"Yes, Chef!": Gender Biases in Hiring in Commercial Kitchens

Jonah E. Kimmel

PSYC 380: Honours Research Project Seminar

Department of Psychology, McGill University

Supervised by Dr. Jordan R. Axt

Submitted to Dr. Anna Weinberg

April 29, 2023

Abstract

One cause of workplace discrimination stems from perceptual biases around an individual's suitability for a role. In these contexts, evaluators may use socially salient but nondiagnostic information about potential candidates, such as race, gender, and age. This work focuses on one possible instance of such discrimination: gender-based biases in professional kitchens. While correlational research shows the existence of gender discrimination in selecting women for positions of kitchen leadership, no experimental research currently exists examining this process within a laboratory setting. We hypothesized that a gender bias exists in the decision-making process for hiring women to positions of power in kitchens, namely that participants would have a higher criterion for selecting women when compared to equally qualified men. Using a task where participants select candidates using objectively more- or less-qualified diagnostic information, paired with non-diagnostic gender information in the form of photographs, we anticipated both culturally prevalent biases against women in leadership roles and individual differences in participant age and gender would drive this effect. However, contrary to expectations, participants (n = 294) held lower and thus more preferential criteria for hiring women than equally qualified men. Potential causes for this result may stem from relying on convenience samples and features of the stimuli used to represent gender. We propose future directions, including representative population sampling and methodological refinement, to clarify and elaborate our findings.

Gender Biases In Hiring in Commercial Kitchens

Implicit biases in hiring against women - that is, judgments of role suitability made with comparatively automatic biases against women – has clear impacts on women's professional life (Moss-Racusin et al., 2012), but also carries implications for individuals' mental health and can even inhibit organizational success by reducing team diversity (Rospenda et al., 2009; "*Women in Business and Management*," 2018). This work focuses on one form of discrimination against women, specifically within professional kitchens. While women constitute half of the culinary school graduates in the United States, female respondents to a correspondence study (n = 337) indicated that 61% of the discrimination they faced resulted in unfair opportunities for hiring and advancement (Woods & Kavanaugh, 1994). Furthermore, audit studies show that while 41% of chefs and cooks in Canada are women, 71% of the lowest-paying kitchen roles of food counter attendants and kitchen helpers are women (Statistics Canada, 2017), and only 6% of the top chef positions in the highest-rated global kitchens are women (*About a Quarter*, 2022). While the roots of such discrimination can be structural (Green, 2003), prior evidence also suggests a more psychological cause, as a meta-analysis of the literature on stereotypical leadership qualities found that the concept of leadership suitability was predominantly masculine (Koenig et al., 2011). These results suggest that the disparity between the number of women in the culinary industry and those that lead it could at least be partly from automatic associations and automatic biases regarding gendered roles in leadership.

Still, no lab-based, experimental research exists examining such biases in hiring and promoting women in kitchen leadership roles. To investigate this question in the present work, we adapted a recent psychological measure called The Judgement Bias Task (JBT; Axt, Nguyen & Nosek, 2018). The JBT is a flexible, easily administered measure that assesses the degree to

which participants incorporate non-diagnostic social information into their judgment and decision-making. Specifically, the JBT pairs multiple pieces of relevant diagnostic information with nondiagnostic social information, requiring participants to select or reject candidates for a proposed position.

Using a Signal Detection Theory (SDT) analysis, the JBT then examines whether members of different social groups (in this case, men and women) receive more lenient treatment in evaluation (quantified by response criterion), despite both groups being equated on objective indicators of applicant quality. In this version of the JBT, applicants for a kitchen opening were shown with diagnostic information like amount of work experience, previous restaurant rating, and culinary school GPA in addition to non-diagnostic information in the form of a photo of the applicant, which communicated the applicant's gender. Applications are scored such that some are more qualified and some are less qualified, though this difference is relatively small, and the difficulty participants have in synthesizing across the four pieces of objective information allows for errors in judgment that may systematically benefit some groups over others. For example, in prior uses of the JBT (Axt & Lai, 2019), more physically attractive applicants received a lower response criterion than less physically attractive applicants, meaning that more attractive applicants were more likely to receive "beneficial errors" (being accepted for a position when one is actually less qualified) and less attractive applicants were more likely to receive "detrimental errors" (being rejected for a position when one is, in fact, more qualified).

As a result, a difference in response criterion between social groups on the JBT can illustrate a preference in the treatment of one group over the other, and applying the task to the issue of gender bias in kitchens could provide more controlled, experimental evidence of discriminatory hiring biases for kitchen leadership that would align more correlational evidence

of disparate hiring of women to positions of power in commercial kitchens from Canada (Statistics Canada, 2017), the United States (*Tipped Over the Edge*, 2012), Spain (Santero-Sanchez et al., 2015), and the United Kingdom (*Women in Hospitality,* 2020*)*.

We hypothesized that participants completing a JBT investigating gender-biased biases in kitchen contexts would be relatively more lenient in their criterion for selecting male candidates and relatively stricter in their criterion for female candidates. Participants may show biased judgment in selecting candidates for several reasons. For one, if participants have pre-existing biases that favour stereotypically masculine traits in leadership (Koenig et al., 2011), they may exhibit a higher criterion and be more stringent for selecting women for a position of leadership. That is, if participants hold beliefs around the required traits needed for professional kitchen work in general, and those traits are stereotypically masculine, for example, perceived confidence or independence (Ehrlinger & Dunning, 2003), that may participants to prefer men over women applicants, even if the two groups are equated on our objective indicators of quality.

In the present study, we investigated this issue by using the JBT and various self-report items concerning explicit stereotypes between gender and leadership generally and the kitchen business specifically. Understanding the mechanisms of discrimination in kitchen hiring that inhibit diversity and the degree to which they are automatic or explicit will permit focused research on specific interventions to address bias.

Materials and Methods

We sought to recruit 300 eligible participants, which provided 93% power to detect a small within-subjects effect of Cohen's d = 0.20 when comparing the criterion values for male versus female applicants, as determined by G*Power 3.1.9.6. In total, 331 participants' data were collected from the Project Implicit (implicit.harvard.edu) participant pool (M_{age} = 39.06, SD =

15.21; 63.6% White, 67.4% women, 68.3% United States citizens). Participants were oversampled as we anticipated excluding some participants due to task performance exclusion criteria. Specifically, participants were excluded from the analysis if they accepted < 20% or > 80% of the applications, indicating a failure to follow instructions to reject approximately half (21 participants were excluded for this criteria; Axt et al., 2018).

Additionally, participants were excluded if they reported a gender other than male or female, as exploratory analyses were concerned with identifying if any differences existed between criteria for male or female applications by male or female participants. Eight participants reporting "other" and eight reporting "null"/non-responding to gender were omitted from the analysis, leaving a final sample of 294 participants. Demographic information was collected at the time of registration with the Project Implicit participant pool.

The study session consisted of two components, first, the Judgment Bias Task and then the self-report measures of gender-based attitudes and stereotypes. See https://osf.io/wc3s2/ for the study's pre-registration of materials, procedure, and analysis strategy.

Judgment Bias Task. Participants were instructed to evaluate candidates for the next round of hiring decisions for the sous-chef position (i.e., assistant head chef) for a restaurant kitchen job. They were asked to accept approximately half of the candidates as most qualified for the job and reject approximately half. Participants then completed one of six possible JBT orders with 64 pre-generated trial stimuli shown in random order. The 64 applications had 32 female and 32 male faces selected from the Chicago Face Database (Ma et al., 2015). The facial stimuli selected were rated high on their perceived Caucasian race, perceived gender identity, low-to-neutral attractiveness, and low dominance. Perceived race and gender were calculated as the proportion of raters who positively categorized a photo divided by the number of raters, and

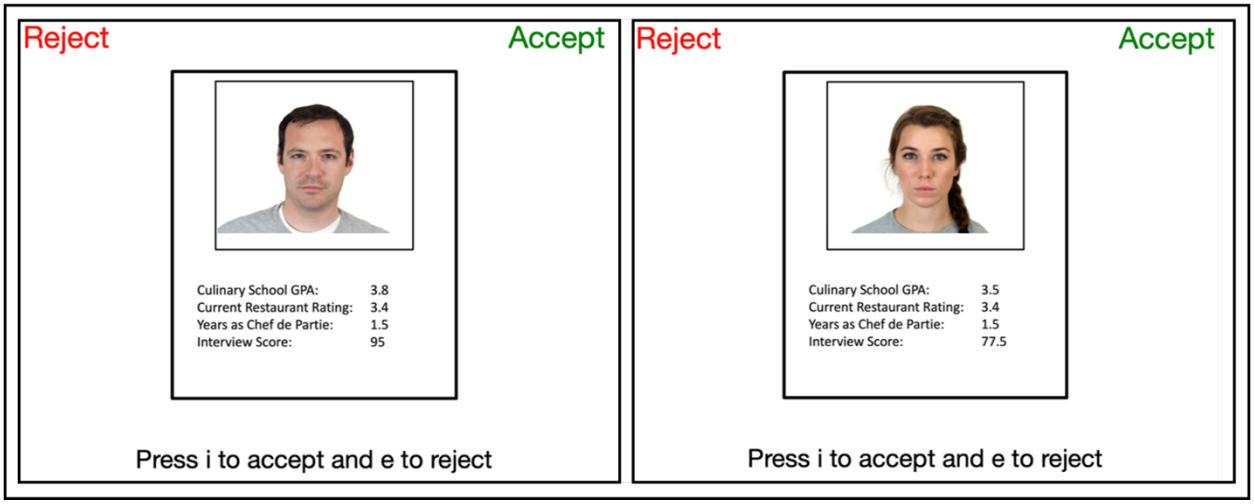
attractiveness and aggressiveness were reported on a 7-point Likert scale (1 = Not at all; 4 = Neutral; 7 = Extremely). Male stimuli were rated unanimously as male-gendered (M = 1.00), high on perceived Caucasian race (M = 0.98, SD = 0.05), low on physical attractiveness (M = 2.88, SD = 0.34), and low in dominance (M = 2.88, SD = 0.37). Female stimuli were rated unanimously as female-gendered (M = 1.00), high on perceived Caucasian race (M = 0.98 , SD = 0.02), neutral on attractiveness (M = 3.77, SD = 0.83), and low in dominance (M = 2.39, SD = 0.44).

Each JBT trial had a photo of either a male or a female candidate and four pieces of qualification information: Culinary School GPA (Scale of 1-4), Current Restaurant Rating (1-4), Years of Experience as a Chef de Partie (0.5-2 years in 0.5-year intervals), and Interview Score (1-100). See Figure 1. Across trials, each face was equally likely to be assigned to a more or less qualified application. Half (16) of the male and female candidates were objectively more qualified, and half were objectively less qualified, with higher or lower total sum qualification scores for the four pieces of hiring information. Applications were created so that no two stimuli had identical hiring information.

To calculate the more- and less-qualified applications, scores were converted to a scale with a maximum possible score of four, such that the maximum possible score for any application was 16. Culinary School GPA and Current Restaurant Rating were already scored out of 4 and thus required no conversion. Years of Chef de Partie experience were multiplied by two, and Interview Scores were divided by 25 to arrive at the 4-point scale. Objectively more qualified application scores summed to 14 out of 16, while less qualified applications summed to 13 out of 16, which follows prior uses of the JBT (Axt et al., 2018).

Figure 1

Examples of Stimuli Used



Note. Stimuli were presented one at a time in random order from a pre-generated order of 64. The stimulus on the left demonstrates a more-qualified male applicant with qualification scores totalling 14 out of 16, and the stimulus on the right shows a less-qualified female applicant with qualification scores totalling 13 out of 16.

At the start of the viewing phase, participants had each of the four pieces of qualification defined to them, so there was clarity about the scoring and significance of each. Participants then viewed all 64 stimuli individually for one second each to familiarize them with the diagnostic information. Next, participants began the selection phase, working through the 64 applications one at a time. They were instructed to either reject the application by pressing the ‘e’ key on their keyboard or accept it by pressing the ‘i’ key. Participants were given 20 seconds to respond per application, after which the stimuli timed out, and the sorting task proceeded to the next application. Participants were not given any trial feedback but were shown a prompt to “Respond Faster” following any timeouts.

Self-report questionnaire. Following the JBT, participants completed a questionnaire with items measuring culinary-related gender stereotypes, performance on the JBT, and any previous professional kitchen experience. Specifically, the questionnaire included two questions about JBT performance, one question on their beliefs on kitchen leadership, two questions of beliefs about kitchen skills, and one about kitchen experience. All questions were scored on a 7-point scale, except for the question on kitchen experience, which had a yes/no response scale. All participants completed the questionnaire items in the order described above.

Task performance included one question on *perceived* task performance: "Which statement best describes your **performance** on the task?" (1 = I was extremely easier on female applicants and extremely tougher on male applicants, 7 = I was extremely easier on male applicants and extremely tougher on female applicants), and one question on *desired* task performance: "Which statement best describes how you **wanted** to perform on the task?" (1 = I wanted to be extremely easier on female applicants and extremely tougher on male applicants, 7 = I wanted to be extremely easier on male applicants and extremely tougher on female applicants).

Beliefs on kitchen leadership were measured with one question: "Which statement best describes your beliefs about kitchen leadership?" (1 = Women are significantly more capable of leading a kitchen than men, 7 = Men are significantly more capable of leading a kitchen than women). Kitchen skills included one question on inherent skills: "Kitchen leadership requires a specific set of inherent skills." (1 = Strongly agree, 7 = Strongly disagree), and one question on learned skills: "Kitchen leadership skills can be learned by anyone." (1 = Strongly agree, 7 = Strongly disagree).

Finally, the participant's kitchen experience was assessed with one question: “Have you ever worked as a professional chef or cook?” (1 = Yes, 2 = No).

Results

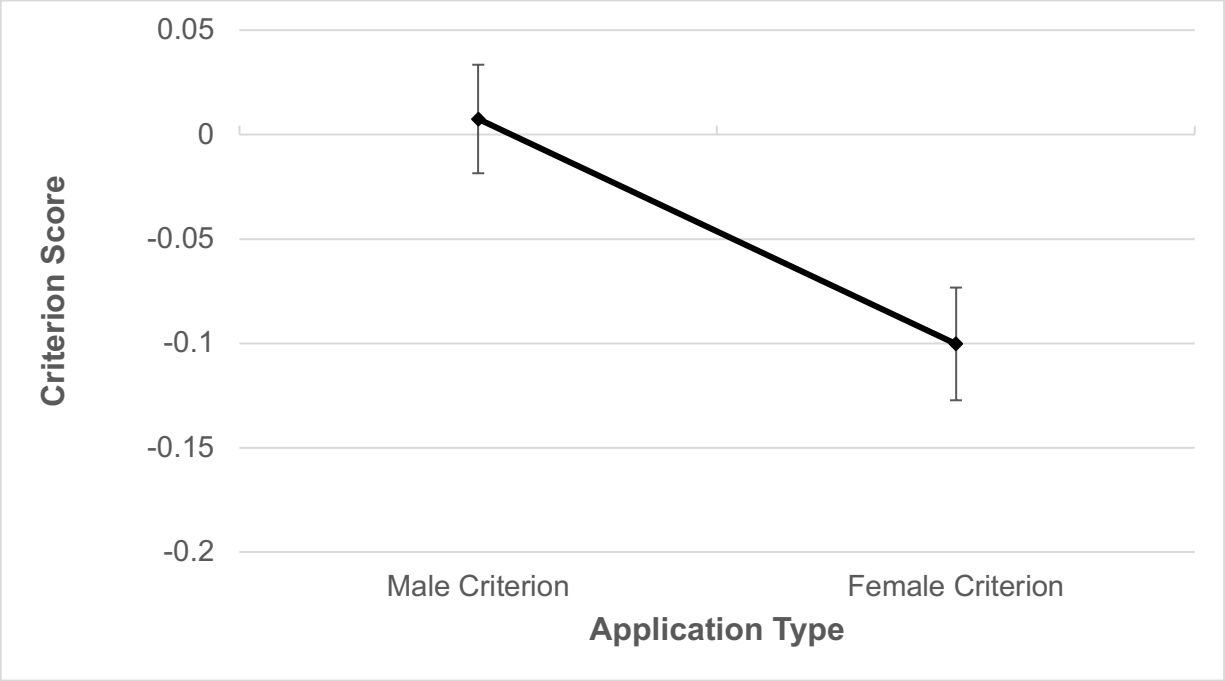
First, we examined differences in the criterion for selecting male versus female applicants among all eligible participants. We then performed exploratory analyses on how differences in criterion related to participant age, gender, perceived task performance, and explicit beliefs on kitchen leadership and kitchen skills.

Total accuracy on the task was 63.4% (*SD* = 6.7), as defined by participants correctly selecting more qualified applications and correctly rejecting less qualified applications, indicating that accuracy was higher than a chance rate, but the task was also not trivially easy. The average acceptance rate of all participants was 51.4% (*SD* = 13.6), which was close to the recommended acceptance rate per the task’s instructions.

Differences in response criterion for male versus female applications were of principal interest, and criterion scores were computed using the same procedure used by Correll et al., 2007. Using a paired-sample *t*-test and contrary to our expectations, female applications (*M* = -0.100, *SD* = 0.47) received a lower response criterion than male applications (*M* = 0.01, *SD* = 0.44), $t(293) = 5.08$, $p < .001$, $d = 0.30$, 95% C.I. [0.07, 0.15]. See Figure 2. This criterion bias indicates favouritism in selecting female versus male applications for a position of kitchen leadership, even though both groups of applicants had equal qualifications. That is, female applicants did not have to be as qualified as male applicants in order to receive a “hire” response from participants. Since less than 30 participants with professional kitchen experience were recruited for participation, the preregistered analysis of investigating gender-based criterion bias among participants with kitchen experience was not investigated.

Figure 2

Differences between criteria for preferential selection of applications paired with male versus female stimuli.



Note. Error bars represent the standard error of the mean for male and female criteria.

Exploratory analyses were conducted by first calculating a single criterion difference score for each participant by subtracting female hiring criterion scores from male hiring criterion scores, with positive difference scores then meaning a greater bias for hiring women over men and negative difference scores representing a bias for hiring men over women ($M = 0.11$, $SD = 0.36$). Difference scores were then correlated with rater age, kitchen experience, explicit measures of desired task performance, perceived task performance, kitchen leadership beliefs, and beliefs on the importance of inherent and learned kitchen skills.

Results found that the criterion bias measure was only correlated with one outcome. Specifically, the criterion difference score was weakly and positively correlated with item on kitchen leadership beliefs "Kitchen leadership skills can be learned by anyone." (1 = Strongly agree, 7 = Strongly disagree), $r(288) = .12$, $p = .037$, 95% C.I. [0.01, 0.23]. This positive correlation indicates that participants with a stronger belief that "not everyone can learn kitchen leadership skills" have a more relaxed criterion for women versus men when hiring for leadership positions in restaurant kitchens. No other correlations with participant demographics or explicit items were observed. See Table 1 for descriptive statistics and bivariate correlations.

Table 1

Pearson Correlations for Demographics, Explicit Beliefs, Kitchen Skills, and Task Performance.

Variable	*n*	*M*	*SD*	1	2	3	4	5	6	7	8	9
1. Kitchen experience [a]	294	1.94	0.24	—								
2. Gender belief on kitchen leadership	292	3.96	0.50	.04	—							
3. Inherent kitchen skills	293	2.38	1.30	.03	-.03	—						
4. Learned kitchen skills	290	2.80	1.59	-.06	.05	-.14*	—					
5. Perceived task performance	290	3.90	0.49	.04	.11	.09	.01	—				
6. Desired task performance	285	3.92	0.34	.11	.08	.08	.04	.33***	—			
7. Participant age	294	39.77	15.15	-.10	-.04	-.07	.14*	.13*	.11	—		
8. Participant gender [b]	294	0.29	0.45	.04	.01	-.01	-.11	.05	.05	-.03	—	
9. Criterion difference score [c]	294	0.11	0.36	.01	-.04	.03	.12*	-.05	.08	.07	-.01	—

[a] 1 = has professional kitchen experience, 2 = does not have professional kitchen experience

[b] 0 = female, 1 = male

[c] Calculated within subjects, $C_{difference}$ = (Male criteria score – Female criteria score)

$^{*}p < .05$. $^{***}p < .001$.

General Discussion

We sought to examine if differences in criterion existed between hiring men or women for kitchen management in a hypothetical hiring task. Based on past research concerning associations between gender and leadership as well as real-world evidence of gender disparities in high-status kitchen positions, we expected to find that men were more likely to be hired than women, regardless of their qualifications. Workers in commercial kitchens commonly come from vulnerable populations: young workers, immigrants, and individuals who have low socioeconomic status, so examining what biases affect their positive career trajectory, which in turn affects earning power and overall mental well-being, is crucial.

Our study results of more leniency towards female applicants are the opposite of what we expected to find based on reports of hiring/promotion discrimination against women (Woods & Kavanaugh, 1994) and the broad empirical evidence illustrating the disproportional representation of women in professional kitchens globally (Statistics Canada, 2017; *Tipped Over the Edge*, 2012; Santero-Sanchez et al., 2015; *Women in Hospitality,* 2020). However, it is worth noting that our sample is from the general population. These results may be partially due to the lack of participant kitchen experience, as only 6.1% of participants reported having such experience. In addition, the research question may have been deduced by participants and resulted in socially desirable response bias, with participants preferentially selecting traditionally disadvantaged groups in a selection task. Indeed, similar effects have emerged with the JBT when using other socially sensitive domains (e.g., race, Axt, Ebersole & Nosek, 2016).

To better understand whether participants' greater leniency for selecting female applications resulted from conscious “affirmative action” of intentionally helping applicants from stigmatized groups, as compared to a more automatic, implicit desire to correct for a

historical disadvantage, we examined the differences in criterion for participants who self-reported unbiased kitchen leadership beliefs and fairness in their JBT performance.

Specifically, after completing the JBT, participants responded to one item on explicit beliefs of kitchen leadership and gender: "Which statement best describes your beliefs about kitchen leadership?" (1 = Women are significantly more capable of leading a kitchen than men, 4 = Men and women are equally capable of leading a kitchen, 7 = Men are significantly more capable or leading a kitchen than women) with participants not endorsing an explicit gender bias with regards to kitchen leadership (M = 3.96, SD = 0.50), $t(291) = -1.42$, $p = .158$, $d = -0.08$, 95% CI [-0.19, 0.03]. For participants who reported no gender bias in the capability between genders for leading a kitchen with a response of 4 (91.1% of respondents), female applicants (M = -0.11, SD = 0.47) received a lower criterion than male applicants (M = 0.01, SD = 0.43), $t(265) = 5.18$, $p < .001$, $d = .32$, 95% CI [0.07, 0.15], an effect size nearly equivalent to that observed with the whole sample. This indicates that even for participants who report no explicit gender bias for kitchen leadership, a preferential criterion difference for women exists.

With regards to the intention to judge both female and male applicants equally, participants responded to one item on perceived task performance: "Which statement best describes your **performance** on the task?" (1 = I was extremely easier on female applicants and extremely tougher on male applicants; 4 = I treated both female and male applicants equally; 7 = I was extremely easier on male applicants and extremely tougher on female applicants). For participants who reported that they selected candidates equally with a response of 4 (87.2% of respondents), female applicants (M = -0.11, SD = 0.48) received a lower criterion than male applicants (M = 0.00, SD = 0.45), $t(252) = 4.75$, $p < .001$, $d = 0.30$, 95% CI [0.06, 0.15], an effect size nearly equivalent to that observed with the whole sample. Finally, examining how

participants desired to rate both male and female applicant groups equally, participants responded to one item on desired task performance: "Which statement best describes how you **wanted** to perform on the task?" (1 = I wanted to be extremely easier on female applicants and extremely tougher on male applicants; 4 = I wanted to treat both female and male applicants equally; 7 = I wanted to be extremely easier on male applicants and extremely tougher on female applicants). For participants who reported wanting to treat both groups of applicants equally with a response of 4 (93.3%), female applicants (M = -0.10, SD = 0.47) received a lower criterion than male applicants (M = 0.02, SD = 0.45), $t(265) = 5.1078$, $p < .001$, $d = .31$, 95% CI [0.07, 0.16], an effect size nearly equivalent to that observed with the whole sample.

In all, these results indicate either that this socially desirable compensation results either from 1) more automatic processes unrelated to explicit desires to promote one gender over the other in kitchen leadership or 2) that participants compensated through affirmative action contrary to the observed underrepresentation of women in kitchen leadership, and then mislead researchers as to their motivations for doing so. It is also possible that some participants had a different understanding of the term "equal treatment" than was intended, as some may have considered favouring female applicants to be a form of "equal treatment" given the existence of interpersonal prejudices and structural disparities that promote men over women in the culinary field. Future directions will consider modifying these items to ask about more fine-grained perceptions of fairness and include items that investigate if participants are aware of disparities in representation for women in positions of power in kitchens.

The unexpected results observed in the current study could have stemmed from the study population, the study methods, or a combination of both. To address these points, future research will need to replicate the proposed study using a population recruited from current kitchen

workers, and ideally composed of participants with prior experience or responsibility for hiring to positions of power. In addition, it may be productive to modify the JBT to include other demographic differences among applicants, so that the notion of gender may not be as consciously salient to participants and create more naturalistic responding.

Relatedly, future studies may want to adjust the approach used to select stimuli for the task. As discussed earlier, the concept of leadership is an inherently male construct (Koenig et al, 2011), and that trait dominance contributes to the emergence of leadership (Kim et al., 2020). In selecting stimuli faces for our trials from the Chicago Face Database (Ma et al., 2015), we focused first on gender and race prototypicality and then sought to equate male and female stimuli as much as possible on ratings of dominance/aggressiveness and attractiveness. Notably, both the male and female stimuli were characterized by low dominance scores out of a possible score of seven, with male scores ($M = 2.88$, $SD = 0.37$) approximately one-half point higher than female scores ($M = 2.39$, $SD = 0.44$). Recent research suggests that perceived dominance correlates positively with perceptions of leadership suitability, but only for men (Kim et al., 2020). As a result, if participants were using this stereotypical association between dominance and leadership as part of their automatic judgment of male applicants, male applications with low dominance stimuli may have been judged with a higher criterion for failing to fit this stereotype. In this sense, trying to minimize differences between male and female stimuli on dominance may have led to an unrepresentative sample of male faces that impacted participants' responses.

In addition, participants from a lay or non-kitchen background may have found the diagnostic information presented particularly unimportant or irrelevant in their own judgment heuristics due to their unfamiliarity with what may qualify an applicant and thus have reverted to more personally salient or relevant non-diagnostic information to create a standard by which to

sort participants. As such, participants faced with socially salient stimuli (like gender) without a clear direction to accept or reject participants of a particular social category may engage in a counteractive strategy to reduce uncomfortable feelings based on preconceptions of gender discrimination during applicant selection, in this case choosing to select applicants from a group that they know to be a target of discrimination. That is, our results may have been due to participants being aware that discrimination against women in management is prominent, and it is generally considered socially undesirable for people to express gender biases publicly. Future studies will benefit from using distractor items in the qualification section of the presented applicant information. For example, the applicant's cooking cuisine type or how many siblings they have, could be listed to draw attention away from applicant photographs and focus their attention more equally on diagnostic information, with the aim of reducing overt compensation towards socially desirable results by preferring women over men in hiring.

Furthermore, the JBT task could be modified to make the research question less salient to participants in the interests of reducing social desirability bias in responses by reducing the overt salience of the proportionality between the gender of participants. This could be accomplished by introducing but controlling for a second race in stimuli photos. We suggest that by adding a second differing social identity to a minority of stimuli photos, it may be possible to make the gender of stimuli photos less salient for explicit decisions based on social desirability. Furthermore, the JBT could be improved by adjusting the number of male applications to female applications to be more representative of their prevalence in real-world kitchen situations. This could be achieved while maintaining the proportion of more and less qualified applications between those two groups, for example, by shifting it from 50% men and women to 30% women

and 70% men, and thus it may be possible in the future mask the research goal of the study from participants while still being able to measure gender biases in judgment.

Gender discrimination in hiring, particularly for positions of power within kitchens, is contemporary and ubiquitous (Statistics Canada, 2017; *Tipped Over the Edge*, 2012; Santero-Sanchez et al., 2015; *Women in Hospitality,* 2020). Differences in criterion response, representing a bias in hiring between men and women, regardless of the benefiting group, evinces continued heuristic biases (Kahneman et al., 1982) that shape meaningful, individual and group outcomes for many people, particularly for members of vulnerable or stigmatized groups. Continued focus on this issue will be important to understand what factors contribute to the underrepresentation of women in positions of power in restaurant kitchens.

References

Axt, J. R., Ebersole, C. R., & Nosek, B. A. (2014). The rules of implicit evaluation by race, religion, and age. *Psychological Science*, *25*(9), 1804–1815. https://doi.org/10.1177/0956797614543801

Axt, J. R., Ebersole, C. R., & Nosek, B. A. (2016). An unintentional, robust, and replicable pro-black bias in social judgment. *Social Cognition*, *34*(1), 1–39. https://doi.org/10.1521/soco.2016.34.1.1

Axt, J. R., & Lai, C. K. (2019). Reducing discrimination: a bias versus noise perspective. *Journal of Personality and Social Psychology*, *117*(1), 26–49. https://doi.org/10.1037/pspa0000153

Axt, J. R., Nguyen, H., & Nosek, B. A. (2018). The judgment bias task: a flexible method for assessing individual differences in social judgment biases. *Journal of Experimental Social Psychology*, *76*, 337–355. https://doi.org/10.1016/j.jesp.2018.02.011

Axt, J. R., Yang, J., & Deshpande, H. (2022). Misplaced intuitions in interventions to reduce attractiveness-based discrimination. *Personality and Social Psychology Bulletin,* 1-14. https://doi.org/10.1177/01461672221074748

Chef's Pencil (2022). About a quarter of chefs are women: but only 6% reach the very top. *Chef's Pencil*. Retrieved October 21, 2022, from https://www.chefspencil.com/female-chefs-6-percent-reach-the-top/

Correll, J., Park, B., Judd, C. M., Wittenbrink, B., Sadler, M. S., & Keesee, T. (2007). Across the thin blue line: police officers and racial bias in the decision to shoot. *Journal of Personality and Social Psychology*, *92*(6), 1006–23.

Ehrlinger, J., & Dunning, D. (2003). How chronic self-views influence (and potentially mislead)

estimates of performance. *Journal of Personality and Social Psychology*. *84*. 5-17. 10.1037//0022-3514.84.1.5.

Green, T. K. (2003). Discrimination in workplace dynamics: toward a structural account of disparate treatment theory. *Harvard Civil Rights Civil Liberties Law Review*, *38*, 91–158.

Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty: heuristics and biases*. Cambridge University Press.

Kim, J.-Y., Hsu, N., Newman, D. A., Harms, P. D., & Wood, D. (2020). Leadership perceptions, gender, and dominant personality: the role of normality evaluations. *Journal of Research in Personality*, *87*. https://doi.org/10.1016/j.jrp.2020.103984

Koenig, A. M., Eagly, A. H., Mitchell, A. A., & Ristikari, T. (2011). Are leader stereotypes masculine? A meta-analysis of three research paradigms. *Psychological Bulletin*, *137*(4), 616–642.

Lai, C. K., Skinner, A. L., Cooley, E., Murrar, S., Brauer, M., Devos, T., Calanchini, J., Xiao, Y. J., Pedram, C., Marshburn, C. K., Simon, S., Blanchar, J. C., Joy-Gaba, J. A., Conway, J., Redford, L., Klein, R. A., Roussos, G., Schellhaas, F. M. H., Burns, M., … Nosek, B. A. (2016). Reducing implicit racial preferences: ii. intervention effectiveness across time. *Journal of Experimental Psychology. General*, *145*(8), 1001–16. https://doi.org/10.1037/xge0000179

Ma, D. S., Correll, J., & Wittenbrink, B. (2015). The Chicago face database: A free stimulus set of faces and norming data. *Behavior research methods, 47*(4), 1122-1135.

Moss-Racusin, C. A., Dovidio, J. F., Brescoll, V. L., Graham, M. J., & Handelsman, J. (2012). Science faculty's subtle gender biases favor male students. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(41), 16474–16479

Ofosu, E. K., Chambers, M. K., Chen, J. M., & Hehman, E. (2019). Same-sex marriage legalization associated with reduced implicit and explicit antigay bias. *PNAS Proceedings of the National Academy of Sciences of the United States of America, 116*(18), 8846–8851. https://doi.org/10.1073/pnas.1806000116

PricewaterhouseCoopers (2020). *Women in hospitality, travel, and leisure 2020: WiH2020 Review* (2020, October 1). Retrieved March 18, 2023, from https://www.pwc.co.uk/hospitality-leisure/documents/women-in-hospitality-travel-and-leisure-final-report.pdf

Restaurant Opportunities Centers United (2014). *The great restaurant divide*. (2014, October 22). Retrieved November 11, 2022, from https://rocunited.org/wp-content/uploads/sites/7/2022/06/GSD_Report_F1-Final.pdf

Restaurant Opportunities Centers United (2012). *Tipped over the edge: gender inequity in the restaurant industry*. (2012, February 12). Retrieved January 22, 2023, from https://rocunited.org/wp-content/uploads/sites/7/2022/10/ROC_GenderInequity.pdf

Rospenda, K. M., Richman, J. A., & Shannon, C. A. (2009). Prevalence and mental health correlates of harassment and discrimination in the workplace: results from a national study. *Journal of Interpersonal Violence*, *24*(5), 819–843. https://doi.org/10.1177/0886260508317182

Santero-Sanchez, R., Segovia-Pérez, M., Castro-Nuñez, B., Figueroa-Domecq, C., & Talón-Ballestero, P. (2015). Gender differences in the hospitality industry: a job quality index. *Tourism Management*, *51*, 234–246. https://doi.org/10.1016/j.tourman.2015.05.025

Statistics Canada. (2017). Women and paid work. *Women in Canada: Gender-based Statistical Report* (No. 89-503-X2015001). https://www150.statcan.gc.ca/n1/pub/89-503-x/2015001/article/14694-eng.pdf

Woods, R. H., & Kavanaugh, R. R. (1994). Gender discrimination and sexual harassment as experienced by hospitality-industry managers. *Cornell Hotel and Restaurant Administration Quarterly*, *35*(1), 16–21. https://doi.org/10.1016/0010-8804(94)90059-0