



bluenotes
GLOBAL 2022

Mitigating Implicit Bias in Student Evals: A Randomized Intervention

Brandon Genetin

US Department of Agriculture
Economic Research Service

Joyce Chen

Ohio State University

Vladimir Kogan

Ohio State University

Alan Kalish

Ohio State University

Roadmap

- **Background and motivation**
- **RCT embedded into actual Ohio State University student evaluation system**
- **Results**
 - Impact on racial and gender disparities
 - Impact on relationship between student grades and course evaluations
- **Implications**

Motivation: Bias in evaluations

- **Growing concern and evidence of disparities in how students evaluate course instructors**
 - Race, gender, non-native English speakers
- **Gaps have been discussed for many years, but recent evidence provides strong evidence of causality**

What's in a Name: Exposing Gender Bias in Student Ratings of Teaching

Lillian MacNell • Adam Driscoll • Andrea N. Hunt

The Teacher

Gender Bias in Student Evaluations

Kristina M. W. Mitchell, *Texas Tech University*

Jonathan Martin, *Midland College*

Motivation: Bias in evaluations

- **Can the survey instrument itself be redesigned to reduce bias?**
- **Important pilot study from Iowa State provided some reason for optimism**

RESEARCH ARTICLE

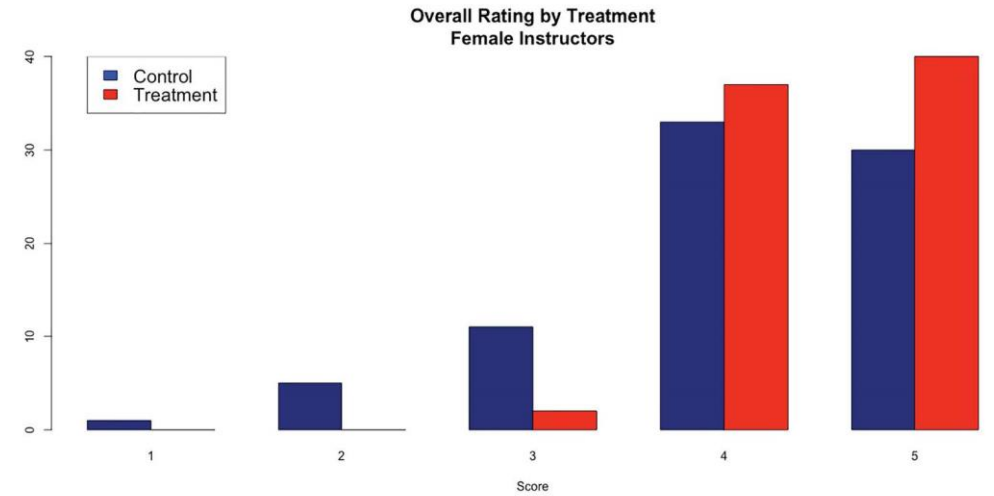
Mitigating gender bias in student evaluations of teaching

David A. M. Peterson^{1*}, Lori A. Biederman², David Andersen¹, Tessa M. Ditonto¹, Kevin Roe³

Motivation: Bias in evaluations

“Student evaluations of teaching play an important role in the review of faculty. Your opinions influence the review of instructors that takes place every year. Iowa State University recognizes that student evaluations of teaching are often influenced by students’ **unconscious** and **unintentional** biases about the race and gender of the instructor. Women and instructors of color are systematically rated lower in their teaching evaluations than white men, even when there are no actual differences in the instruction or in what students have learned.

As you fill out the course evaluation please keep this in mind and make an effort to resist stereotypes about professors. Focus on your opinions about the content of the course (the assignments, the textbook, the in-class material) and not unrelated matters (the instructor’s appearance).”



Motivation: Bias in evaluations

- **Iowa State study had significant limitations, however**
 - **Small sample sizes, limited to four biology and political science courses**
 - **Unclear if treatment changed how students rated instructors or just reduced participation among low raters**
 - **Combined two treatments: (1) priming bias and (2) increasing salience of stakes**
 - **If this works, do effects apply to other subgroups (e.g., race)?**

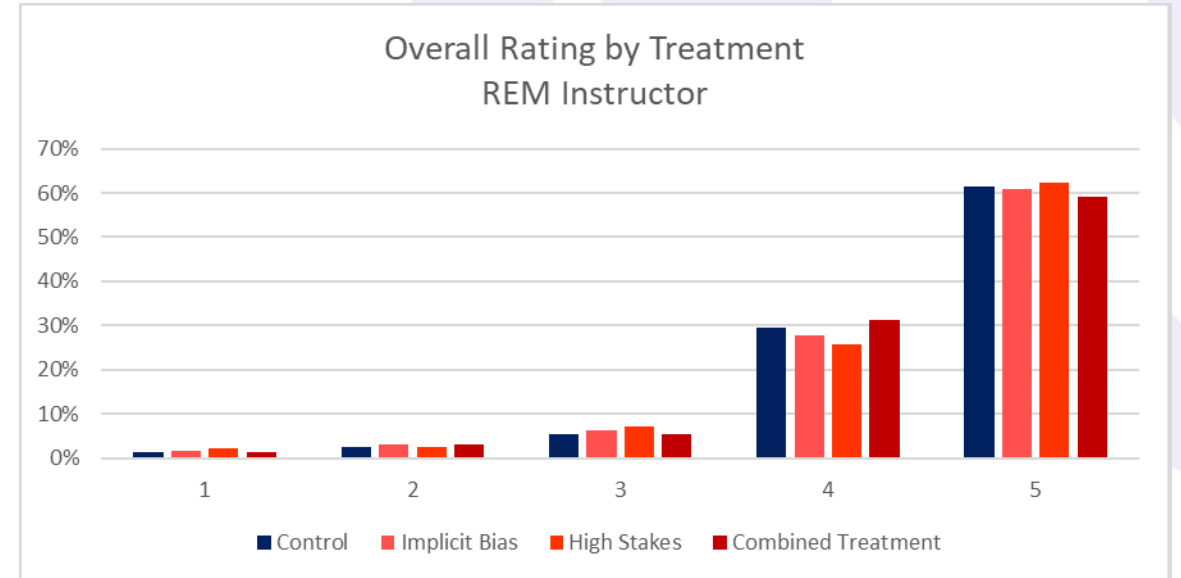
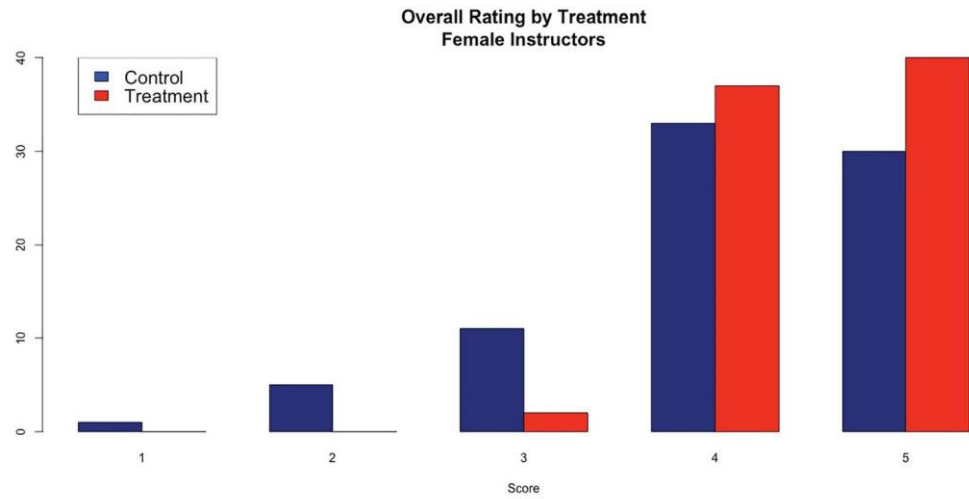
OSU RCT

- **Study design**
 - Run during Spring 2021
 - 400 participating instructors (16% of total), 700 classes
 - 19,000 evaluations by 14,000 unique students
 - (Modified) Within-class randomization
- **Implementation challenges**
 - Ensuring that students in multiple participating classes saw consistent version of survey
 - Examine both “intent to treat” and “treatment on treated,” due to technological limitations of mobile app

OSU RCT

- **Unique strengths**
 - Extensive administrative data on *both* student and instructor characteristics
 - Information on other factors that may influence evaluations (e.g., end-of-course grades)
- **Limitation**
 - Due to COVID context, nature of courses, significant share of online courses

Results

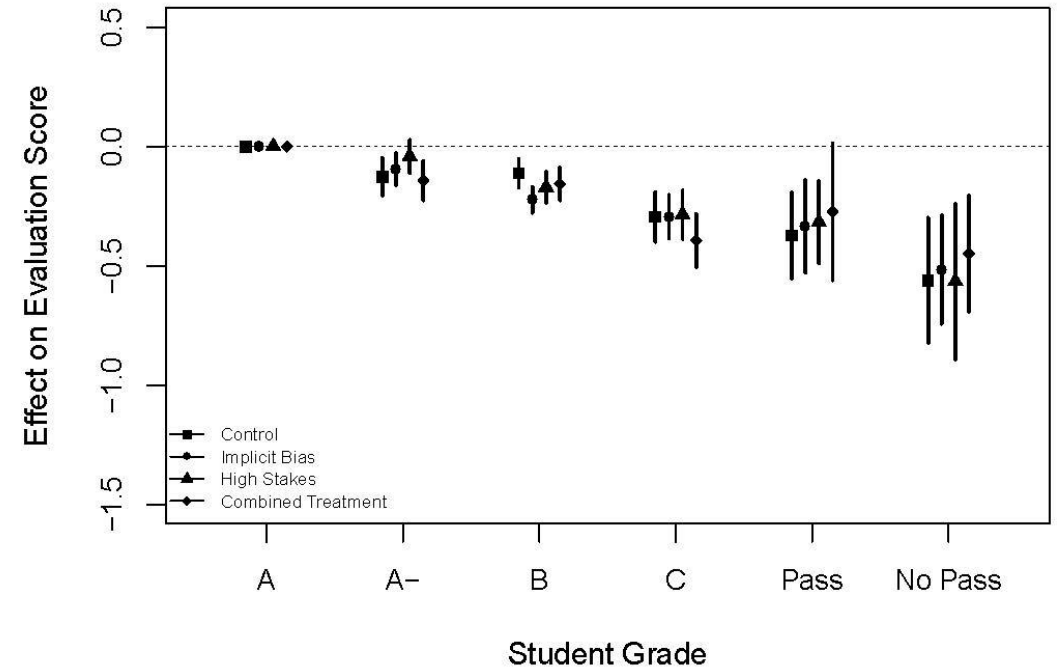
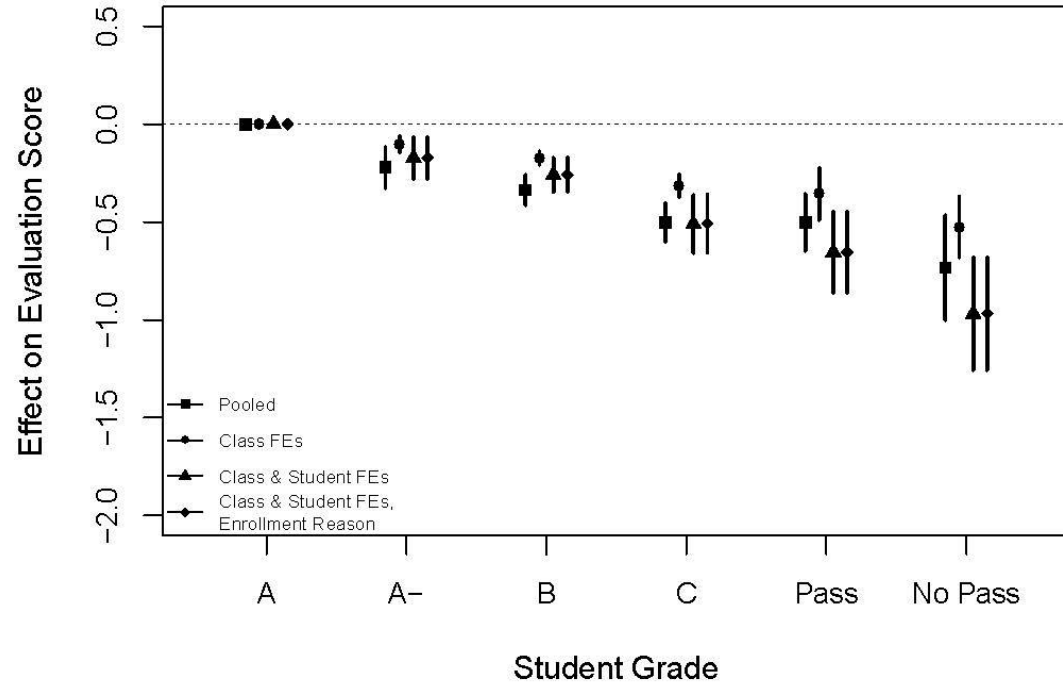


OSU RCT

- **Additional questions**

- **Since adoption, OSU research has found that course grades were strongest predictor of evaluations, but unclear implications**
- **Our data allows us to examine proposed mechanisms**
- **Can rule out “better teaching=higher grades” (via class fixed effects) and student sorting (via student fixed effects)**
- ***Did treatments reduce the impact of student grades on evaluations?***

OSU RCT



Implications and Discussion

- **Importance of replication before broadscale adoption!**
- **No cheap, easy, quick fix for bias concerns!**
- **Most current focus is on bias, but should remember broader worries about construct validity and incentives (grade inflation)**
 - **Not obvious that interventions that address one issue also address the other**
- ***Caveat: Current analysis focuses only on numerical scores, not open-ended comments***