdf>

Horowitz, M., Haynor, A. L., & Kickham, K. (2023). "Undeserved" grades or "underserved" students? Faculty anxieties and eroding standards in the corporate university. *Higher Education Politics & Economics*, 9(1), 43-84. <https://files.eric.ed.gov/fulltext/EJ1377743.pdf>

Jephcote, C., Medland, E., & Lygo-Baker, S. (2021). Grade inflation versus grade improvement: Are our students getting more intelligent? *Assessment & Evaluation in Higher Education*, 46(4), 547-571. <https://doi.org/10.1080/02602938.2020.1795617>

Kiziltas, Y. (2024). An important problem we need to face in schools: Inflated grades and grade inflation. *European Journal of Education*, 59(4), e12744. <https://doi.org/10.1111/ejed.12744>

Kostal, J. W., Kuncel, N. R., & Sackett, P. R. (2016). Grade inflation marches on: Grade increases from the 1990s to 2000s. *Educational Measurement: Issues and Practice*, 35(1), 11-20. <https://doi.org/10.1111/emip.12077>

Lane, C., Schrynemakers, I., & Kim, M. (2020). Examining the academic effects of developmental education reform: Faculty perceptions from a large,

public, urban university. *Community College Enterprise*, 26(1), 27-57.
<https://www.schoolcraft.edu/cce/community-college-enterprise>

Maamari, B. E., & Naccache, H. S. (2022). The impact of grade inflation on teachers' evaluation: A quantitative study conducted in the context of five Lebanese universities. *Journal of Global Education and Research*, 6(2), 192-205.
<https://files.eric.ed.gov/fulltext/EJ1347283.pdf>

Mirazchiyski, P. V., & Klemencic Mirazchiyski, E. (2023). Students' and teachers' perceptions of students' academic outcomes in Slovenia: Evidence from REDS data. **Large-scale Assessments in Education*, 11*(23).
<https://doi.org/10.1186/s40536-023-00173-9>

Morrison, A. (2021). Just deserts? Grade inflation and desert-based justice in English higher education. *British Journal of Educational Studies*, 69(4), 437-451.
<https://doi.org/10.1080/00071005.2020.1832957>

Park, B., & Cho, J. (2023). How does grade inflation affect student evaluation of teaching? *Assessment & Evaluation in Higher Education*, 48(5), 723-735.
<https://doi.org/10.1080/02602938.2022.2126429>

Robbins, J., Firth, A., & Evans, M. (2018). Improving work-based assessment: Addressing grade inflation numerically or pedagogically? *Practitioner Research in Higher Education*, 11(1), 80-86.
<https://files.eric.ed.gov/fulltext/EJ1180127.pdf>

- Ternikov, A., & Blyakher, M. (2024). Grade inflation and grading process: Does faculty workload matter? *Journal of Applied Research in Higher Education*, 16(5), 1937-1955. <https://doi.org/10.1108/JARHE-01-2023-0024>
- Ullah, H., Qureshi, A. W., & Ali, R. (2024). Grade inflation in Pakistan higher education: Teachers' perspective. *Journal of Further and Higher Education*, 48(7), 700-713. <https://doi.org/10.1080/0309877X.2024.2386455>
- Vítecková, K., & Houdek, P. (2025). The rise of business education, the ESG revolution, and the limited impact on students' values. *Interchange: A Quarterly Review of Education*, 56(2), 143-157. <https://doi.org/10.1007/s10780-025-09540-7>
- Yalçın, Ü. D. (2023). Analyzing grade inflators: Some metaethical issues. *Journal of Academic Ethics*, 21(4), 649-668. <https://doi.org/10.1007/s10805-023-09473->