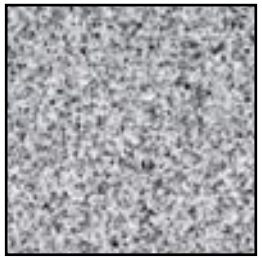
UNIVERSITY OF GEORGIA
Franklin College of Arts and Sciences

Investigating the impact of post-exam grade-adjustment practices in introductory physics

Nicholas T. Young[1], Rebecca L. Matz[2], Eric F. Bell[3], Caitlin Hayward[2]
[1] Department of Physics and Astronomy, University of Georgia
[2] Center for Academic Innovation, University of Michigan
[3] Department of Astronomy, University of Michigan



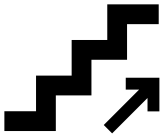
Georgia Physics and Astronomy Education Research
Franklin College of Arts and Sciences
UNIVERSITY OF GEORGIA

CENTER FOR ACADEMIC INNOVATION
UNIVERSITY OF MICHIGAN

We thank Erin Murray for her assistance in accessing and preparing the data for this project.

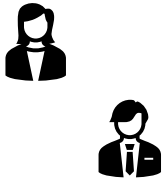
Grades matter

Financial Aid



Majors

Graduation



Careers

Ost 2010; Simmons and Heckler 2020

Grades have patterns of inequities

Matz et al. 2017, Whitcomb, Cwik, and Singh 2021, Castle et al. 2024

Inequities as a result of course structure

"We conclude that there are likely systemic biases, in introductory physics classes, that act against some underrepresented demographic groups. These biases are easily seen by comparing outcomes between different systems of teaching and assessment and these biases ***can likely be removed with appropriate structural changes at the level of a course*** *that—importantly—do not impact the educational standards of the course."* – Webb & Paul

Webb and Paul, 2023

Evidence of structural changes

- Reduce weight of exams in the final grade to reduce DFW rate (Simmons and Heckler 2020)

Evidence of structural changes

- Reduce weight of exams in the final grade to reduce DFW rate (Simmons and Heckler 2020)
- Grading directly on 4.0 scale rather than percent scale could reduce equity gaps by 20-25% (Paul and Webb 2022)

Evidence of structural changes

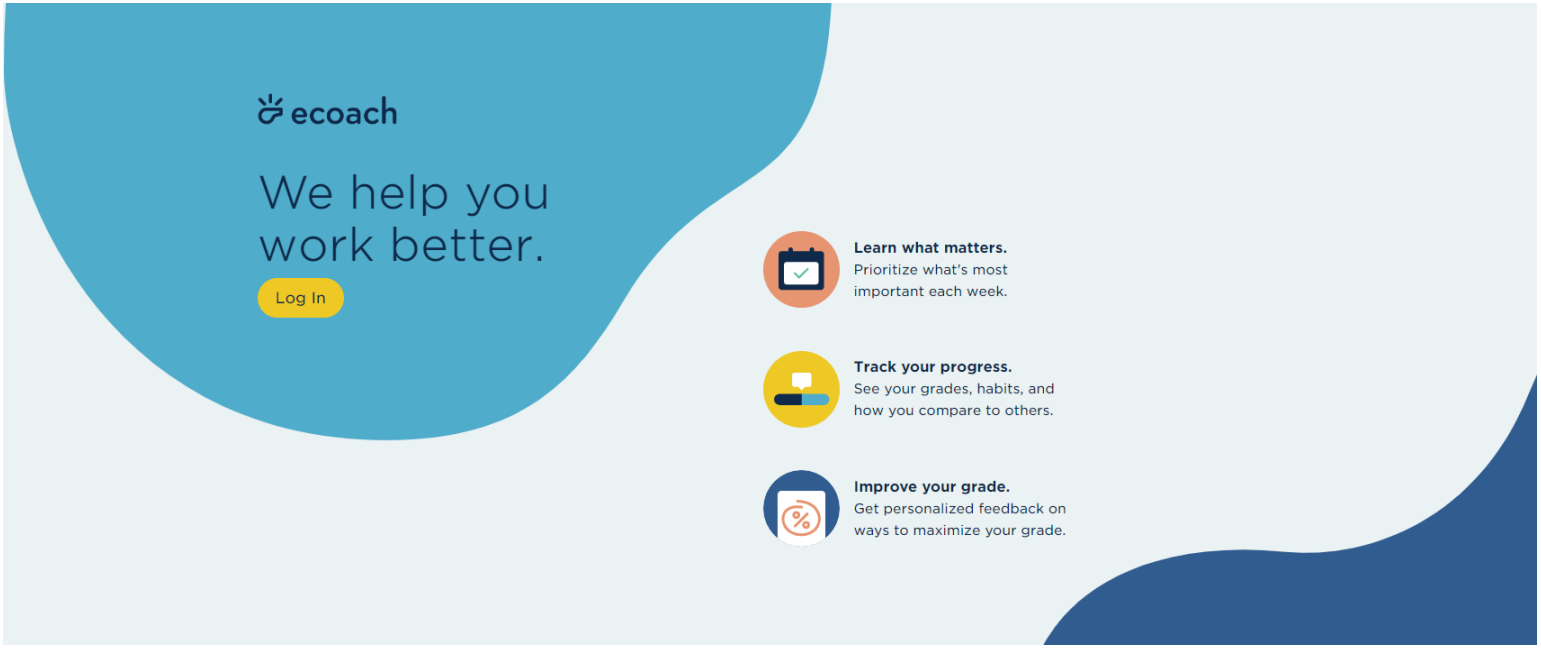
- Reduce weight of exams in the final grade to reduce DFW rate (Simmons and Heckler 2020)
- Grading directly on 4.0 scale rather than percent scale could reduce equity gaps by 20-25% (Paul and Webb 2022)
- Optional exam retakes may exacerbate score disparities between students with different social identities (Supriya et al. 2024)
- What about total exam grades?

Evidence of structural changes

- Reduce weight of exams in the final grade to reduce DFW rate (Simmons and Heckler 2020)
- Grading directly on 4.0 scale rather than percent scale could reduce equity gaps by 20-25% (Paul and Webb 2022)
- Optional exam retakes may exacerbate score disparities between students with different social identities (Supriya et al. 2024)
- What about total exam grades?

How does changing the weight each exam has in the final grade affect equity gaps?

Data: ECoach



https://ecoach.ai.umich.edu

Data: ECoach

- Winter 2017 – Fall 2019
- Grades:
  - Midterm exam 1
  - Midterm exam 2
  - Midterm exam 3
  - Final exam

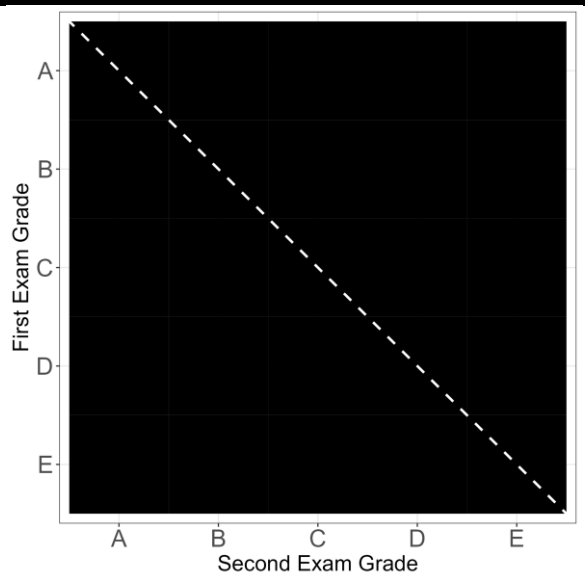
Data: ECoach

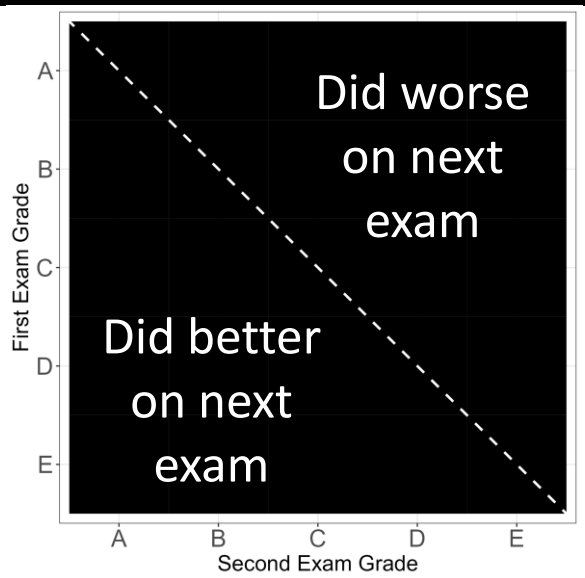
- Winter 2017 – Fall 2019
- Grades:
  - Midterm exam 1
  - Midterm exam 2
  - Midterm exam 3
  - Final exam
- 2290 Students

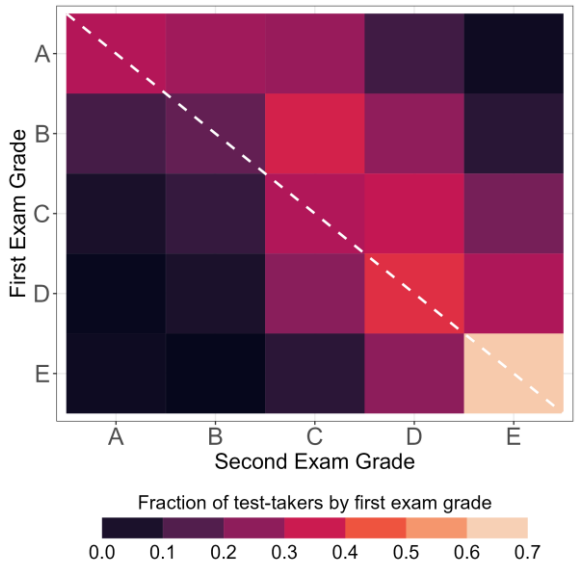
Data: ECoach

- Winter 2017 – Fall 2019
- Grades:
  - Midterm exam 1
  - Midterm exam 2
  - Midterm exam 3
  - Final exam
- 2290 Students

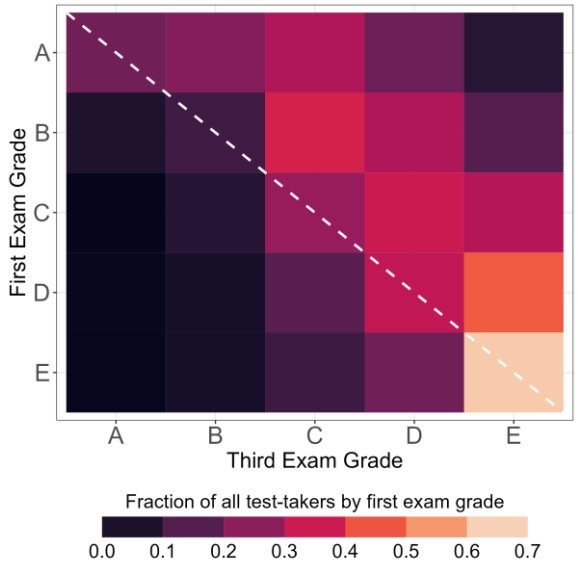
- Exams make up 36% of final grade
- Final exam makes up 16% of final grade

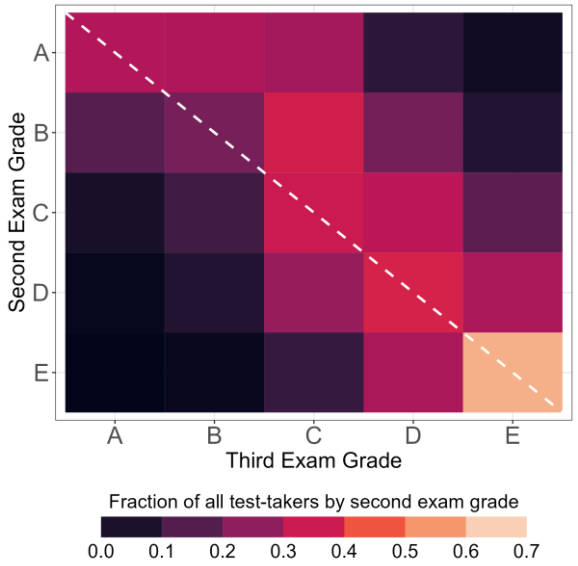

A
B
C
D
E
First Exam Grade
A
B
C
D
E
Second Exam Grade


Did worse on next exam
Did better on next exam
First Exam Grade
A
B
C
D
E
Second Exam Grade



Most students earn similar grades across exams (or do worse)





Nicholas Young

How could we determine total exam score?

- Average all three exam scores (default)

How could we determine total exam score?

- Average all three exam scores (default)
- Modifications:
  - Drop the lowest exam score

How could we determine total exam score?

- Average all three exam scores (default)
- Modifications:
  - Drop the lowest exam score
  - Replace the lowest exam score with the final exam score if the student does better on the final

How could we determine total exam score?

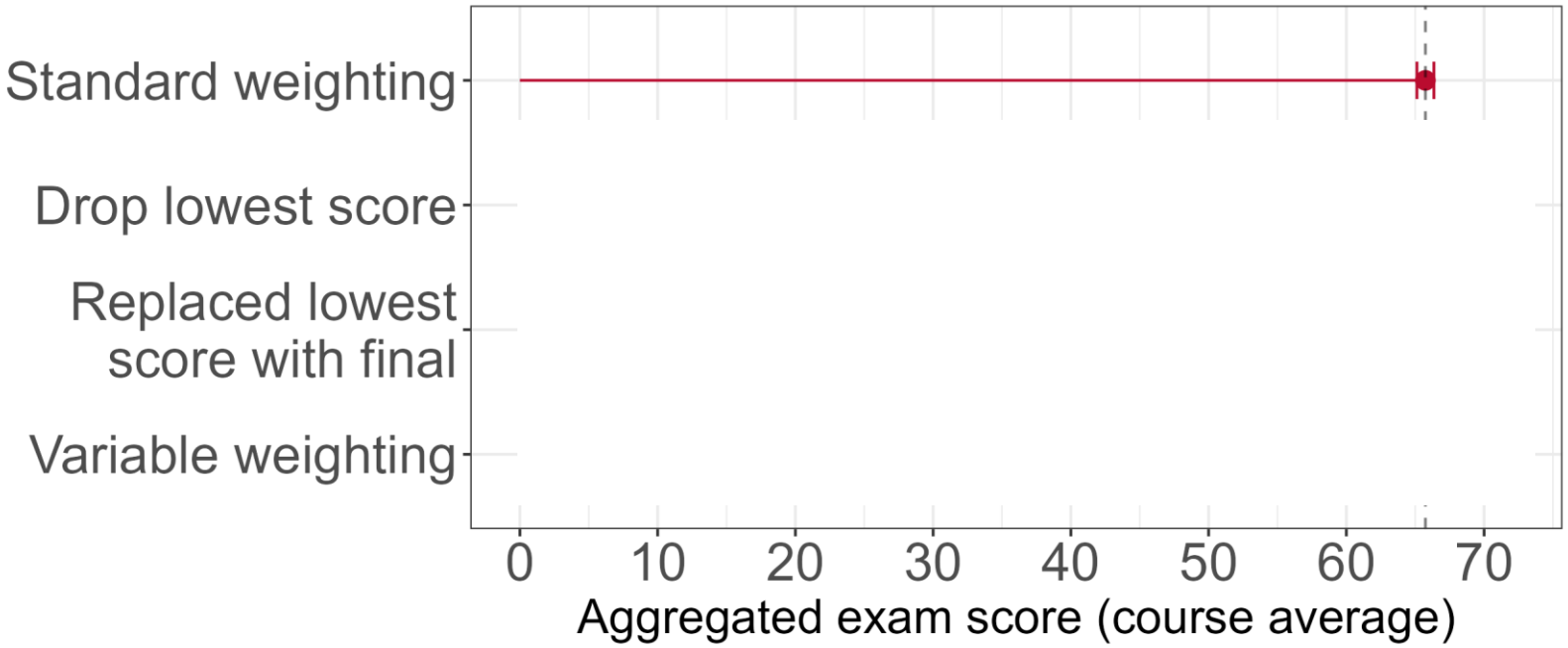
- Average all three exam scores (default)

- Modifications:
  - Drop the lowest exam score
  - Replace the lowest exam score with the final exam score if the student does better on the final
  - Have the best exam score count for 50% more and have the worst exam score count for 50% less

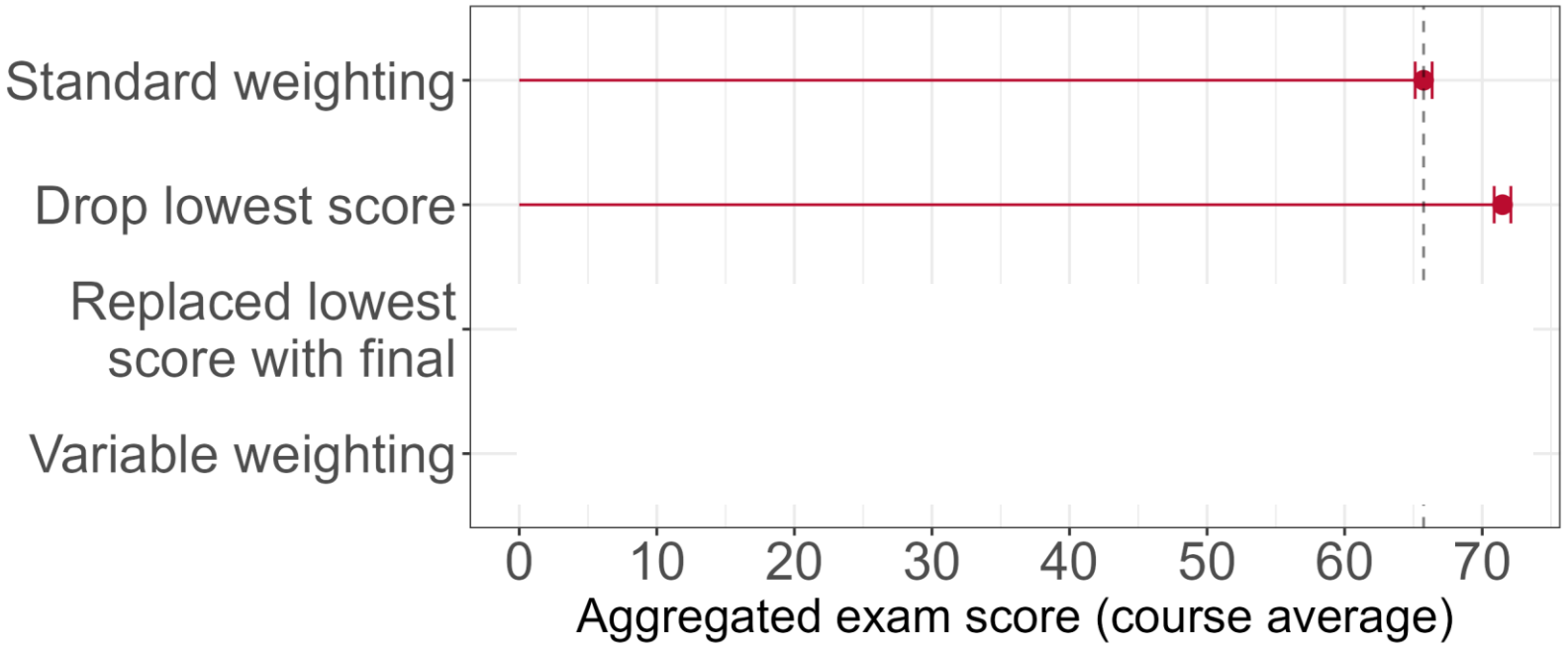
How could we determine total exam score?

- Average all three exam scores (default)
- Modifications:
  - Drop the lowest exam score
  - Replace the lowest exam score with the final exam score if the student does better on the final
  - Have the best exam score count for 50% more and have the worst exam score count for 50% less
- Let's apply these to our data!

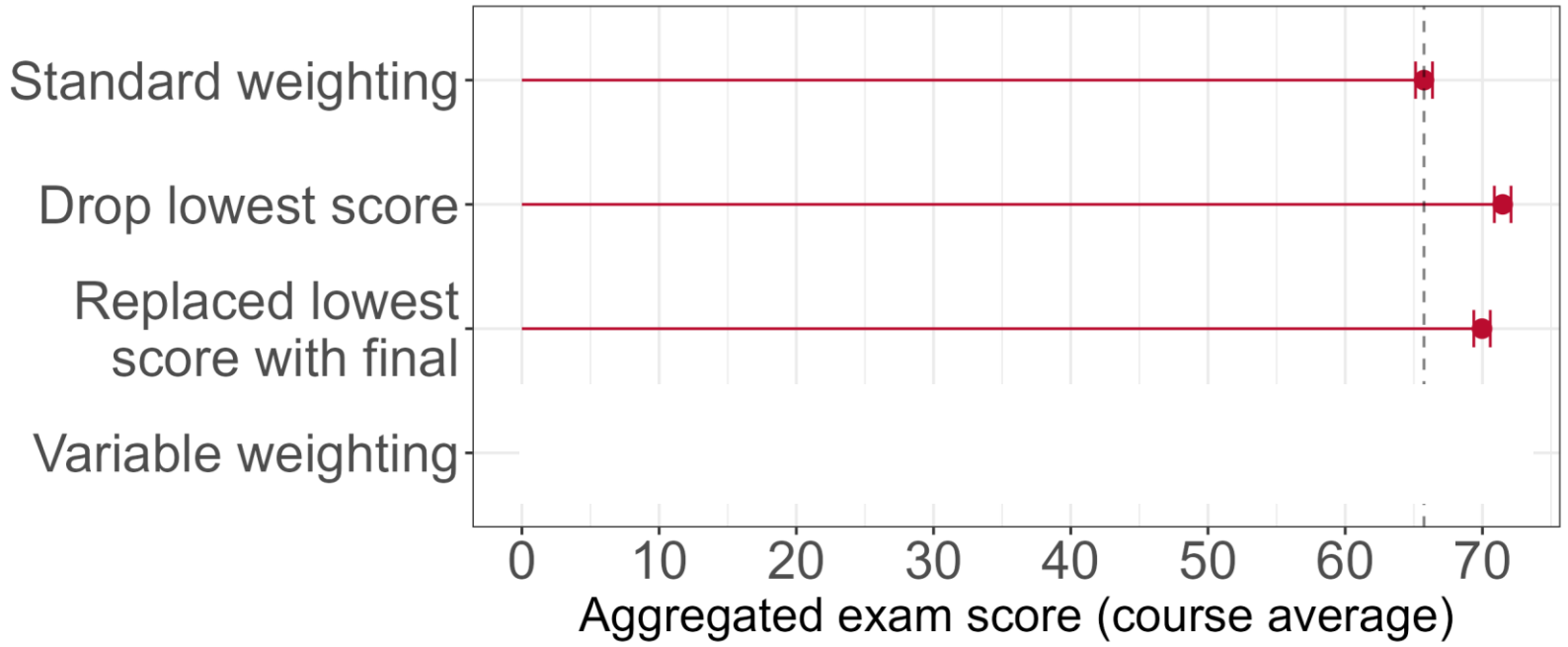
Average total exam score with no modifications



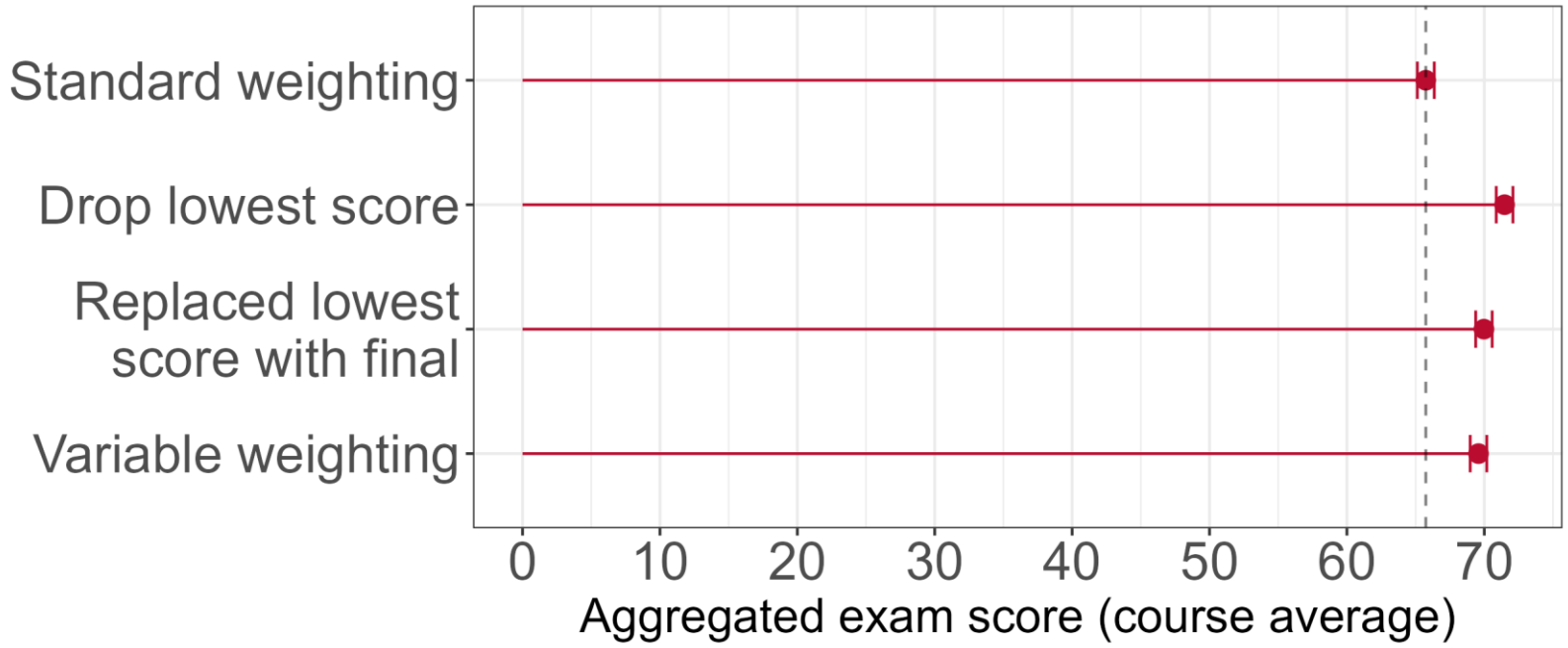
Increase in exam scores



Also increase, but not as much



Also increase, but again, not as much



Dropping the lowest exam score increased the total exam score the most.

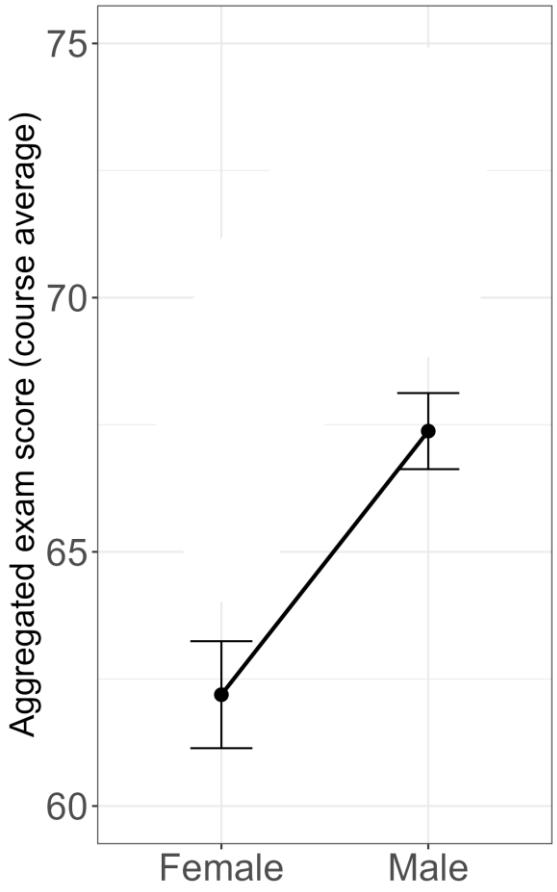
Dropping the lowest exam score increased the total exam score the most.

Replacing the lowest exam score with the final exam score and weighing each exam differently in the total exam score also increased the total exam score, but not as much.

What about equity though?

Salehi et al (2019)

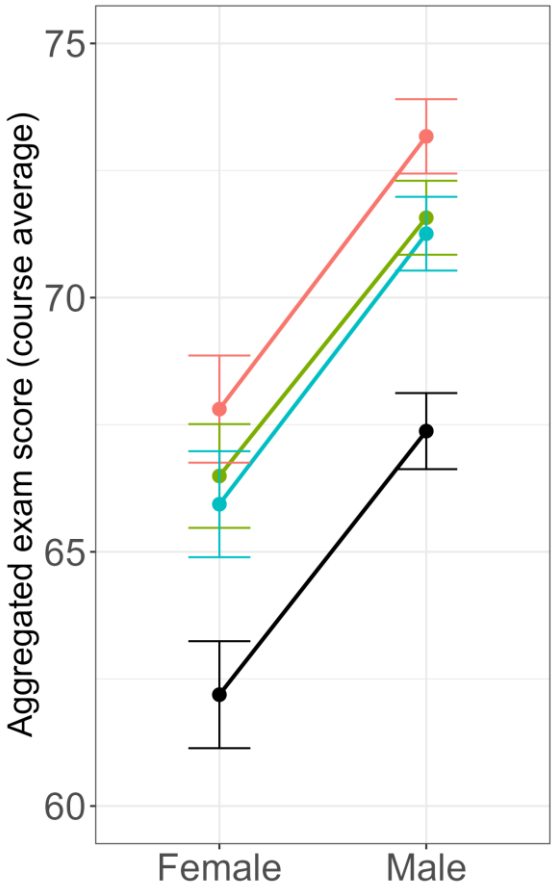
Nicholas Young



Standard weighting

Nicholas Young

None of the modifications seem to close the equity gaps.



Drop lowest score

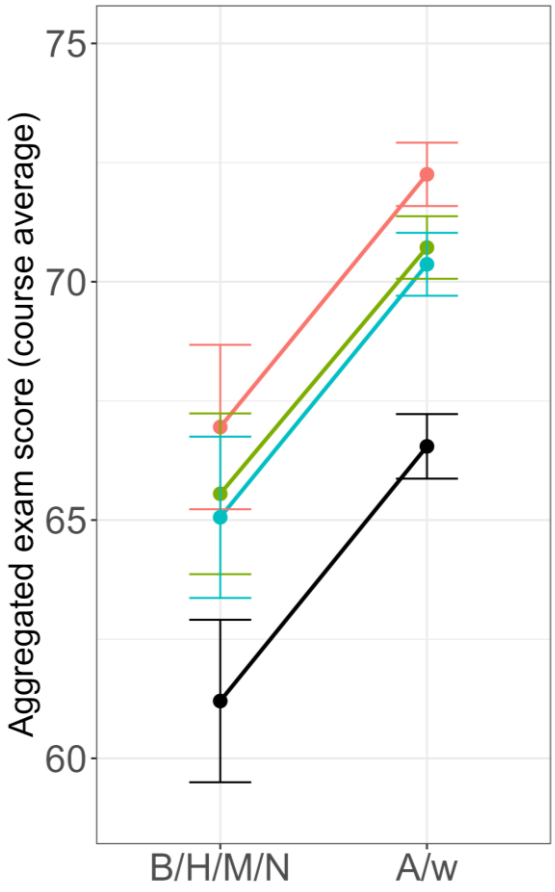
Replace lowest score with final

Variable weighting

Standard weighting

Nicholas Young

None of the modifications seem to close the equity gaps here either.



Drop lowest score

Replace lowest score with final

Variable weighting

Standard weighting

Nicholas Young

Takeaways

- Students tend to earn the same letter grade or a lower grade across exams.

Takeaways

- Students tend to earn the same letter grade or a lower grade across exams.
- Grading practices that reduce the impact of the lowest exam grade increase grades.
  - Dropping the lowest exam grade resulted in the greatest grade increase.

Takeaways

- Students tend to earn the same letter grade or a lower grade across exams.
- Grading practices that reduce the impact of the lowest exam grade increase grades.
  - Dropping the lowest exam grade resulted in the greatest grade increase.
- None of the three alternative grading practices affected equity gaps on exams though.

Takeaways

- Students tend to earn the same letter grade or a lower grade across exams.
- Grading practices that reduce the impact of the lowest exam grade increase grades.
  - Dropping the lowest exam grade resulted in the greatest grade increase.
- None of the three alternative grading practices affected equity gaps on exams though.
- contact: nicholas.young@uga.edu
- slides: bit.ly/