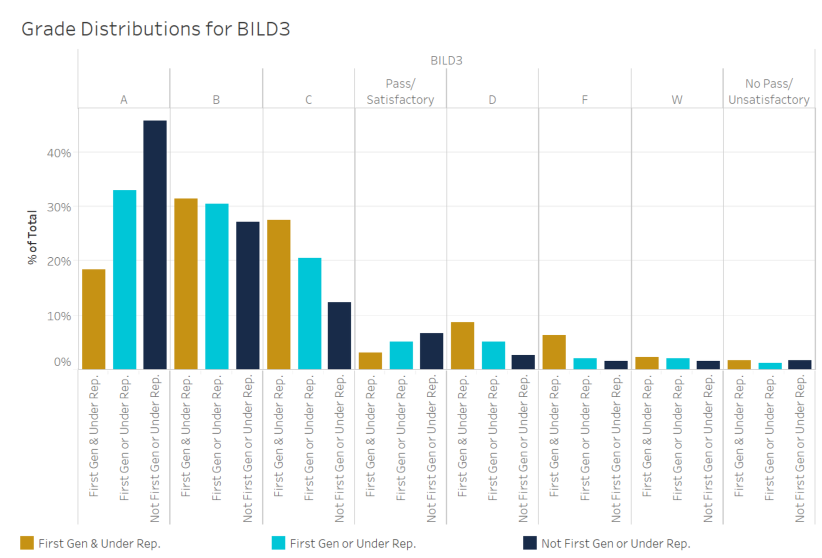
Keefe Reuther
kdreuther@ucsd.edu
University of California, San Diego
School of Biological Sciences
Department of Ecology, Behavior & Evolution

Intentionally Setting More Equitable Grading Policies: A Machine Learning Approach

The Issue

First-generation and racially minoritized undergraduate students typically face higher rates of Ds, Fs, and Withdrawals (DFWs) and lower average grades (López, Santelices, & Taveras, 2023). These disparities are often attributed to a complex interplay of historical and deeply rooted systemic policies and social structures (Feldman, 2023). Addressing and rectifying these inequities is a significant concern in Discipline-Based Education Research (DBER) today (Hsu et al., 2021). I suggest that optimizing the weighting of assessments in final grade calculations could mitigate the extent of these disparities within grade distributions. To the author's knowledge, this is the first instance of grading policies being mathematically optimized to minimize grading inequity.



The Assumption

For instructors, there is an acceptable range of graded assessment weights, dropped assessments, and minimum grades that accomplish their pedagogical goals. For example, few instructors would weight a final exam as 1% or 99% of students' final grades. *However, the exact parameters chosen (e.g., the final exam is 30% of the grade) is likely not an evidence-based, optimized decision.*

The Goal

Provide instructors with clear methods to lower grading inequities in the classroom by modification to grading policies, by providing optimized assessments weights, number of dropped assessments, and minimum scores.

Methods

1. Collect demographic and grading data
- Canvas survey. Usually hard to receive from institution.
- Does not require IRB approval if data is only used internally for educational purposes.

2. Choose your variables
- How will you divide students demographically?
- What metric will you try to minimize?

3. Clean gradebook data
- Convert exported Canvas gradebook into a csv file with columns representing demographic information and assessment grades.
- Convert scores to proportions (range of 0-1)

4. Choose your approach
- Solution 1: Machine learning optimization (upper right of poster)
- Solution 2: Interactive web application (lower right of poster)

5. Amend grading policies
- Use optimized grading policies to assign student grades.
- Make sure to renormalize grade cut-offs to prevent grade inflation.

Results & Directions

All results are based on demographic and grade data collected from a lower-division course, Organismic and Evolutionary Biology (BILD 3), at UCSD in the Spring of 2023. Below are some general findings from the application of both Solution 1 and Solution 2 to the course data:

- Grade inequity decreases when summative exams are weighted less in the final grade.
- Grade inequity decreases when small- and mid-level assessments are given more weight.
- Stricter lecture attendance policies have been associated with reduced grade inequity.
- Minimum scores had little effect on median grade differences; however, they likely significantly influence grade inequity in D/F/W rates (Feldman, 2023).

These conclusions may not reflect broader patterns. A more extensive and varied set of course data is necessary to effectively use this technique to identify best practices.

Many thanks to…

Members of the School of Biological Sciences Assessment Committee: Janni Pedersen, Melinda Owens, Jim Wilhelm, Stanley Lo, Liam O'Connor Mueller, Sarah Stockwell, and Laurie Smith. The wonderful community of teaching faculty at UCSD, T3PN, and SABER for giving me the support to do what brings me joy.

SOLUTION #1: Apply a machine learning optimization function to find the grading policies

Optimization functions operate by identifying the optimal mix of parameters—such as grade weights, dropped assessments, and minimum grades—that result in the smallest calculated value for the median grade difference between student groups (Bilal et al., 2020; Ronkkonen et al., 2005)

Define boundaries of pedagogically acceptable weights

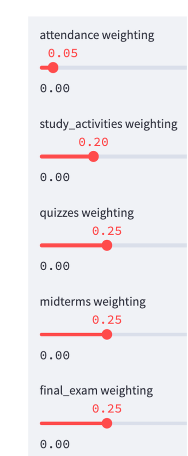
```
# Define custom weight bounds for each assignment group
custom_weight_bounds = {
    "attendance": (0.01, 0.15),
    "study_activities": (0.1, 0.3),
    "quizzes": (0.1, 0.4),
    "midterms": (0.2, 0.6),
    "final_exam": (0.15, 0.5)
}

# Define custom drop bounds for each assignment group
custom_drop_bounds = {
    "attendance": (0, 8),
    "study_activities": (0, 4),
    "quizzes": (0, 3),
    "midterms": (0, 1)
}
```

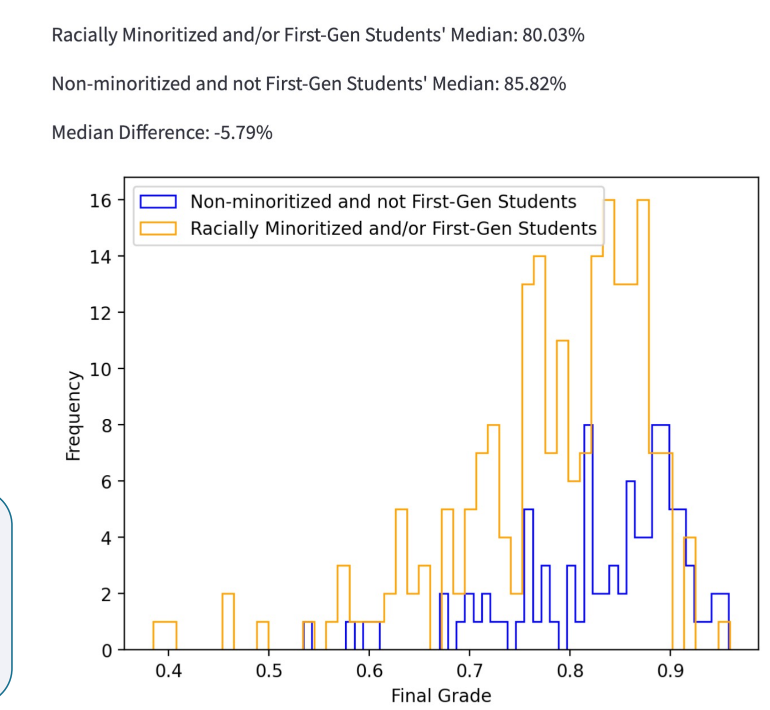
Define machine learning optimization function: differential evolution

```
# Perform the optimization using differential evolution
result = differential_evolution(
    constrained_extended_objective,
    bounds,
    args=(df, assignment_groups),
    strategy='best1bin',
    popsize=15,
    workers=num_cores - 1,
    tol=0.01,
    # maxiter= none,
    init='latinhypercube',
    mutation=(0.5, 1),
    recombination=0.7)
```

Before optimization

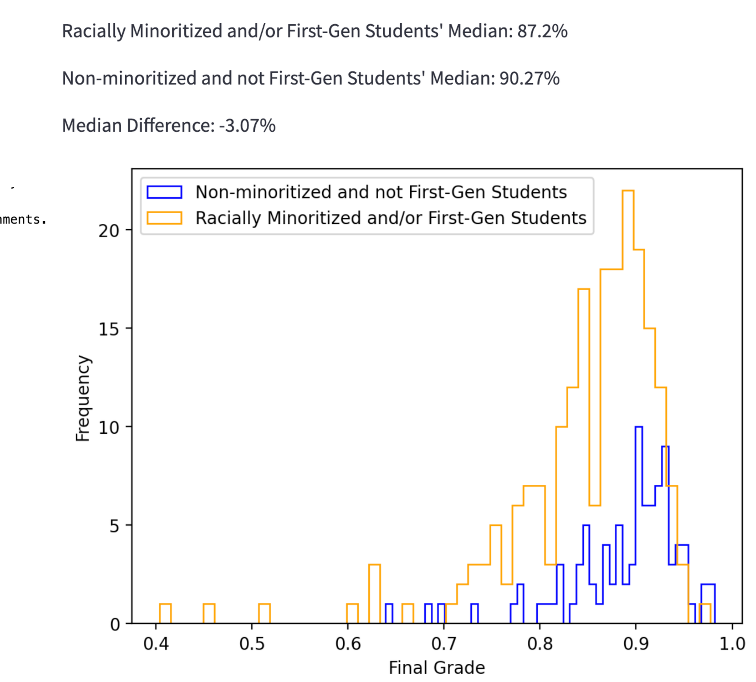


No dropped assessments or grade minimums



After optimization

```
Optimization complete.
Extracted optimized weights and number of dropped assignments.
Optimized Weights:
attendance: 0.14980499738586497
study_activities: 0.2990414793058371
quizzes: 0.20070249313280317
midterms: 0.20017372427424096
final_exam: 0.1502773009009463
Total Sum of Optimized Weights: 0.9999999949996925
Optimized Dropped Scores:
attendance: 1
study_activities: 4
quizzes: 3
midterms: 1
Optimized Minimum Scores:
attendance: 0.09291509170548955
study_activities: 0.10881478011574841
quizzes: 0.048749851928674356
midterms: 0.16948180662025178
final_exam: None
Minimum Median Difference: 0.03054298599550842
Script execution completed.
```



Please note that because optimization may result in grade inflation, it is essential to renormalize grade cutoffs to ensure that the distribution of A, B, C, D, and F grades remains consistent.

SOLUTION #2: An interactive web app that shows how changing grading policies affects grading inequity. Try it out using the the QR code below!

Are you interested in applying this concept with your own data? You can access all the code at https://github.com/keefereuther/grading_equity. Alternatively, you can email me at kdreuther@ucsd.edu for more information.

Note: the grading data in the app is NOT REAL to protect student privacy. It is only a proof of concept to show what you can do with your own data.

References:
- López, María José, Maria Veronica Santelices, and Carmen Maura Taveras. "Academic Performance and Adjustment of First-Generation Students to Higher Education: A Systematic Review." *Cogent Education* 10, no. 1 (December 31, 2023): 2209484.
- Feldman, J. (2023). *Grading for Equity: What It Is, Why It Matters, and How It Can Transform Schools and Classrooms.* Corwin Press.
- Ronkkonen, J., Kukkonen, S., & Price, K. V. (2005). Real-parameter optimization with differential evolution. *2005 IEEE Congress on Evolutionary Computation, 1*, 506-513 Vol.1
- Bilal, Pant, M., Zaheer, H., Garcia-Hernandez, L., & Abraham, A. (2020). Differential Evolution: A review of more than two decades of research. *Engineering Applications of Artificial Intelligence, 90*, 103479.
- Hsu, J. L., Chen, A., Cruz-Hinojoza, E., Dinh-Dang, D., Roth-Johnson, E. A., Sato, B. K., & Lo, S. M. (2021). Characterizing Biology Education Research: Perspectives from Practitioners and Scholars in the Field. *Journal of Microbiology & Biology Education, 22*(2), 10.1128/jmbe.00147-21.

This poster was edited for clarity and conciseness using ChatGPT 4.0. The ideas, opinions, and facts presented are my own, for which I take full responsibility.