

Analyzing Implicit Gender Bias of Gender Dominated Fields using Eye Tracking

BEAU BRIERRE, Clemson University, USA

ANDY YANG, Clemson University, USA

1 ABSTRACT

When looking to identify the most qualified individual for a task, many factors come into consideration. A particularly notable factor when assessing a new individual seems to be whether the person being assessed is male or female[1]. In this experiment, participants will be shown pictures of a male and female in the same role and determine which individual is more qualified. The eye movements of the participants will be tracked as they verbally identify which of the two is more qualified, looking to find a correlation between gender and perceived qualification levels. This paper will discuss the studies qualification identification research and interpret the data of the results we collected. As a disclaimer, this paper will refer to "gender" and "sex" interchangeably. It is understood that there is a scientific difference, but for the sake of simplicity, only the sexes of male and female will be considered. Studies on intersex and gender-fluid individuals will have to be done in a future experiment.

ACM Reference Format:

Beau Briere and Andy Yang. 2021. Analyzing Implicit Gender Bias of Gender Dominated Fields using Eye Tracking. In *CPSC 4120: Eye Tracking and Methodologies*, August 18–December 10, 2021, Clemson, SC. ACM, New York, NY, USA, 6 pages. <https://doi.org/10.1145/nnnnnnnn.nnnnnnnn>

2 KEYWORDS

Gender Bias, Duration, Fixation, Qualified

3 INTRODUCTION

In modern culture, certain fields seem to be dominated by a select sex[3, 4, 5, 6]. After years of this culture, the resulting effect could be an unconscious bias that certain genders are more qualified in their respective roles. Some examples of this bias would include the notion that males do well as leaders, chess players, and pilots, among other competitive fields[7]. On the other hand, this would imply that females are more suitable for gardening, nursing, parenting, cooking, teaching, and other more nurturing roles[6].

Research on implicit gender bias exists yet is still a subject that needs development. One example of a study done revealed that men are subconsciously more likely to be seen as brilliant than women[8]. This study was able to assess the subconscious by having participants perform a sorting task under time pressure, resulting in "brilliant" and "male" being paired together in the researchers response key. Eye tracking research is also limited, yet a study in the past has shown that participants tend to look towards "emergent leaders" with a longer duration and more fixations[1]. By utilizing eye tracking, it is possible to view the fixations and duration's

of participants as they decide whether a male or female is more qualified for a role.

This experiment hopes to analyze whether implicit gender bias can be seen while identifying the more qualified individual in a role. To get a solid data foundation, it is imperative that any perceived bias by the participants should be from gender and nothing else. Because of this, the stimuli shown are selected for being as similar as possible with the only main difference being the gender of the two images shown. Without knowing any further information about the people shown in the stimuli, and by controlling other factors such as body language and dress of the stimuli, any correlation found in the data makes the suggestion of an implicit gender bias stronger.

4 HYPOTHESES

4.1 Duration

The figure which is verbally identified as more qualified will receive a longer duration than the other figure. (If there is a male and female engineer the one verbally selected by the participant will receive higher duration).

4.2 Fixation

The figure which is verbally identified as more qualified will receive more fixations than the other figure. (If there is a male and female engineer the one verbally selected by the participant will receive more fixations).

5 METHOD

5.1 Participants

5 male and 5 female Clemson students were recruited by word of mouth. The average age of these participants was 20.2 years old. Most of the participants were college juniors and seniors but one was a college freshman. Some participants wore glasses that corrected their abnormal vision to normal vision, the others already had normal vision. All 10 participants sat in front of the computer monitor, and their eyes were tracked with the Gazepoint eye tracker that samples the eyes at 60Hz. Once the participants were seated, their eyes were calibrated to the screen.

5.2 Materials

All subjects were seated in front of a Dell 1680x1050 LCD monitor driven by an Alienware Alpha mini desktop PC with Intel i7-6700T CPU and Nvidia GTX960 GPU. Subjects' eyes were being tracked via a Gazepoint eye tracker connected to the mini desktop PC via USB that samples the eyes at 60Hz.

CPSC 4120, August 18–December 10, 2021, Clemson, SC

© 2021 Association for Computing Machinery.

This is the author's version of the work. It is posted here for your personal use. Not for redistribution. The definitive Version of Record was published in *CPSC 4120: Eye Tracking and Methodologies*, August 18–December 10, 2021, Clemson, SC, <https://doi.org/10.1145/nnnnnnnn.nnnnnnnn>.

5.3 Experimental Design

This experiment was a factorial design experiment where the participant chose between the more qualified of two sexes in the same role. The data collected during the experiment was the length of time spent viewing either image, the number of times the participants eyes left and returned to either image, and the verbal answer given by the participant.

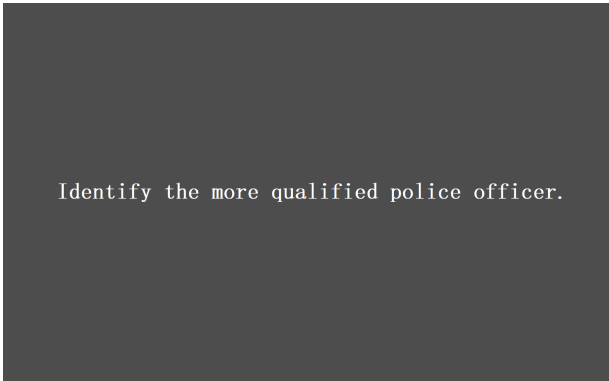
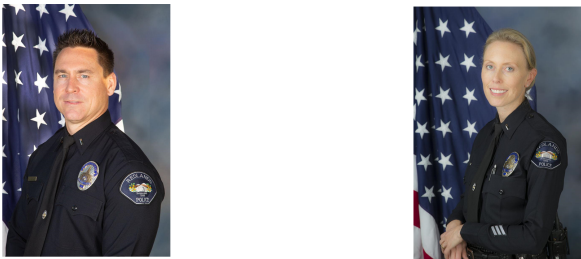


Fig. 1. Instructions preceding the image of police officers. A slide similar to this preceded every image, the only variation in each was the word for what is about to be shown.



Identify the more qualified police officer.

Fig. 2. An image of police officers



Fig. 3. An image of nurses



Fig. 4. An image of chess players



Fig. 5. An image of chefs



Fig. 6. An image of professors



Fig. 7. An image of kpop stars

5.4 Procedure

Before starting the experiment, participants were given the chance to ask any questions or raise concerns. After the participants questions and concerns were answered by the experimenter, they were given an information sheet that gave greater detail of the coming experiment and made their consent to participating official. Once consent was confirmed, the participants filled out a survey to account for demographics, this included data such as age and gender. After the survey was completed and collected, the participant was briefed on the task they were about to perform and gave verbal confirmation of their understanding the instructions. Then the eye tracker was calibrated and the experiment began. As the participant verbally determined the more qualified individual the experimenter recorded the responses. When the experiment finished, the participant was asked if they have any questions, the experimenter would answer these. The participant was then thanked and the data was analyzed.

6 RESULTS

6.1 Data Analysis

After all data was collected, 6 tables were created. Two of these were on gaze duration's of male and female participants, two others were

on fixations of male and female participants, and the last two were on the verbal answers of male and female participants. This data was entered into Microsoft Excel and tables were created based on those values.

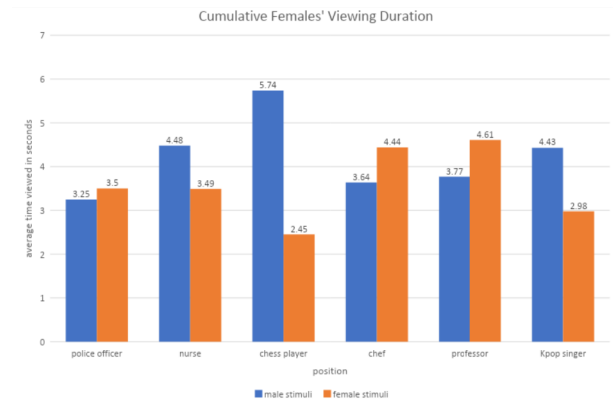


Fig. 8. Female participant average duration

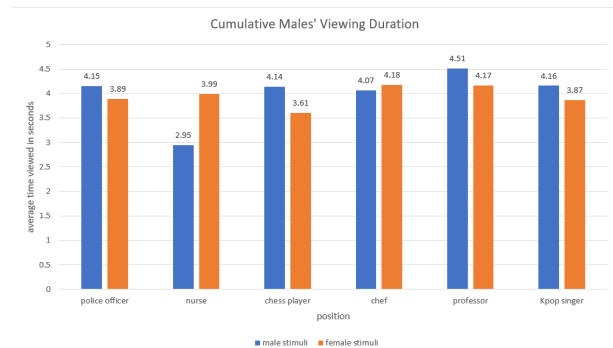


Fig. 9. Male participant average duration

Regarding gaze duration: Police officers and professors average gaze duration is higher on stimuli that have the same gender. For nurses, the average gaze duration is higher on stimuli that have the opposite gender. For chess players and Kpop singers, the average gaze duration for male stimulus is higher for both male and female participants. For chefs, the average gaze duration for female stimuli is higher for both male and female participants.

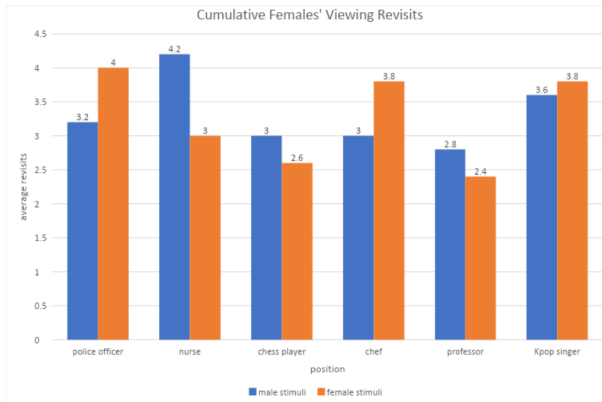


Fig. 10. Female participant average re-visitations

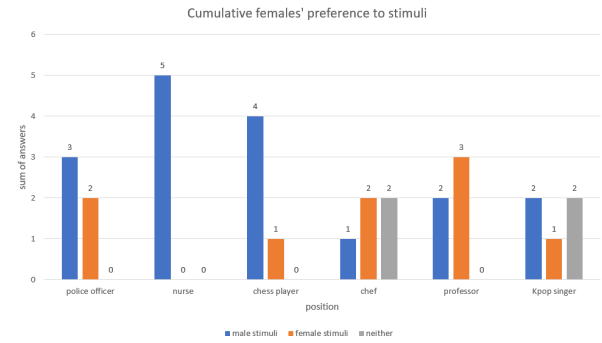


Fig. 12. Female participant total verbal answers

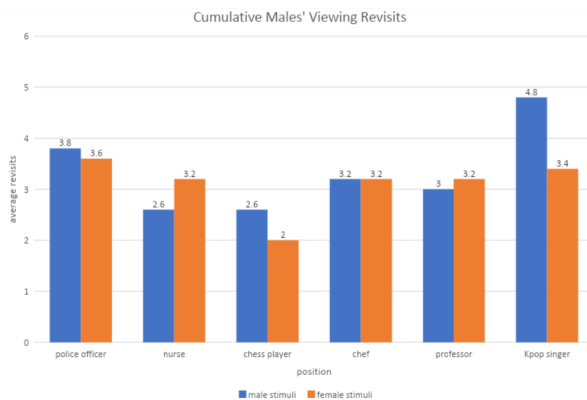


Fig. 11. Male participant average re-visitations

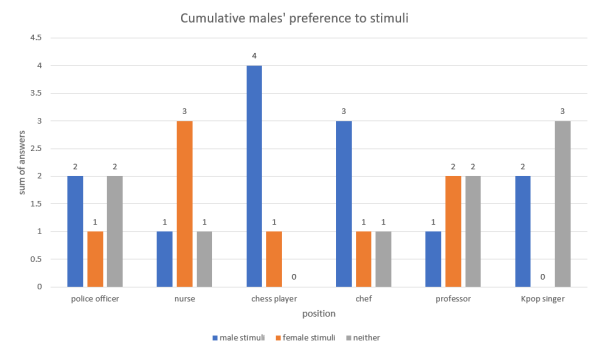


Fig. 13. Male participant total verbal answers

Regarding fixations: Police officers and Kpop singers average number of revisits is higher on stimuli that have the same gender. For nurses and professors, the average number of revisits is higher on stimuli of the opposite gender. For chess players, the average number of revisits for the male figure is higher for both male and female participants. For chefs, the average number of revisits for the female figure is higher for female participants and even for male participants.

Regarding verbal answers: For chefs, the sum of answers is higher on stimuli that have the same gender. For nurses, the sum of answers is higher on stimuli that have the opposite gender. In fact, every female participant preferred the image of the male nurse. For police officers, chess players, and Kpop singers, the sum of answers for male stimuli is higher for both male and female participants. For professors, the sum of answers for female stimulus is higher for both male and female participants. Nevertheless, both chef and Kpop singer stimuli received 2 neutral answers among female participants, and police officer, nurse, chef, professor, and Kpop singer stimuli received 2, 1, 1, 2, and 3 neutral answers among male participants. Overall, Kpop stimulus received the most neutral answers among all participants. This is probably due to the fact that most participants did not know much about Kpop prior to the experiment.

6.2 Hawthorne Effect

The Hawthorne Effect is the explanation for changed behavior resulting from the participant knowing that they are being observed [2]. The Hawthorne Effect may have influenced the participants to verbally answer in response to what they believe the experimenter wants to hear instead of what they actually think. For example, one of the male participants answered "neither" for every section except for chess player, of which he gave the outlier answer of "Woman". The opposite of this could be the general female answer for nurse, every female participant said the male nurse is more qualified than

the female nurse. This could imply either the female participants very much prefer the male nurse, or that they answered this way because they thought it was the "right" answer for the experiment. Although there were some significant results in the data, care should be taken while analyzing since this may be a factor and the sample size is only 10 students.

6.3 Discussion

The data of this study suggests that certain roles carry implicit bias in the minds of University Students. This bias seems to be dependant on whether or not the participant is male or female, the most notable example of this would be the nurse and chess player roles. When viewing the nurse role it is worth distinguishing eye tracking and verbal data. Through eye tracking it can be noted that males looked at the female nurse longer and had more fixations than the male nurse. Female participants reversed this trend by looking at the male nurse longer and with more fixations. The verbal results correlated directly with the eye tracking data, male participants largely preferred the female nurse and female participants largely preferred the male nurse.

The chess player role is dominated by the image of the male figure in every aspect of data. Both male and female participants heavily preferred the male chess player by verbal confirmation.

The data on police officer seems to slightly favor supporting a bias towards the male officer. However the eye tracking data and verbal confirmations do not carry much significance one way or the other. Similarly, the data on the Kpop singer seems to favor the male figure slightly more. However, this is not very strongly weighted to one side or the other. This is most likely due to the participants not knowing much about Kpop before coming into the experiment.

Not every role had this significant of a difference in eye-tracking data versus verbal answer. Roles such as professor and chef do not seem to carry a heavy statistic significance. One explanation for this could be that these roles could be seen as either male or female dominated depending on the level performed. Elementary education is saturated with female teachers[3] while STEM professors are generally male[4]. A head chef in a restaurant is typically a male dominated role[5] while the head of the kitchen in the house has traditionally been the female member of the family[6]. Since either gender can be seen as successful completing these roles, the participants visual data seems to suggest that the participants do not favor one significantly more than the other. The verbal data for professor seems to support this as well, the difference between male and female professor preference boils down to one person each. However, the male participants seem to have a verbal preference towards the male chef despite visual data that seems even towards both images, female participants do not seem to share this same verbal opinion.

7 CONTRIBUTIONS TO KNOWLEDGE

This study has contributed to the knowledge base of implicit gender bias by strengthening the position that gender bias exists in males and females. The correlations between the eye tracking data and the verbal answers expands the concept that an individual is more likely to look towards the individual they believe is more qualified. For

the first proposed hypothesis, the roles that have statistical significance support the concept that participants have a longer duration towards the individual they deem is more qualified. For the second proposed hypothesis, the roles that have statistical significance also support the concept that participants have more fixation towards the individual they deem is more qualified. However, not all the roles in this experiment support the proposed hypotheses. For example, the data of the police officer role from the female participants falsified both of the proposed hypotheses. The verbal answers for police officers is opposite from the gaze duration and number of fixations of the participants. This most likely happened because the participant sample size is too small or the participants were recruited by word of mouth.

8 RECOMMENDATIONS FOR FUTURE STUDIES

In this study, some of the roles were partial, chef and professor as mentioned in discussion. To avoid getting data that makes the result statistically insignificant, be more extreme with the roles shown.

When recruiting participants, try to select random individuals without word of mouth. One potential problem of this study is that the participants were recruited by word of mouth, meaning they knew the experimenter prior to participating in the experiment. By recruiting strangers, the risk of skewed data through the Hawthorne Effect decreases and there may be more clear-cut results.

When recruiting participants, recruit more than 10 so that the data carries even more statistical weight.

Give the participants more time to generate a verbal answer, but make it mandatory to choose so that neither is not an option.

Look into gender-fluid and intersex biases with an eye tracking component.

9 REFERENCES

- [1]Gerpott, F. H., Lehmann-Willenbrock, N., Silvis, J. D., & Vugt, M. V. (2017, November 20). In the eye of the beholder? an eye-tracking experiment on emergent leadership in team interactions. *The Leadership Quarterly*. Retrieved October 14, 2021, from <https://www.sciencedirect.com/science/article/pii/S1048984317305003?via%3Dihub>
- [2]Wickström, G., Bendix, T. (2000). The "Hawthorne effect"—what did the original Hawthorne studies actually show?. *Scandinavian journal of work, environment health*, 26(4), 363–367.
- [3]Coe - characteristics of Public School Teachers. (n.d.). Retrieved December 1, 2021, from <https://nces.ed.gov/programs/coe/indicator/clr>.
- [4]Jensen, S. (2019). *Women in STEM: Strategies and Recommendations for Academic Women and Institutional Leaders*. Theses and Dissertations Retrieved from <https://scholarworks.uark.edu/etd/3473>
- [5]Chef demographics and statistics in the US - zippia.com. (n.d.). Retrieved December 1, 2021, from <https://www.zippia.com/chef-jobs/demographics/>.
- [6]Brenan, M. (2021, November 20). Women still handle main household tasks in U.S. Gallup.com. Retrieved December 1, 2021, from <https://news.gallup.com/poll/283979/women-handle-main-household-tasks.aspx>.

[7]Men are still more likely than women to be perceived as leaders, study finds. University at Buffalo. (2018, August 6). Retrieved December 2, 2021, from <http://www.buffalo.edu/news/news-releases.host.html/content/shared/mgt/news/men-still-more-likely-than-women-perceived-leaders-study-finds.detail.html>.

[8]ScienceDaily. (2020, July 2). Implicit bias against women: Men more likely than women to be seen as brilliant. ScienceDaily. Retrieved December 2, 2021, from <https://www.sciencedaily.com/releases/2020/07/202007200000.htm>.