

Final Paper: Analyzing Discrimination in Workplace Recruitment

Zeying Guo

May 15th, 2024

DACSS 690E

Abstract

This research experiment is designed to explore the impact of gender markers in resumes on hiring decisions to uncover any biases. In this experimental design, we modified resumes with names that sound White or Black in male/female gender hints to assess how these elements influenced evaluations of qualifications and employment prospects. The analysis of data from 102 individuals revealed no differences in assessments based on these markers. This experimental findings suggest that racial and gender cues do influence hiring choices as resumes with names tended to receive ratings for qualifications and chances of being hired. Removing information from resumes helped reduce bias indicating that omitting details could promote fairer hiring practices.

Introduction

Bias and discrimination have always been a topic that people talk about. In recent years, there has been increased focus on bias and discrimination in recruitment processes. Subtle biases against groups often unconscious can lead to unequal opportunities and perpetuate systemic inequalities in the workplace. This research is designed to investigate whether gender or race

cues in resumes impact hiring decisions. The central question is "Do gender indicators in resumes affect how candidates are assessed for qualifications and likelihood of being hired?"

Recognizing the presence and scope of these biases is essential for fostering equitable recruitment procedures. Identifying discrimination in the hiring process allows us to shape policies that support diversity and inclusion in the workplace. This experiment research project is driven by real world events and ongoing societal conversations surrounding equality making it a timely and pertinent study.

Literature Review

Previous research conducted by Bertrand and Mullainathan (2004) illustrated bias in hiring practices, where resumes with white sounding names received more callbacks compared to those with black sounding names. Their field experiment showed that applicants with sounding names had a 50% chance of receiving a callback than those with black sounding names despite having identical qualifications.

Recent studies like the one by Kline, Rose and Walters (2024) have further emphasized existing biases by revealing that white female names are favorably evaluated over names. This particular study sent out resumes with varying gender cues to employers showing that white male and female names received the most callbacks while black male and female names received the least.

Moreover, research indicates that biases can vary across industries. Gassam Asare (2024) observed that name discrimination was particularly pronounced in sectors such as the industry.

An approach to address biases in recruitment involves anonymizing resumes by omitting names and demographic details as proposed by Gassam Asare in 2024. Various research studies have also investigated the efficacy of this method, in minimizing bias. For instance findings indicate

that using anonymized resumes can lead to fairer hiring choices, indicating that eliminating data may help counteract implicit biases (Gassam Asare, 2024).

Hypotheses

H1: Resumes with Black-sounding or female-sounding names will receive lower ratings for qualifications and hiring likelihood compared to those with White-sounding or male-sounding names.

H2: Anonymizing resumes will result in reducing bias.

Theoretical Justification

These hypotheses were made to rely on the online research studies stated in Literature Review section, it is theorized that resumes bearing names associated with female identities may lower ratings for qualifications and hiring prospects in comparison to those linked with White or male identities. Additionally, the act of anonymizing resumes is expected to lead to ratings across all applications indicating a reduction in bias.

Experimental Design

Manipulation of Independent Variable

This experimental design was set up as a 2 x 2 design. The key focus of this research was the details included in resumes, specifically the racial and gender indicators conveyed through names. Except for the names shown on the resumes, all other parts in the resumes remain the same, which is designed to control in the experiment, no other factors would affect the result.

Resumes were varied to represent gender groups; white male, white female, black male, black

female and anonymized (no name). Additionally, a control group, with anonymized resumes was used to examine the impact of removing information. This approach helped us analyze how racial and gender cues influence evaluations of resumes addressing our inquiry into biases in hiring practices.

Measurement of Dependent Variables

The dependent variables measured were the assessment of qualifications and the likelihood of hiring both evaluated on a 5-point Likert scale. Participants were tasked with rating each candidates' qualifications and their probability of being hired based on the resume they reviewed. To ensure clarity on what candidates were being evaluated, there is an actual job description for an entry level software engineer position, from IBM was included before they review the resume. The specific survey questions posed were:

1. Qualification Rating; "On a scale of 1 5 how would you rate this candidates' qualifications?"

(On a scale of 1, to 5 with 1 being highly qualified and 5 being not all qualified)

2. How probable is it that you would hire this candidate? (From 2 as not likely to 6 as highly likely)

Sample

This experimental design research mainly focused on how racial and gender cues impact individuals while also taking into account how factors, like age, gender and ethnicity of participants could influence the results. By ensuring that these variables were evenly spread across all groups through statistics and balance tests we were able to evaluate their potential effects.

To gather participants, for this experiment study, we used a survey platform called Lucid Theorem and compensated each respondent with \$1.50 for completing the survey.

The sample consisted of 102 individuals, with demographic information as follows:

- **Gender:** 100% Female
- **Age:** Mean = 52.26, SD = 15.78
- **Ethnicity:** 76.47% White, 10.78% Black, 0.98% American Indian or Alaska Native, 0.98% Chinese, 3.92% Some other race, 0.98% Prefer not to answer, 0.98% Other
- **Hispanic Origin:** 91.18% No, 2.94% Mexican, 2.94% Prefer not to answer, 0.98% Nicaragua, 0.98% Spain, 0.98% Puerto Rican
- **Education:** 31.37% High school graduate, 18.63% Completed some college, but no degree, 17.65% bachelor's degree, 14.71% associate's degree, 8.82% Master's or professional degree, 3.92% Other post-high school vocational training, 2.94% Some high school or less, 1.96% None of the above
- **Household Income:** Varied, with the most common categories being less than \$14,999 (16.84%) and \$15,000 to \$19,999 (9.47%)
- **Political Party:** 20.59% Strong Democrat, 19.61% Independent - neither 14.71% Other - neither 14.71% Strong Republican

We can find from the percentage distribution in the demographic information analyze, that it is not balanced distributed. We are especially focusing the gender; 100% female were taking this survey that might influence the result heavily. The distribution, across categories such as ethnicity, Hispanic origin, education level, household income and political party affiliation shows

varying percentages indicating diversity, within the surveyed group. The survey may lack information potentially leading to decreased confidence and trust in the data analysis outcomes of this experiment.

Manipulation Check Questions

Participants were asked questions to confirm their perception of the intended gender signals.

They were inquired:

"Based on the resume what gender do you believe this applicant represents?" (Male, Female, Not Sure)

Debrief Statement

Thank you for participating in our research. The focus of this study was to examine the impact of gender/race indicators on resumes, on hiring decisions. We intentionally kept this purpose undisclosed. We hope that the findings from this study will promote hiring practices by highlighting the effects of biases. All participants' personal information is kept privacy for the personal information safety consideration.

Results

Qualification Ratings

The mean, max, min and standard deviation value of qualification ratings for each treatment group were as follows:

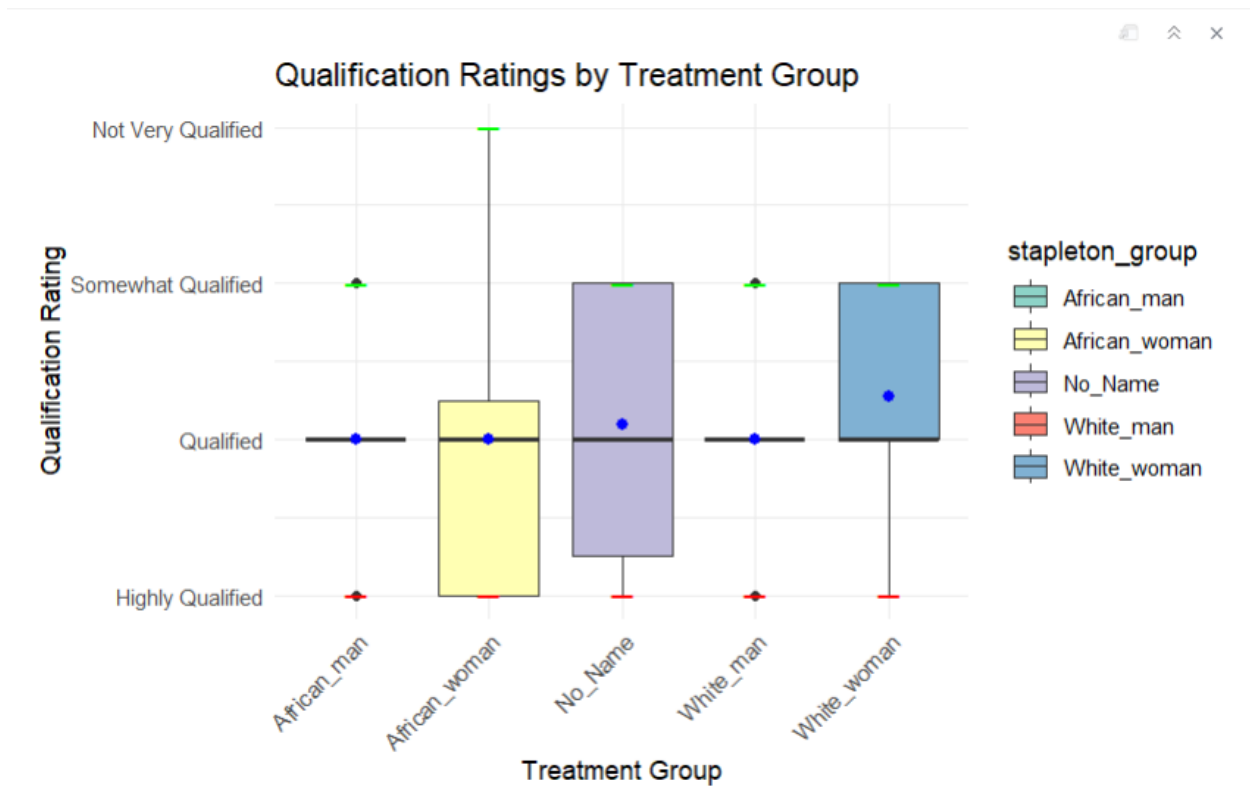
stapleton_group <chr>	mean_qual <dbl>	min_qual <dbl>	max_qual <dbl>
African_man	2.000000	1	3
African_woman	2.000000	1	4
No_Name	2.090909	1	3
White_man	2.000000	1	3
White_woman	2.272727	1	3
NA	NaN	Inf	-Inf

A tibble: 6 × 3

stapleton_group <chr>	mean_qual <dbl>	sd_qual <dbl>
African_man	2.000000	0.6488857
African_woman	2.000000	0.8944272
No_Name	2.090909	0.8111773
White_man	2.000000	0.6172134
White_woman	2.272727	0.7862454
NA	NaN	NA

6 rows

The boxplot for qualification ratings by treatment group shows the distribution of ratings for each group. Blue dots mean the mean value, green dots show the min value and red dots is where the max value stands.



As above graph and tables shown, African women is more likely to receive “Not very qualified” option for this job position, but all resumes received similar average rating for this job position. No obvious bias observed from the data visualizations.

Hiring Likelihood

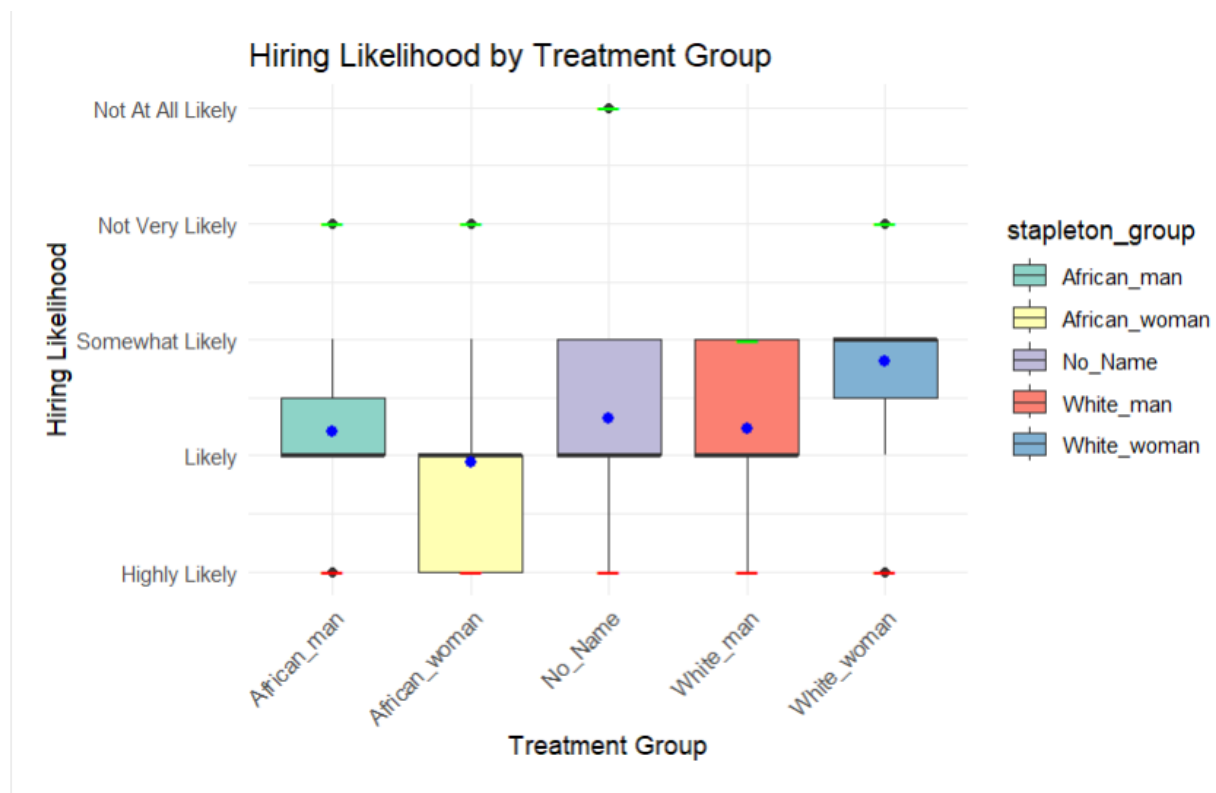
The mean, max, min and standard deviation value of hiring likelihood ratings for each treatment group were as follows:

stapleton_group <chr>	mean_hire <dbl>	min_hire <dbl>	max_hire <dbl>
African_man	2.210526	1	4
African_woman	1.937500	1	4
No_Name	2.318182	1	5
White_man	2.227273	1	3
White_woman	2.818182	1	4
NA	NaN	Inf	-Inf

stapleton_group <chr>	mean_hire <dbl>	sd_hire <dbl>
African_man	2.210526	0.7132825
African_woman	1.937500	0.8539126
No_Name	2.318182	0.9454837
White_man	2.227273	0.6853444
White_woman	2.818182	0.8738629
NA	NaN	NA

6 rows

The boxplot for hiring likelihood by treatment group shows the distribution of ratings for each group. Blue dots mean the mean value, green dots show the min value and red dots is where the max value stands.

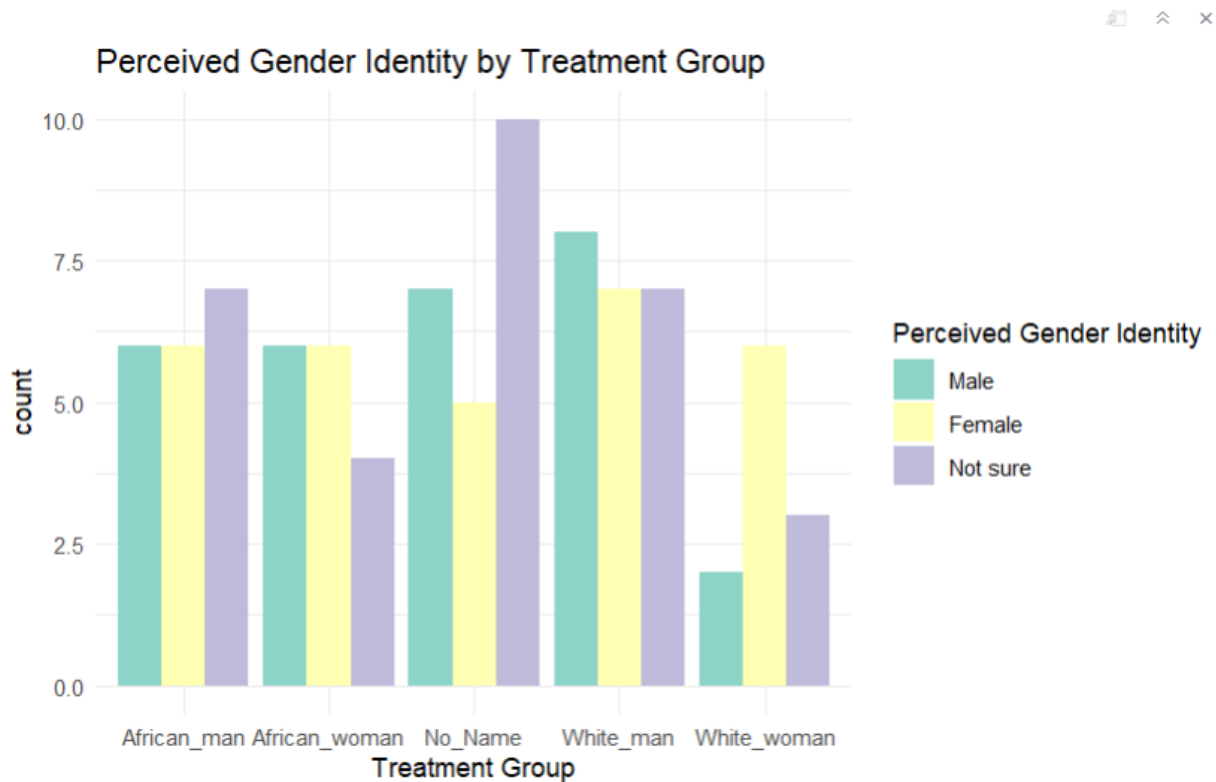


As above graph and tables shown, no name on the resume is more likely to receive “Not at all likely hiring” option for this job position, but all resumes received similar average rating for this job position. However, African women resume receive the higher average rating in hiring

possibilities in this job position, white women resume has slightly lower rating score on hiring possibilities.

Perceived Gender Identity

The perceived gender identity of candidates across treatment groups is shown in below figure:



From this graph, it is interesting to see that the participants has not much accurate sense of the gender of the name written on the resume. This result would affect the results of this experiment design heavily and is leading this experiment result into a large point of inaccuracy.

Statistical Analysis

ANOVA for Qualification Ratings

The ANOVA results for qualification ratings showed no statistically significant differences between the treatment groups ($F = 0.325$, $p = 0.861 > 0.05$).

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
stapleton_group	4	0.73	0.1813	0.325	0.861
Residuals	86	48.00	0.5581		

ANOVA for Hiring Likelihood

The ANOVA results for hiring likelihood showed no statistically significant differences between the treatment groups ($F = 1.972$, $p = 0.106 > 0.05$).

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
stapleton_group	4	5.23	1.3080	1.972	0.106
Residuals	85	56.37	0.6632		

Chi-Square Test for Perceived Gender Identity

The chi-square test for perceived gender identity by treatment group was not statistically significant ($\chi^2 = 4.7039$, $p = 0.7887 > 0.05$).

Pearson's Chi-squared test

```
data: table(data$gender_identity, data$stapleton_group)
X-squared = 4.7039, df = 8, p-value = 0.7887
```

Interpretation of Results

The results did not support Hypothesis 1, which Hypothesis 1 is rejected by the result. In the data analysis section, except for doing the ANOVA test for qualification rating and hiring possibilities, I also used Tukey HSD test to satisfy the result getting from ANOVA test. Both the

ANOVA and Tukey HSD tests failed to reveal distinctions between names that sounded Black versus White or between names that sounded female versus male.

The results for Hypothesis 2 were partially supported. While the mean ratings for the anonymized resumes (No Name group) were more uniform and did not show significant deviations compared to other groups, the ANOVA results did not show statistically significant differences.

Conclusion

This experiment is designed to test the race or gender bias on the resumes during job hiring position. As the gender identity question result evaluated, shows that participants could not correctly recognize the gender and race based on the name written on the resume. This finding leads to the result of this experiment is not valued for trust, and the result would not be accurate at all.

However, beside the above consideration, the analysis findings did not show any impact of the cues on qualification ratings or hiring likelihood in the resumes. This implies that in this study group factors, like race and gender did not play a role in how participants assessed candidate qualifications or their chances of being hired. These results were surprising, because considering the existing research on biases influencing hiring decisions.

This experiment result is contrary to the expectations. The outcomes did not demonstrate significant difference in the ANOVA test analysis of qualification ratings or hiring probabilities based on manipulated signals. This discovery indicates that in the experimental setting, the assessments made by participants were not noticeably swayed by the gender signals presented in the resumes.

Improvement to Future Studies

Upcoming research should consider sample sizes, diverse industry scopes and more nuanced hints to gain a deeper understanding of how biases play out in recruitment processes.

Furthermore, delving into the effects of intersectionality (such as combining race and gender with identity aspects) could offer insights.

In future trials, it would be advantageous to enlarge the sample size while ensuring diversity among participants could strengthen the generalizability of results, also would provide more balanced distribution in the participants' demographic result. This improvement would help the result of the experiment be more trust and confidence. The main improvement should be finding a more effective means to make sure that participants understand the gender and race of the candidates' name in the resumes. For this purpose, explicit label for race and gender might help this progress. Additionally, investigating roles beyond entry level software engineering positions might shed light on whether biases differ across industries or job categories.

References:

Gassam Asare, J. (2024, April 17). New Research Reveals Resumes With Black Names

Experience Bias In The Hiring Process. Forbes. Retrieved from

<https://www.forbes.com/sites/janicegassam/2024/04/17/new-research-reveals-resumes-with-black-names-experience-bias-in-the-hiring-process/?sh=345a2d0694bd>