

2024-10-26

# Implicit bias and RateMyProfessor

---

J. Zhu, C. Jia, S. Chen, H. Zhou, D. Mello-Goldner. 2024. "Implicit bias and RateMyProfessor"

<https://hdl.handle.net/2144/49760>

*"Downloaded from OpenBU. Boston University's institutional repository."*

## Abstract

This study explored the impact of similarities in gender and race between students and professors on students' rating of the professors' teaching quality. Using a simulated *RateMyProfessors.com* profile with and without AI-generated professor photos, we assessed student perceptions of teaching quality. Female students rated profiles higher than male students. Asian students rated all profiles lower. Profiles without photos were rated higher, pointing to visual cues influencing implicit biases. This research contributes to the understanding of how implicit biases may shape educational outcomes and emphasizes the importance of equitable practices in academia.

## Introduction

Annamma and Morrison (2018) emphasized that marginalized groups often face perceptions of being problematic, leading to disparities in resource allocation and classroom dynamics. Addressing implicit biases is crucial for promoting inclusivity and equity, as these biases can significantly impact academic opportunities and outcomes (Staats, 2016). Our hypothesis was that students will be more likely to assign higher ratings when interacting with professors who share similar attributes, such as gender or the same racial background.

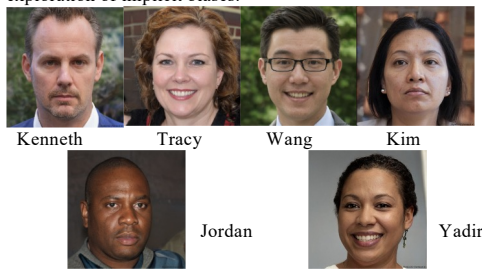
## Method

### Participants

Eighty-six students at Boston University.

### Materials and Procedure

Participants received instructions explaining the study's intent to assess perceptions of teaching quality and class engagement after viewing (fictional) *RateMyProfessor.com* ratings, half of which included AI-generated photos. After giving informed consent, they reviewed six profiles, with or without photos, and answered questions related to each professor, followed by demographic questions, including gender, race, and class year. When done they were debriefed about the study's true purpose emphasizing the exploration of implicit biases.



Kenneth Tracy Wang Kim  
Jordan Yadira

## Results

### Ratings by Gender of Participants

Female participants rated the profiles higher ( $M=4.01$ ) than male participants ( $M=3.76$ ),  $F(1, 84)=3.98, p=.049, \eta_p^2=.05$ . There also was a main effect for profiles,  $F(5, 420) = 3.98, p=.049, \eta_p^2=.05$ . Bonferroni post hoc tests showed Yadira was rated higher than all other profiles besides Wang, and Kenneth was rated lower than Wang and Yadira. There was no interaction of gender of participants and profiles,  $F(5, 420) = .63, p = .676, \eta_p^2 = .01$  (see Table 1).

**Table 1**  
Mean Ratings for Profiles by Participant Gender

| Gender of Participants | Profiles |        |        |       |        |        |
|------------------------|----------|--------|--------|-------|--------|--------|
|                        | Kenneth  | Tracy  | Wang   | Kim   | Jordan | Yadira |
| Male ( $n=18$ )        | 3.63     | 3.83   | 3.83   | 3.67  | 3.76   | 3.84   |
| Female ( $n=68$ )      | 3.84     | 4.01   | 4.13   | 3.93  | 3.93   | 4.19   |
| Marginal Means         | 3.79c    | 3.97bc | 4.07ab | 3.88c | 3.90c  | 4.12a  |

Note. Max. rating = 5.00. Marginal means with different subscripts are significant at the  $p<.05$  level. Means in color reflect a match between participants' and professors' gender.

### Ratings by Race of Participants

A mixed two-way ANOVA found a significant main effect for rating by race of participants,  $F(3,45)=5.08, p=.004, \eta_p^2=.25$ . Post hoc tests found Asian participants ( $M=3.65$ ) rated profiles significantly lower than Hispanic/Latinx ( $M=4.42$ ), White ( $M=4.01$ ), and Black/African American participants ( $M=4.09$ ). There was also a significant main effect by profile identity,  $F(5, 225)=3.91, p=.002, \eta_p^2=.080$ . The result of the post hoc LSD test indicated that Yadira's profile was rated significantly higher than all other profiles except Wang. There was no interaction,  $F(15, 225)=.84, p=.638, \eta_p^2=.053$  (see Table 2).

**Table 2**  
Means Ratings By Participants' Race

| Profile             | Race  |        |       |          |        | Profile Marginal Means |
|---------------------|-------|--------|-------|----------|--------|------------------------|
|                     | White | Black  | Asian | Hispanic | Others |                        |
| Kenneth             | 3.85  | 4.10   | 3.35  | 4.25     | 3.44   | 3.63b                  |
| Tracy               | 4.15  | 3.90   | 3.70  | 4.52     | 4.03   | 3.92b                  |
| Wang                | 4.15  | 4.18   | 3.89  | 4.48     | 3.97   | 4.04ab                 |
| Kim                 | 3.69  | 3.97   | 3.56  | 4.40     | 3.69   | 3.70bc                 |
| Jordan              | 3.96  | 3.97   | 3.58  | 4.21     | 4.05   | 3.80c                  |
| Yadira              | 4.27  | 4.44   | 3.83  | 4.58     | 4.49   | 4.10a                  |
| Race Marginal Means | 4.10a | 4.09ab | 3.65b | 4.42a    | 3.94   |                        |

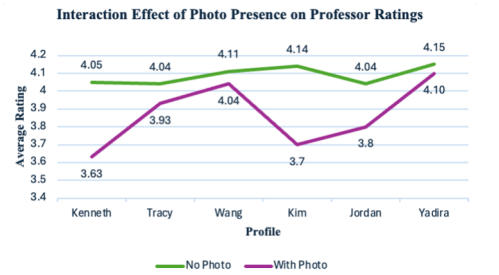
Note. Max. rating=5.00. Only participants who received photos were included ( $N=52$ ). The "Other" racial group was not included in the analysis. Means in color reflect a match between the race of the participant and the professor.

### Ratings by Photo Presence

A mixed two-way ANOVA was conducted to examine the effect of photo presence on students' average ratings for professors. There was a significant main effect on average ratings by profile photo presence,  $F(1, 84)=4.80, p=.031, \eta_p^2=.05$ .

Profiles received a higher rating when no photo was present ( $M=4.09$ ) compared to when a photo was included ( $M=3.87$ ). Also, there was also a significant main effect on average ratings by different profile photos,  $F(5, 420)=7.98, p<.001, \eta_p^2=.08$ . Tukey post hoc test revealed that Kenneth's profile ( $M=3.84$ ) was rated significantly lower than Wang's ( $M=4.07$ ), while the average ratings of Yadira's profile ( $M=4.13$ ) was higher than Kenneth's, Tracy's ( $M=3.98$ ), Kim's ( $M=3.92$ ), Jordan's ( $M=3.92$ ). A significant interaction found that when no photo was present, Yadira's profile ( $M=4.13$ ) received the highest average ratings, but when photos were included, Tracy's profile with a photo ( $M=4.00$ ) had a higher average rating compared to Kenneth's ( $M=3.84$ ), and Wang's profile ( $M=4.08$ ) is rated positively regardless of photo presence,  $F(5, 420)=5.46, p<.001, \eta_p^2=.06$  (see Figure 1).

**Figure 1**  
Interaction Effect of Photo Presence on Professor Ratings



## Discussion

Our study found bias impacting perceived teaching quality and interactions with professors. While students didn't show a strong preference for professors of the same gender, they generally viewed female professors more positively. Contrary to our hypothesis, students didn't consistently rate their experiences higher when interacting with professors of similar racial backgrounds. Surprisingly, we discovered that the presence of a professor's photo influenced average ratings, highlighting the role of implicit biases in shaping perceptions and evaluations, especially if the person was SMILING.

The use of convenience sampling with a small sample size (101 respondents) and a skewed demographic representation (predominantly Asian and White, majority female) may affect the generalizability of our findings. Future research should employ larger, more diverse samples using random sampling methods to address these limitations and ensure more robust and applicable results.

This research is essential for colleges aiming to create an inclusive and equitable educational environment. By uncovering the impact of racial and gender biases on student evaluations of professors, colleges can identify and address biases within their evaluation systems.