Check for updates

Original Research Article

Are Black Restaurant Servers Tipped Less Than White Servers? Three Experimental Tests of Server Race Effects on Customers' Tipping Behaviors

Cornell Hospitality Quarterly
1–15
© The Author(s) 2021
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/19389655211036652
journals.sagepub.com/home/cqx
SAGE

Zachary W. Brewster[1], Kenneth Gourlay[1], and Gerald Roman Nowak III[2]

Abstract

A limited number of published studies have presented evidence indicating that restaurant customers discriminate against Black servers by tipping them less than their White coworkers. However, the cross-sectional, localized, and small samples that were analyzed in these extant studies do not support any unqualified claim that consumer racial discrimination in tipping practices is a widespread phenomenon. Thus, in an effort to further clarify the relationship between restaurant servers' race and customers' tipping practices, we present results from three survey experiments designed to assess the causal effect of servers' race on customers' tipping intentions. In three independent, demographically diverse, and relatively large samples of U.S. consumers, we found no evidence to conclude that all else being equal consumers discriminate against Black restaurant servers by tipping them less than comparable White servers. Furthermore, the null effects of servers' race on customers' tipping practices were not found to be sensitive to variation in service quality, dining satisfaction, servers' sex, customers' sex, or customers' race. Our results challenge the generalizability of the previously observed server race effects on customers' tipping practices and point toward the need for future research that aims to advance our understanding of the conditions under which customers' tipping practices are sensitive to the perceived race of their server. The implications of our results for restaurant operations and directions for future research are also discussed.

Keywords

tipping; race; discrimination; restaurant; server; consumer discrimination

Introduction

Do restaurant customers discriminate against Black waiters and waitresses by tipping them less, on average, than their White counterparts? The answer to this question has important implications not only for the earnings and livelihood of Black workers[1] but also for the restaurants that employ them. In fact, the answer to this question has important implications for employers of tipped workers, more generally. If the answer is yes, and there are both theoretical and empirical reasons to believe that this may be the case, tipping as a system of remuneration may be in violation of Title VII of the Civil Rights Act of 1964 prohibiting employers from discriminating against employees as a result of their race, color, religion, sex, or national origin. In short, and as others have suggested (see Ayres et al., 2005; Brewster & Lynn, 2014; Lynn et al., 2008), reliance on customers' voluntary relinquishment of tips as a means of compensating employees, while a facially neutral business practice, may nonetheless produce a racially disparate impact on tipped employees' earnings and as such constitute a Civil Rights employment violation.

This interpretation of the law has gained the attention of legal scholars (see Kline, 2016; Nelson, 2020; Wang, 2014), and while it has not yet been tested in the courts, that may soon change. Recently, it was reported that the not-for-profit organization One Fair Wage (OFW, 2020) filed a federal Civil Rights violation complaint with the equal employment opportunities comission (EEOC) against Darden Restaurant Inc. alleging that the "company's practice of paying tipped workers a sub-minimum wage causes them to suffer more sexual harassment than non-tipped workers and leads employees of color to earn less in tips than their White co-workers" (Carman, 2020).[2] According to Saru Jayaraman, President of OFW, "Workers of color are segregated into lower-tipping restaurants, but even

[1]Wayne State University, Detroit, MI, USA
[2]Michigan State University, East Lansing, USA

Corresponding Author:
Zachary W. Brewster, Associate Professor, Department of Sociology, Wayne State University, 2272 Faculty/Administration Building, Detroit, MI 48202, USA.
Email: zbrewster@wayne.edu

when they get to the same restaurant, same shift, and same section, they're tipped less because of implicit [racial] bias" (Press, 2020).[3] Although additional details have not yet been made public, if the EEOC determines the complaint to have merit and it goes to litigation, OFW's legal council may present in support of their claim the results from three extant published studies that have reported evidence of racial discrimination in consumers' tipping practices.

In the first of these studies, Ayres et al. (2005) analyzed 1,066 tips given to 12 taxicab drivers in New Haven, Connecticut and found that Black drivers received one third less in gratuities than their White coworkers (Ayres et al., 2005). Moreover, Black drivers in this study were shown to be 80% more likely to be *stiffed*, or left no tip, than their White counterparts. Although the researchers controlled for the effects of a host of factors that have been shown to affect customers' tipping decisions, they were not able to rule out the possibility that the observed driver race effect on tipping was the result of interracial differences in service quality provided by taxi drivers. In other words, it is possible that the lower tips given to Black taxi drivers in this study reflected unobserved interracial differences in service skills (e.g., service quality) rather than consumer discrimination.

In an effort to conceptually replicate these results, Lynn et al. (2008) subsequently surveyed consumers ($n = 140$) leaving a full-service restaurant in the Southern region of the United States and asked them to report their servers' race (Black $n = 44$, White $n = 96$), appearance, friendliness, attentiveness, promptness, and the size of tip that they left at the end of the encounter. Net of the effects of service quality (e.g., servers' appearance, friendliness, attentiveness, and promptness), Black and White restaurant customers were found to tip Black servers less than they did White servers ($b = -3.25$, $p < .02$). Furthermore, the server race effect in this study was shown to be moderated by customers' perceptions of service quality and dining party size such that the White/Black disparity in tips was larger among bigger groups of diners and when service quality was rated highly.

In a third study, Brewster and Lynn (2014) surveyed a relatively larger sample of restaurant consumers ($n = 394$) as they left a full-service restaurant located in a large city in the Midwest. Once again, Black and White customers were observed to on average tip Black servers less than they did White servers ($b = -1.50$, $p < .05$), and this was true even though Black servers were on average perceived to have provided higher quality service. Results from subsequent exploratory analyses suggest that the seller race effect on tipping that was observed in this study was driven by the lower tips that Black consumers left Black servers, in contrast to White servers, when they were not satisfied with their dining experience.

Although customers may in fact tip Black servers on average less than White servers, as these studies' results suggest, such a conclusion is premature for a host of reasons. Each of the aforementioned studies analyzed small and localized samples that included only a limited number of servers who were working in a single restaurant establishment. The generalizability of these results is thus unknown. More importantly, the cross-sectional nature of the data analyzed in each of these studies precludes causal inferences regarding the relationship between service providers' race and customers' tipping practices. The relationships observed in these studies could have been the byproduct of the relationship between tipping practices and any number of unobserved customer, server, or restaurant characteristics. In short, the *cross*-sectional, localized, and small samples that were analyzed in these studies undermine any unqualified claim that customers' on average tip Black servers less than White servers because of their perceived race.

In response to these limitations, in this article, we introduce to the literature three additional estimates of the relationship between restaurant servers' race and customers' tipping practices. However, deviating from these extant studies, we offer the first experimental tests of whether customers do *on average* tip Black restaurant servers less than comparable White servers. Results from three survey experiments that were administered in independent, demographically diverse, and relatively large samples of U.S. consumers produced no evidence to conclude that consumers on average discriminate against Black restaurant servers by tipping them less than comparable White servers. Furthermore, the (null) effects of servers' race on customers' tipping intentions were not found to be sensitive to variability in service quality, dining satisfaction, servers' sex, customers' sex, or customers' race. Below, we outline the details and results of each of our experimental tests of the effects of servers' race on customers' tipping intentions. We conclude by discussing directions for future research and the applied implications associated with our inability to replicate previous observational results indicating that customers discriminate against Black servers by tipping them less than their White counterparts.

Study 1

Sample

This study analyzes survey data collected from an online sample of Survey Sampling International's (SSI) U.S. consumer panel. Stratified sampling was used to obtain an approximately equal number of responses from Asian, Black, Hispanic, and White consumers. In total, 914 respondents completed the survey. However, 49 cases were omitted because they did not answer one or more of the tipping questions that were part of the survey experiment assessing the effects of servers' race on customers' tipping practices. Additional participants ($n = 163$) were omitted from the analysis because they provided illogical and/or questionable

tip amounts across the within-subject dissatisfaction conditions (see description of within-subject experiment below).[4] These procedures resulted in a geographically diverse (48 states and the District of Columbia) final analytic sample of 702 U.S. consumers. As a result of our stratified sampling approach, Black ($n = 151$), Asian ($n = 164$), Hispanic ($n = 158$), and White ($n = 202$) participants were all well represented in our sample. Males and females were equally represented in our sample and 52% of the participants reported to have earned a bachelor's, master's, professional, or doctoral degree. The average respondent is 43 years of age ($SD = 13.13$) and reported a household income of between US$50,000 and US$60,000.

Experimental Manipulations

Embedded in the survey was a within- and between-subject experiment. Following a nondescript picture of a restaurant dining room, respondents randomly received a short hypothetical dining scenario involving a server who was a White male, Black male, White female, or Black female.[5] Participants' perceptions of their servers' race and sex were manipulated by using African American–sounding or White-sounding female and male names (e.g., Tamika and Jamal, Emily and Brad; see Bertrand & Mullainathan, 2004). In each scenario, the size of the bill was held constant (US$45.50). Immediately following the vignette participants were asked to report, in dollars and cents, how much they would tip their server if they were (a) satisfied with both the food and quality of service, (b) dissatisfied with both the food and quality of service, (c) satisfied with the food but dissatisfied with the quality of service, and (d) dissatisfied with the food but satisfied with the quality of service. Thus, tip amount was measured across four within-subject levels of dissatisfaction, two between-subject levels of servers' race (White, $n = 357$; Black, $n = 345$), and two between-subject levels of servers' sex (male, $n = 352$; female, $n = 350$). The average tip amount across the four levels of satisfaction/dissatisfaction ranged from US$0.00 to US$23.00 with a mean of US$5.81 ($SD = 2.33$).[6]

Results

To assess the role that servers' race plays in customers' tipping practices, we estimated a repeated-measures general linear model (GLM) predicting the amount (in dollars and cents) a respondent would reportedly tip across our four hypothetical experimental conditions involving satisfactory or dissatisfactory dining experiences. In this model, dissatisfaction is the within-subjects factor, and the hypothetical server's race and sex are the experimentally manipulated between-subjects factors. As shown in Table 1, a test of between-subjects effects produced by this analysis revealed no significant differences in tip amounts that respondents report they would leave to Black and White ($F = 0.061$, $p = .805$) nor male and female ($F = 0.251$, $p = .617$) servers. Furthermore, between-subject differences in tipping intentions were not found to be sensitive to the interactive effects of our manipulation of servers' race or sex ($F = 0.005$, $p = .946$). Finally, while the within-person manipulation of dissatisfaction with the food and/or service quality had a predictable strong effect on respondents' tipping intentions ($F = 864.77$, $p < .001$), this effect did not significantly vary as a function of the hypothetical servers' race ($F = 1.75$, $p = .168$), sex ($F = 0.364$, $p = .728$), or the interactive effects of servers' race and sex ($F = 0.624$, $p = .560$). We also estimated supplemental repeated-measures GLMs to assess whether our manipulations of servers' race and sex interact with respondents' race and sex to explain unique variance in within or between-subject tip amount. These models produced no reliable evidence that the effect of servers' race on customers' tipping intentions is conditioned by participants' race or sex.

In sum, this within- and between-subjects survey experiment produced no evidence that Black servers are on average tipped less than White servers. However, we unfortunately did not include an assessment of the effectiveness of our experimental manipulation of servers' race (or sex) in this study. If our use of Black- and White-sounding names failed to unequivocally signify servers' race (or sex), it might explain our inability to replicate the results of previous field studies (see Brewster & Lynn, 2014; Lynn et al., 2008). In response to this source of uncertainty, we present results from a second survey experiment that was also designed to assess the causal effect of servers' race (and sex) on consumers' tipping intention. In contrast to Study 1, the following experiment included a manipulation check to confirm that our respondents were attentive to the name of the server in the vignette and correctly associated the name to the racial category of Black or White.

Study 2

Sample

Participants for this study were recruited using Amazon Mechanical Turk. *MTurk* is an open-market website that allows organizations and individuals to act as *requesters* when they offer compensation to so-called *workers* in exchange for the performance of a variety of *human intelligence tasks* (*HITs*). Although not without concerns (e.g., see Aruguete et al., 2019; Rouse, 2015; Smith et al., 2016), researchers are increasingly using MTurk to conduct large sample studies at an economical rate. Moreover, MTurk samples tend to be more representative of the U.S. population than conventional convenience samples and are particularly well suited for experimental research (e.g., Berinsky et al., 2012; Casler et al., 2013; Hauser & Schwarz, 2016).

Table 1.
Study 1 Between and Within-Subject Effects of Servers' Race on Customers' Tipping Behaviors (n = 702).

Variable	Type III Sum of Squares	*df*	Mean Square	*F*	Significance
Between-subject effects					
Constant	94,039.84	1	94,039.84	4,319.50	.000
Black server	1.32	1	1.32	0.061	.805
Female server	5.46	1	5.46	0.251	.617
Black Server × Female Server	0.099	1	0.099	0.005	.946
Error (dissatisfaction)	55,196.15	698	21.77		
Within-subjects effects					
Dissatisfaction	8,205.70	2.33	3,515.48	864.77	.000
Dissatisfaction × Black Server	16.57	2.33	7.10	1.75	.168
Dissatisfaction × Female Server	3.45	2.33	1.48	0.364	.728
Dissatisfaction × Black Server × Female Server	5.92	2.33	2.54	0.624	.560
Error (dissatisfaction)	6,623.25	1,629.25	4.07		

Note. Results from a Mauchly's sphericity test indicated that the sphericity assumption underlying univariate tests of within-subject effects was violated. As such, Huynh–Feldt Epsilon corrections were used to adjust the degrees of freedom.

A total of 1,658 MTurk "workers" initiated participation in our study by clicking on the link to our questionnaire. However, 104 of these participants were omitted from the analysis because they reported that they were not living in the United States. An additional 198 respondents were omitted from the analysis because they failed to report how much they would tip the server in the hypothetical vignette (n = 120) or reported a questionably valid tip amount (n = 78).[7]

These procedures resulted in a large (n = 1,356) and demographically diverse sample of U.S. consumers (50 states, District of Columbia, and Puerto Rico). Males and females were equally represented in our sample. Nearly 70% of the sample self-reported to be non-Hispanic White, 10.6% as Asian, 10.1% as non-Hispanic Black or African American, 8% as Hispanic or Latino, and 2.7% of another race/ethnic group. Nearly 60% of the sample reported to have earned a bachelor's, master's, professional, or doctoral degree. The average respondent is 38 years of age (SD = 12.12) and reported a household income of roughly US$50,000.

Experimental Manipulations

The effect of servers' race on customers' tipping practices is further explored using a 2 (Server Race) × 2 (Server Gender) × 3 (Service Quality) between-subject factorial survey experiment. Specifically, respondents were asked to read a hypothetical dining vignette wherein the server's race (Black/White), server's gender (male/female), and service quality (excellent/average/poor) were manipulated. One of the 12 conditions was randomly displayed to each of our participants. Following the vignette, participants were asked a variety of questions related to their attitudes toward and likely behavioral reactions to the hypothetical dining scenario including how much, in dollars and cents, they would tip the server if their bill was $47.50. Reported tip amounts ranged from US$0.00 to US$20.55 with a mean of US$5.33 ($SD$ = 3.80). Servers' race and gender were manipulated with the use of first names. Names in this study were chosen based on the results of Gaddis' (2017) research on public perceptions of race and gender that are presumed to be associated with first and surnames. In the Gaddis (2017) study, MTurk participants were asked to consider a variety of names and for each indicate the race or ethnicity that they perceive the name to be associated. We selected two male and two female names that were "correctly" identified as being associated with Black or White Americans by more than 90% of Gaddis' (2017) participants. Thus, the hypothetical dining scenario that our respondents were asked to consider included a server named Hunter (White male, n = 354), Claire (White female, n = 328), Dashawn (Black male, n = 341), or Lakisha (Black female, n = 333). As shown in Table 2, service quality (poor, average, and excellent) was measured by manipulating the attentiveness, promptness, and friendliness of the waiter/waitress in the hypothetical dining scenario.

Manipulation Checks

To confirm that our respondents were cognizant of our manipulation of the servers' race and gender we asked them to recall the name of the server depicted in the hypothetical dining scenario. Of the 1,354 respondents that answered this question nearly all of them (98.6%) were able to correctly recall the name of the server portrayed in the vignette, they had read. At the very end of the survey, we also asked respondents to reflect back to the hypothetical dining scenario that they had read earlier and indicate, "if they had to guess,"

Table 2.
Study 2 Service Quality Manipulations.

Excellent Service Condition (n = 459)
Imagine that you are dining with a friend at a new corporate-owned family restaurant on a Friday night during its grand opening weekend. Upon walking in the door, you are promptly seated at a clean table. After you are seated, your server [Hunter/Claire/Dashawn/Lakisha] takes **no time at all to greet you.** When [Hunter/Claire/Dashawn/Lakisha] arrives at your table, [he/she] seems **extremely happy and engages you in playful conversation about the weather.** After you place your order, you wait **mere moments** before [Hunter/Claire/Dashawn/Lakisha] brings you your drink. Your food takes a long time to arrive at your table. When it finally does arrive, you are disappointed with your meal. Halfway through eating, your drink is empty. [Hunter/Claire/Dashawn/Lakisha] **brings you another beverage without having to be asked. Just as** you've finished your meal, [Hunter/Claire/Dashawn/Lakisha] brings you the check and you notice that the total bill is US$47.50. You place your payment down and [Hunter/Claire/Dashawn/Lakisha] **is right there to collect it**. After you've paid the bill, [Hunter/Claire/Dashawn/Lakisha] **cheerfully thanks you, wishes you a "fantastic day," invites you back**, and walks away.
Average Service Condition (n = 483)
Imagine that you are dining with a friend at a new corporate-owned family restaurant on a Friday night during its grand opening weekend. Upon walking in the door, you are promptly seated at a clean table. After you are seated, your server [Hunter/Claire/Dashawn/Lakisha] takes **only a little bit of time to greet you.** When [Hunter/Claire/Dashawn/Lakisha] arrives at your table, [he/she] seems **friendly and engages you in some obviously scripted conversation about the weather**. After you place your order, you wait for **a little bit** before [Hunter/Claire/Dashawn/Lakisha] brings you your drink. Your food takes a long time to arrive at your table. When it finally does arrive, you are disappointed with your meal. Halfway through eating, your drink is empty. [Hunter/Claire/Dashawn/Lakisha] **brings you another beverage after being asked. A few minutes after** you've finished your meal, [Hunter/Claire/Dashawn/Lakisha] brings you the check and you notice that the total bill is US$47.50. You place your payment down and [Hunter/Claire/Dashawn/Lakisha] **collects it shortly thereafter**. After you've paid the bill, [Hunter/Claire/Dashawn/Lakisha] **thanks you** and walks away.
Poor Service Condition (n = 426)
Imagine that you are dining with a friend at a new corporate-owned family restaurant on a Friday night during its grand opening weekend. Upon walking in the door, you are promptly seated at a clean table. After you are seated, your server [Hunter/Claire/Dashawn/Lakisha] **takes a very long time to greet you**. When [Hunter/Claire/Dashawn/Lakisha] arrives at your table, [he/she] **seems annoyed and makes no attempt to converse with you**. After you place your order, you wait for **what seems like an eternity** before [Hunter/Claire/Dashawn/Lakisha] brings you your drink. Your food takes a long time to arrive at your table. When it finally does arrive, you are disappointed with your meal. Halfway through eating, your drink is empty. [Hunter/Claire/Dashawn/Lakisha] **never brings you another beverage. Fifteen minutes after** you've finished your meal, [Hunter/Claire/Dashawn/Lakisha] brings you the check and you notice that the total bill is US$47.50. You place your payment down and **must flag [Hunter/Claire/Dashawn/Lakisha] down to collect it**. After you've paid the bill, [Hunter/Claire/Dashawn/Lakisha] **says nothing to you** and walks away.

Note. The service quality manipulations are represented in bold.

whether their server was likely a "White male," "Black male," White female," "Black female," Hispanic male," or a "Hispanic female." Seventy-seven percent (n = 1,043) of our participants correctly "guessed" the race and gender of the server depicted in the vignette thus indicating that our manipulation of servers' race and gender was successful.[8] In an effort to assess the success of our manipulation of service quality, we asked respondents to rate on a 7-point scale (0 = terrible, 7 = excellent) the service provided by Hunter (Claire, Dashawn, or Lakisha) in the hypothetical dining scenario. The average service rating in the poor service condition was 1.00 (SD = 1.49), in the average service condition, it was 2.81 (SD = 1.41), and in the excellent service condition it was 4.62 (SD = 1.41). All mean differences were statistically significant ($p < .001$), thus indicating that our manipulation of service quality was also effective.[9]

Results

As expected, and as shown in Table 3, we observe a strong main effect of our manipulation of service quality (F = 145.63, $p < .001$) on participants' reported tip amount. However, we find no evidence that Black servers are, on average, tipped less than their White counterparts (F = 0.509, p = .476) and this is the case across poor, average, and excellent service quality conditions (F = 1.49, p = .227). The model did produce a significant interaction effect between servers' race and servers' sex on tip amount (F = 5.18, p = .023). An analysis of the estimated marginal means indicated that respondents would tip the hypothetical White waitress (US$5.37) nominally better than the comparable Black waitress (US$5.08). In contrast, our respondents reported that they would leave Black waiters more (US$5.61) than comparable White waiters (US$5.04). The

Table 3.
Study 2 Between-Subject Effects of Servers' Race on Customers' Tipping Behaviors (n = 1,356).

Variable	Type III Sum of Squares	*df*	Mean Square	*F*	Significance
Corrected model	3,580.68	11	325.52	27.34	.000
Constant	37,584.74	1	37,584.74	31.57.07	.000
Service quality	3,467.35	2	1,733.67	145.63	.000
Black server	6.06	1	6.06	0.509	.476
Female server	3.17	1	3.17	0.266	.606
Service Quality × Black Server	35.38	2	17.69	1.49	.227
Female Server × Black Server	61.71	1	61.71	5.18	.023
Service Quality × Female Server	7.57	2	3.79	0.318	.728
Service Quality × Black Server × Female Server	26.47	2	13.23	1.11	.329
Error	16,000.27	1,344	11.91		
Total	58,120.80	1,356			
Corrected total	19,580.96	1,355			

three-way interaction involving service quality, servers' race, and sex was also not significant (F = 0.318, p =.728) in this model. Finally, we estimated supplemental GLMs to assess whether the effects of servers' race on tip amount differ as a function of customers' race and/or sex. We found no reliable differences in the effect of servers' race (Black/White) on reported tip amount across Asian, Black, Hispanic, White, and "other" respondents (F = 0.210, p < .993), nor did we find a reliable difference in the effect of servers' race on reported tip amount across male and female respondents (F = 1.73, p = .188).

The substantive conclusions derived from these results corroborate those from Study 1 and cast an additional layer of doubt on the claim that customers' discriminate against Black servers by tipping them less than Whites. However, while our manipulation check in Study 2 confirmed that our use of racially distinct first names was effective in that respondents overwhelmingly reported that Hunter/Claire and Dashawn/Lakisha were likely associated with White and Black individuals, respectively, racially distinct names may nonetheless be insufficient to induce an effect on customers' tipping intentions. That is, a racially suggestive name may not trigger a cognitive and behavioral response in respondents that is equivalent to seeing a server in the flesh. In an effort to assess whether visual cues of race have a more pronounced effect than racialized names, we present results from a third survey experiment that manipulated servers' race and sex with photographs of a White and Black waiter and waitress.

Study 3

Sample

Participants for this study were also recruited using Amazon Mechanical Turk.[10] A total of 540 MTurk "workers" initiated participation in our study by clicking on the link to our questionnaire. However, 25 of these participants were omitted from the analysis because they took the survey twice (n = 4) or completed it in (n = 21) 120 s or less (median = 319 s). An additional 98 cases were omitted from the analysis because they either failed to report a tip in one or more of the within-subject satisfaction/dissatisfaction conditions or provided questionable tip amounts.[11] These procedures resulted in 417 demographically diverse U.S. consumers (50 states, District of Columbia, and Puerto Rico). Males (52%) and females (48%) were equally represented in our sample. Nearly 70% of the sample self-reported to be non-Hispanic White, 12% as Asian, 11% as non-Hispanic Black or African American, 6% as Hispanic or Latino, and 2.4% of another race/ethnic group.

Experimental Manipulations

We asked respondents to imagine that they had taken a friend or family member out for dinner at a local restaurant and showed them a photograph of a hypothetical waiter/waitress that we indicated had been assigned to their table. The photograph that was randomly displayed was of a White male (n = 101), White female (n = 99), Black male (n = 94), or Black female (n = 123). The four models that were depicted in the photographs were selected from Adobe Stock.[12] The White and Black male models were each depicted wearing formal attire (e.g., bowtie), smiling, and holding a tray with a glass of wine and a salad. The Black and White female models were photographed smiling, in the same (casual) restaurant, in similar outfits, and in similar poses (e.g., holding checkbook).[13] Like Study 1, following the vignette, respondents were asked to report, in dollars and cents, how much they would tip the server if their bill was US$57.54 and they were (a) satisfied with both the food and quality of service, (b) dissatisfied with both the food and quality of service, (c) satisfied with the food but

Table 4.
Study 3 Between- and Within-Subject Effects of Servers' Race on Customers' Tipping Behaviors (n = 417).

Variable	Type III Sum of Squares	*df*	Mean Square	*F*	Significance
Between-subject effects					
Constant	105,497.26	1	105,497.26	2,136.38	.000
Black server	53.16	1	53.16	1.08	.300
Female server	118.93	1	118.93	2.41	.121
Black Server × Female Server	118.34	1	118.34	2.40	.122
Error (dissatisfaction)	20,394.51	413	49.38		
Within-subjects effects					
Dissatisfaction	8,084.00	1.91	4,244.19	684.23	.000
Dissatisfaction × Black Server	1.56	1.91	.818	0.132	.867
Dissatisfaction × Female Server	16.83	1.91	8.83	1.42	.242
Dissatisfaction × Black Server × Female Server	16.97	1.91	8.91	1.44	.239
Error (dissatisfaction)	4,879.48	789.65	6.20		

Note. Results from a Mauchly's sphericity test indicated that the sphericity assumption underlying univariate tests of within-subject effects was violated. As such, Huynh–Feldt Epsilon corrections were used to adjust the degrees of freedom.

dissatisfied with the quality of service, and (d) dissatisfied with the food but satisfied with the quality of service. The average tip amount across the four within-subject levels of satisfaction/dissatisfaction ranged from US$0.00 to US$20.00 with a mean of US$7.96 (SD =3.52).

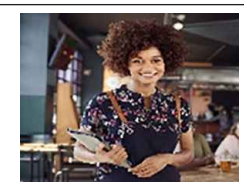
Results

To assess the role that servers' race plays in customers' tipping practices, we estimated a repeated-measures GLM predicting the amount (in dollars and cents) a respondent would reportedly tip across our four hypothetical experimental conditions involving satisfactory or dissatisfactory dining experiences. As shown in Table 4, a test of between-subjects effects produced by this analysis revealed no significant differences in tip amounts that respondents report they would leave to Black and White ($F = 1.08$, $p = .300$) nor male and female ($F = 2.41$, $p = .121$)[14] servers. Furthermore, between-subject differences in tipping intentions were not found to be sensitive to the interactive effects of our manipulation of servers' race and sex ($F = 2.40$, $p = .122$). Finally, although the within-person manipulation of dissatisfaction with the food and/or service quality had a predictable strong effect on respondents' tipping intentions ($F = 684.232$, $p < .001$), this effect did not significantly vary as a function of the hypothetical servers' race ($F = 1.32$, $p = .867$), sex ($F = 1.42$, $p = .242$), or the interactive effects of servers' race and sex ($F = 1.44$, $p = .239$). We also estimated supplemental repeated-measures GLMs to assess whether our manipulations of servers' race (and sex) interact with respondents' race or sex to explain unique variance in within or between-subject tip amount. These models produced no reliable evidence that the effect of servers' race (or sex) on customers' tipping intentions is conditioned by participants' race or sex.

Discussion

In three survey experiments, each administered to relatively large and demographically diverse samples of U.S. consumers, we were unable to replicate results from previously published correlational studies on consumer racial discrimination in tipping behaviors. Specifically, in contrast to previously published results (Ayres et al., 2005; Brewster & Lynn, 2014; Lynn et al., 2008), we find no evidence that ceteris paribus consumers discriminate against Black servers by tipping them less than comparable White servers. Our findings suggest that this is the case for both waitresses and waiters and across experimentally manipulated within and between-subject levels of dining satisfaction (Studies 1 and 3) and service quality (Study 2). Although we did observe a significant interaction effect between servers' race and servers' sex on tip amount in Study 2, the nature of this effect was more advantageous than detrimental to Black servers' tipped income. Although Black waitresses were tipped nominally less than White Waitresses, the interaction effect was largely driven by the significantly better tips that respondents reported they would leave Black waiters relative to comparable White waiters. Although not significant, the nature of this interaction effect was replicated in our third experiment wherein servers' race and sex were manipulated with photographs. Furthermore, we found no evidence to suggest that the effects of servers' race on tipping behaviors differed across male and female respondents or across Asian, Black, Hispanic, or White respondents.[15] In short, and as summarized in Table 5, we find no evidence to suggest that consumers' tipping decisions are on average sensitive to their server's race (or sex[16]). In fact, all else being equal, the mean difference in tips that the respondents in our studies

Table 5.
Summary of Results From Three Survey Experiments Testing the Causal Effect of Servers' Race (and Sex) on Restaurant Customers' Tipping Behaviors.

Study 1 (bill = US$45.50)	White Male (n = 173)	Black Male (n = 179)	White Female (n = 184)	Black Female (n = 166)
Satisfied with food and service	US$8.36 (18.4%)	US$8.55 (18.8%)	US$8.21 (18.0%)	US$8.44 (18.5%)
Satisfied with service/dissatisfied with food	US$6.16 (13.5%)	US$6.12 (13.5%)	US$6.33 (13.9%)	US$5.96 (13.1%)
Satisfied with food/dissatisfied with service	US$4.96 (10.9%)	US$4.82 (10.6%)	US$4.71 (10.4%)	US$4.73 (10.4%)
Dissatisfied with food and service	US$3.91 (8.6%)	US$3.78 (8.3%)	US$3.85 (8.5%)	US$3.76 (8.3%)
Study 2 (bill = US$47.50)	**White Male (n = 354)**	**Black Male (n = 341)**	**White Female (n = 328)**	**Black Female (n = 333)**
Excellent service	US$6.54 (13.8%)	US$7.89 (16.6%)	US$6.99 (14.7%)	US$6.82 (14.4%)
Average service	US$5.64 (11.9%)	US$5.65 (11.9%)	US$5.73 (12.1%)	US$5.55 (11.7%)
Poor service	US$2.95 (6.2%)	US$3.27 (6.9%)	US$3.40 (7.2%)	US$2.87 (6.0%)
				
Study 3 (bill = US$57.54)	(n = 101)	(n = 94)	(n = 99)	(n = 123)
Satisfied with food and service	US$11.00 (19.1%)	US$11.61 (20.2%)	US$10.72 (18.6%)	US$10.68 (18.6%)
Satisfied with service/dissatisfied with food	US$9.21 (16.0%)	US$10.00 (17.4%)	US$8.78 (15.3%)	US$8.76 (15.2%)
Satisfied with food/dissatisfied with service	US$6.05 (10.5%)	US$6.93 (12.0%)	US$6.20 (10.8%)	US$5.97 (10.4%)
Dissatisfied with food and service	US$5.01 (8.7%)	US$6.30 (10.9%)	US$5.56 (9.7%)	US$5.14 (8.9%)
Tip amount (percent) mean of means	US$6.34 (12.51%)	US$6.81 (13.37%)	US$6.41 (12.65%)	US$6.24 (12.32%)

reported that they would leave a hypothetical Black (n = 1,236, US$6.42) and White ($n$ = 1,239, US$6.29) restaurant server was only 13 cents ($CI_{95\%}$ = −0.607, 0.347). After taking into account sample size differences across our experiments, the meta-estimate of the server race effect on tipping was for all practical purposes reduced to zero (mean difference = US$0.05, 95% CI = −0.263, 0.154).[17]

Although our results and substantive conclusions diverge from those of Lynn et al. (2008) and Brewster and Lynn (2014), they are nonetheless consistent extant studies that have included a measure of servers' race as a non-focal variable in their analyses. For instance, Lynn and McCall (2009) surveyed a large sample of current and former restaurant servers (n = 1,606) and asked them to indicate on a 7-point scale (1 = much larger than most others' tips; 7 = much smaller than most others' tips) how the tips they earn compared to those earned by their coworkers. The authors found no difference in White and non-White servers' perceptions of their tips relative to the tips of other servers in their restaurant. Given the overrepresentation of Whites employed as servers or bartenders, the coworkers that many non-White respondents in this study were juxtaposing their wages against were likely to be White. In another large sample (n = 954) of current and former restaurant servers, Brewster (2015) failed to

observe a meaningful difference in White and non-White respondents' reported average tip percentage that they receive from customers at their restaurant. Furthermore, Parrett (2011) analyzed survey data ($n = 295$) collected from patrons as they left five Virginia restaurants and found no effect of servers' race on the customers' tipping behavior. Finally, using a consumer diary methodology, Lynn et al. (2012) analyzed data on 466 dining experiences nested within 122 consumers (students) and found no evidence to suggest that consumers tip non-White severs differently than White servers and this was the case across within-subject ratings of service quality. Although the racially homogeneous nature of the samples analyzed in these studies[18] precluded the authors from comparing server or customer reported tips across White and Black servers, when coupled with the experimental results that we report in the current research, these extant findings casts an added layer of doubt on the generalizable nature of consumers' proclivities to tip Black servers less than their White counterparts.

Our use of a vignette experimental design allowed for a controlled assessment of unconditional and conditional causal effects of servers' race on consumers' tipping intentions in three relatively large and demographically diverse samples of U.S. consumers. Of course, it would have been preferable to do such a controlled test using real-life service encounters involving actual White and Black restaurant servers. Needless to say, however, a test using real-world dining scenarios that controlled for all environmental, social, and psychological confounds would have been quite difficult (impossible) to execute for a variety of reasons. As such, the high internal validity that we achieve with our design comes at a potential cost to the external validity of our results. In other words, it is possible that our respondents' actual tipping decisions would differ from the decisions that we capture with our measures of customers' tipping intentions under hypothetical, controlled, and manipulated conditions. If this is the case, it might explain why we were not able to replicate Lynn et al. (2008) and Brewster and Lynn's (2014) results. However, notwithstanding this as a possibility, we have no reason to believe that our measures of customers' tipping intentions under hypothetical conditions are not reliable proxies or strong predictors of how consumers would actually tip in comparable real-life scenarios.

That extant research has documented a strong relationship between what people say they would do under hypothetical conditions and what they actually do if confronted with a similar situation (e.g., see Brauer et al., 2019; Hainmueller et al., 2015; Murphy et al., 1986) gives us some confidence in the external validity of our results. Our confidence is bolstered by the fact that our respondents' tipping intentions were observed to be sensitive to our within and between-subject manipulations of service quality/satisfaction—relationships that have been widely documented in the extant tipping literature (see Banks et al., 2018). Furthermore, in both the SSI and MTurk samples, Black participants reported that they would leave the hypothetical server a significantly smaller tip than White participants.[19] That we were able to replicate a Black–White difference in tipping that has been observed in many previous studies and that have used a variety of different methodologies (see Banks et al., 2018; Brewster & Nowak, 2019; Lynn, 2006) lends further support to the validity of our participants' reported tipping decisions under hypothetical conditions. Although we have no strong reasons to believe that our results are not generalizable, we of course recognize that the external validity of our results is an outstanding empirical question. Thus, we hope that additional estimates of the effects of service providers' race on customers tipping behaviors are generated in forthcoming studies utilizing different sampling approaches and innovative methods of observation.

Although our results call into question the generalizability of previously documented server race effects on customers' tipping practices, they should not be interpreted as an indication that Black servers are never tipped less than their White coworkers simply because of their perceived race. In fact, while the results reported by Lynn et al. (2008), and Brewster and Lynn (2014) may reflect Type I errors[20] it is nevertheless the case that Black servers (or taxi drivers, see Ayres et al., 2005) in these studies were observed to be tipped less than their White counterparts even after controlling for a wide range of potential confounding factors. Furthermore, anecdotes of racially motivated mistreatment of Black restaurant servers surface routinely in the popular press. For instance, a table of White customers at an Olive Garden in Evansville, Indiana reportedly refused to be waited on by a Black server.[21] Had the Olive Garden manager not acquiesced to the consumer's racist demands by re-assigning a White server to the table, it is reasonable to assume that the table would not have tipped this Black server fairly. In a similar event, four customers at a Kentucky Applebee's stiffed their Black waitress and left a note on a paper napkin that read: "We don't tip Black people."[22] Just recently, a Black waiter working at a Buffalo Wild Wings in Northern Kentucky was stiffed on a US$68.41 bill and in lieu of a signature on the credit card slip, the customer wrote—*"All Lives Matter."*[23]

In short, our results do not undermine the assumption that there are consumers who will, under some conditions, consciously or unconsciously discriminate against Black servers by tipping them less simply because they are perceived to be Black. This assumption is not only buttressed by the extant studies and anecdotes cited above but also by a large body of social science research that has documented the enduring nature of racial prejudice (implicit and explicit) and discrimination in America across a variety of social milieus. As such, we encourage future research that aims to identify and test a wide range of individual, interactional,

and environmental conditions that conjoin with servers' race to produce such discriminatory consumer behaviors. Research in this vein should be cognizant that consumers' propensities to racial discrimination in their tipping practices may not always be disadvantageous to Black servers. That is, if our assumption that some customers, under some conditions, do discriminate against Black servers by tipping them less than White servers is correct, our results imply that other customers may tip Black servers' more than they otherwise would have tipped them had they been White. As such, studies that aim to tease out the countervailing conditions under which Black servers may be the beneficiaries of better tips vis-à-vis White servers are also needed if the role that servers' race plays in customers' tipping decisions is to be fully understood.

We also encourage research that moves beyond the direct role that customers may play in producing differences in tip earnings across Black and White restaurant servers. In the restaurant milieu, there are multiple points of contact aside from servers' interactions with customers that could logically contribute to a racial differential in earnings that have not to date been scrutinized. For instance, restaurant managers may consciously or unconsciously bolster the tips of White servers, at the expense of their Black coworkers, by scheduling them to work the busiest shifts (e.g., weekends) or by assigning them larger and/or more lucrative sections of the restaurant (and/or less favorable, e.g., distance to kitchen, see Kim et al., 2017 see also Alexander v. Casino Queen, Inc.).[24] In fact, just recently two Capital Grille employees filed a complaint with the EEOC alleging such discriminatory treatment. One of the complainants, a Latinx server working in a Washington D.C. Capital Grille, alleges that she was "consistently assigned to sections of the restaurant known to generate less in tips, which management referred to as "Section 8" or "my low-income world" (Carman, 2020).[25] Future research may also find that Black servers' earnings are sometimes undermined by hosts who seat them less often and/or with customers who are known to be below average tippers (e.g., Black customers, see Carman, 2020), by bussers who are slow to clear their tables thereby inhibiting them from "turning" them quickly (see Gatta, 2002), and/or by cooks who control the timing and quality of the food that is delivered to their tables (see Wilson, 2019). In short, we encourage future research that considers a much broader range of factors that may contribute to interracial differences in restaurant servers' tip earnings than have been considered to date.

Notwithstanding these avenues for further research, the fact that we did not experimentally observe a meaningful difference in the average size tip that customers reported they would leave Black and White servers has important implications for the operations of restaurant establishments. In an era wherein restaurant owners and operators are increasingly being asked to consider replacing voluntary tipping with alternative compensation policies, it is important to scrutinize extant empirical evidence used to advance the narrative of both advocates and opponents of voluntary tipping. To this end, the current studies' results suggest that race-based discrimination in customers' tipping practices need not occupy a dominant place in all restaurant operators' deliberations about replacing voluntary tipping with alternative remuneration systems. This is not to suggest that there are not good reasons to consider abolishing voluntary tipping (Lynn, 2017) but rather that racial discrimination in customers' tipping practices by itself may not be a particularly strong reason for many, or most, restaurants to do so.

Similarly, our results should dampen concerns that voluntary tipping, a facially neutral compensation policy, may render establishments vulnerable to Title VII disparate impact litigation and liability by producing a wage differential for Black and White servers (see Brewster & Lynn, 2014; Kline, 2016; Lynn et al., 2008; Nelson, 2020; Wang, 2014). In addition to other challenges, for a plaintiff to maintain a disparate impact case under Title VII, the courts would need to determine that there is sufficient statistical evidence to establish the plaintiff's claim (see Kline, 2016; Nelson, 2020; Wang, 2014). Given our inability to experimentally replicate research by Ayres et al. (2005), Lynn et al. (2008), and Brewster and Lynn (2014), it is questionable whether a court could be convinced by the existing body of published scholarship, *alone*, that employers' practice of relying on customers' voluntary gratuities to compensate servers produces a generalizable discriminatory outcome in Black and White servers' wages.

Of course, this situation could change as the body of literature in this area continues to develop and the individual and restaurant-level boundary conditions of customers' propensities to discriminate racially in their tipping practices become better understood. Furthermore, a plaintiff who litigates a particular claim might be able to produce sufficient evidence that Black servers in their specific restaurant (or chain of restaurants) suffer from lower tip earnings than their White coworkers (see Kline, 2016). If such evidence were to be accepted, employers could encounter difficulties establishing that relying on voluntary tipping to compensate their servers is job related for the position in question and consistent with business necessity, and, therefore, have difficulty defending the practice (see Kline, 2016; Nelson, 2020; Wang, 2014). As such, we echo Lynn et al. (2008, p. 1057) in encouraging restaurants to be proactive and protect themselves from any potential threat of disparate impact litigation by systematically monitoring the tips that their White and non-White employees receive. At minimum, restaurant operators could easily track the credit card tips that are left to their wait staff and in the event that such efforts unveil a racial difference that cannot readily be accounted for the restaurant should implement a redistributive tipping policy (e.g., tip pooling), add automatic gratuities to customers' checks, or abolish the practice of voluntary tipping

altogether in their establishments (see Nelson, 2020 for additional, and innovative, strategies that could be considered to address racial disparities in tip earnings). Given the potential cost associated with litigation, restaurant operators have little to lose by being proactive on this front.

In closing, we encourage the restaurant industry to actively address what is likely to be a greater threat to the earnings of Black restaurant servers/bartenders than is posed by consumer racial discrimination in tipping—the underrepresentation of Black and Brown-skinned servers working in lucrative high-end dining establishments.[26] The underrepresentation of Black servers in fine dining establishments is, in part, an unfortunate outgrowth of the racially and economically segregated neighborhoods where people live. Fine dining restaurant establishments are simply less likely to be a viable employment option for many Black restaurant workers vis-à-vis their White peers. Although operators of fine dining establishments cannot solely be expected to address this source of Black servers' underrepresentation in their establishments, there are things that operators can and should be doing in an effort to further diversify the makeup of their wait staff.

First and foremost, fine dining operators (in particular) need to review their hiring, placement, and promotion practices and implement policies/procedures intended to ensure that there is not a conscious or unconscious preference for White applicants over equally qualified and deserving Black applicants. The need to do so is clear. Restaurant Opportunities Centers United, for instance, recently conducted 400 matched pairs audit tests for discriminatory hiring practices in fine dining restaurant establishments located in New York, Chicago, Detroit, and New Orleans. Results from these audit tests demonstrated that White applicants were more likely to be interviewed and if interviewed were treated better by management and extended more time to speak than were applicants of color. Moreover, White applicants were found to be twice as likely to be offered a position as a server or bartender compared with equally (or better) qualified applicants of color applying for the same position within the same fine dining establishment (see, for example, Bendick et al., 2010; Jayaraman, 2016).

In addition to categorical exclusion, applicants of color have been shown to be the victims of racial channeling, a term used to describe the process by which they are steered into lower paying, non-customer contact, positions (e.g., busser, runner, dishwasher) rather than being considered for the more lucrative serving or bartending position that they applied for or inquired about (Pager et al., 2009). Finally, once employed in black-of-the-house, people of color encounter limited opportunities for promotion into the coveted high paying front-of-the-house positions (Jayaraman, 2013, 2016) In short, intervening where restaurant consumers discriminate against Black servers and bartenders by tipping them less than their White counterparts is not likely to be a sufficient course of action, by itself, to attenuate notably the aggregate racial wage disparity that has been observed among tipped restaurant workers. To do so, fine dining establishments will necessarily have to take a leading role by taking steps to ensure racial equality in their hiring, placement, and promotion processes. We hope that this research encourages such leadership.

Acknowledgements

The authors are grateful to Lu-in Wang and anonymous CQ referees for their helpful feedback on earlier drafts of this article.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, or publication of this article.

ORCID iD

Zachary W. Brewster https://orcid.org/0000-0002-2096-7767

Notes

1. Based on an analysis of Census data, One Fair Wage (OFW) estimates that White male tipped restaurant workers earn an average hourly wage of US$16.48 per hour. Compared with their White male counterparts, Black males, White females, and Black females earn on average US$2.76 (or 16.75%), US$4.48 (or 27.18%), and US$4.79 (or 29.07%) less per hour, respectively (see https://onefairwage.com/wp-content/uploads/2020/07/LegacyOfSlavery_4.pdf).
2. https://www.washingtonpost.com/news/voraciously/wp/2020/09/30/parent-company-of-olive-garden-violates-the-civil-rights-act-with-its-tipping-policies-activists-say/?outputType=amp
3. https://jacobinmag.com/2020/10/tipped-restaurant-workers-waiters-coronavirus-subminimum-wage
4. Reported tip amounts were considered entry errors, nonserious, exaggerated, or otherwise questionable estimates if they exceeded 50% of the hypothetical bill size ($n = 17$). The within-subject experiment design also included two data quality checks. First, response patterns were considered illogical and cases omitted if a respondent reported a higher tip amount in the dissatisfied with food and service condition than in the satisfied with the food and service condition. Second, response patterns were considered illogical and cases omitted if a respondent reported a lower tip amount under the dissatisfied with only food/service condition than in the dissatisfied with both the food and service condition. The 163 cases that were omitted as a result of data quality concerns were distributed similarly across the White ($n = 91$) and Black ($n = 72$) experimental conditions ($\chi^2 = 1.31$, $p = .252$). As such, the substantive conclusions derived from our analysis should not be sensitive to the omission of these cases.
5. The prompt read as follows: "Image that you take a friend or family member to the above-pictured restaurant for dinner. After being seated at a table, a [waiter/waitress] in [his or her] mid-20s greets you by saying, 'Hello, my name is Jamal/

Brad/Tamika/Emily and I will be taking care of you tonight.' After you finish eating [Jamal/Brad/Tamika/Emily] delivers your bill and you see that the total is US$45.50."

6. In a previously published analysis of these experimental data, Brewster and Brauer (2017) demonstrated that the within-person effect of dissatisfaction (or the closely related concept of service quality) on tip amount differed significantly across Black and White respondents. Specifically, the authors showed that the within-person effect of dissatisfaction on respondents tipping intentions was moderated by customers' race such that the tips of Black customers were more sensitive to dissatisfactory conditions than their White counterparts. While not the focus of Brewster and Brauer's (2017) study, the authors did include the manipulation of servers' race in their analysis as a covariate. Results showed that there was not a significant difference in reported tips across the Black and White server conditions. In this study, Brewster and Brauer (2017) analyzed data provided only by Black and White participants and did not test for any multiplicative effects involving servers' race on reported tip amount. The current study analyzes a more race/ethnic diverse sample of consumers (Black, Asian, White, and Hispanic) and importantly we also test for multiplicative effects involving our manipulation of servers' race.
7. Reported tip amounts were considered invalid (e.g., entry errors, non-serious, exaggerated, or otherwise questionable estimates) if they exceeded 50% of the hypothetical bill size (US$47.50). The 198 cases that were omitted because of missing information of data quality concerns were distributed similarly across the White ($n = 104$) and Black ($n = 94$) experimental conditions ($\chi^2 = .334$, $p = .558$). As such, the substantive conclusions derived from our analysis should not be sensitive to the omission of these cases.
8. Participants' in the White server experimental condition (82.6%) were more likely to correctly guess the race/sex of the hypothetical server than those in the Black server condition (71.2%) and this difference was statistically significant ($\chi^2 = 24.53$, $p < .000$). This difference may reflect social desirable reporting. In an effort to present themselves as people whose judgments are not influenced by racial stereotypes, in this case, those that are associated with first names, participants asked to consider a scenario involving Dashawn or Lakisha may have consciously refrained from identifying these names as "Black names." In an effort to assess whether the results we present are sensitive to this potential source of reporting bias we estimated our general linear model (GLM) after omitting respondents who failed this manipulation check. The substantive conclusions derived from this model were the same as those that we present.
9. The zero-order correlation between our experimental measure of service quality and respondents' rating of the service depicted in the hypothetical scenario was .71 ($p < .000$).
10. What we present here is a component of a larger survey experiment that was designed to test the effects of wearing a mask on restaurant servers' tipped incomes. The experiment included a mask manipulation across a White, Black, and Hispanic male and female. As such, the original experiment included 12 randomized conditions. In this article, we present results from an analysis of only the conditions that included photographs of a Black or White male or female who was not wearing a mask ($n = 540$).
11. Reported tip amounts were considered entry errors, non-serious, exaggerated, or otherwise questionable estimates if they exceeded 50% (US$28.77) of the hypothetical bill size. In accordance with Study 1, tipping response patterns were also considered illogical and cases omitted if a respondent reported a higher tip amount in the dissatisfied with food and service condition than in the satisfied with the food and service condition. Second, response patterns were considered illogical and cases omitted if a respondent reported a lower tip amount under the dissatisfied with only food/service condition than in the dissatisfied with both the food and service condition. The 123 cases that were omitted because of missing information or data quality concerns were distributed similarly across the four experimental conditions ($\chi^2 = 1.02$, $p = .797$). As such, the substantive conclusions derived from our analysis should not be sensitive to the omission of these cases.
12. https://stock.adobe.com/license-terms
13. Despite the photographic similarities across the Black and White male (female) conditions, it is nonetheless possible that our manipulation of race is confounded with other individual differences between the male (female) models. Although we are not able to unequivocally rule this possibility out we did ask respondents to report on a 7-point scale their perception of the hypothetical servers' friendliness (1 = extremely unfriendly, 7 = extremely friendly), attractiveness (1 = extremely unattractive, 7 = extremely attractive), and overall appearance (1 = dislike a great deal, 7 = like a great deal). On average, respondents' ratings of the photographed Black and White males were comparable for our measures of friendliness ($p = .368$), attractiveness ($p = .073$), and overall appearance ($p = .932$). On average, respondents likewise rated the Black and White females similarly in terms of friendliness ($p = .931$), attractiveness ($p = .609$), and overall appearance ($p = .338$). These results give us added confidence that the Black/White males and Black/White females depicted in the photographs were indeed perceived to be comparable with the exception of their race. Nonetheless, future research should consider manipulating the race with photographs of the same person to avoid confounding race with other individual differences (e.g., FaceApp, see https://www.neogaf.com/threads/faceapp-apparently-now-has-race-filters-i-never-got-the-humor-in-these.1417700/page-6).
14. The effect of servers' sex on tip amounts in this study should be interpreted with caution. The glass of wine and the formal attire (e.g., bowtie) worn by the male models likely functioned to signal that the scenario was in a fine dining restaurant whereas the females were photographed in what clearly appears to be a casual restaurant. In short, these associations may be the outcome of the confounding of servers' sex with other factors that may be associated with customers' tipping behaviors (e.g., restaurant expensiveness, décor, etc.).
15. We also estimated supplemental models to assess whether the effects of servers' race on customers' tipping intentions were conditioned by respondents' age, income, education, or region of residence. Controlling for the effects of our manipulations of dining satisfaction/service quality and servers' sex, these models produced no evidence to suggest that the effects of

servers' race on respondents' tipping intentions are sensitive to these demographic characteristics. In fact, the only interaction effect that approached statistical significance was that between servers' race and respondents' education (b = .254, p =.09) in Study 2. As the educational level of respondents in Study 2 increased they were more likely to report that they would tip the hypothetical Black server nominally more than the comparable White server. In short, the null effect of servers' race on tipping intentions that we document across three survey experiments is robust across respondents' age, region of residence (Studies 1–3), income, and education (Studies 1 and 2).

16. While not our focus, the results of our studies also offer to the extant literature three experimental tests of server sex effects on customers' tipping behaviors. On average, respondents in each of the three studies reported that they would tip male and female servers comparably. These results are consistent with those of many other studies that have observed null effects of servers' sex on customers' tipping behaviors (e.g., Conlin et al., 2003; Kim et al., 2017; Lynn et al., 2008, 2011, 2012; Lynn & McCall, 2016; Lynn & Simons, 2000; Parrett, 2011; Seiter & Weger, 2010). However, a recent meta-estimate of the servers sex effect on tipping that was based on the results from nine studies that collectively included nearly 1,400 cases does suggest that female serves may on average be tipped nominally better than their male counterparts (Cohen's d = .06; see Banks et al., 2018). In sum, if waitresses are in fact on average tipped better than waiters, our results coupled with those of previously published studies call into question the meaningfulness of the magnitude of this difference. We also estimated supplemental GLM models in an effort to replicate extant scholarship showing that customers' tip opposite sex servers better than same sex servers (Collin et al., 2003; Lynn et al., 2008; Lynn & McCall, 2016). Lending credence to these extant studies, a marginally significant (F = 3.66, p =.056) interaction effect between servers' sex and customers' sex was observed in Study 1. Estimated marginal means indicated that male respondents reported that they would tip a female (US$5.89) nominally more than they would tip a male (US$5.61) server. In contrast, female respondents reported that they would tip a male (US$6.00) more than a female (US$5.60) server. However, we were not able to replicate this pattern in Study 2 (F =.482, p =.488) or Study 3 (F = .165, p =.685).
17. As a percent of the bill, the unweighted and weighted meta-estimate of the server race effect in our studies was .002 (95% confidence interval [95% CI] = −.012, .007) and .001 (95% CI = −.005, .003), respectively.
18. For instance, 92% and 91% of the servers in Lynn and McCall's (2009) and Brewster's (2015) surveys of restaurant servers, respectively, were White. Of the 295 consumers in Parrett's (2011) research, only 15 were waited on by a non-White server. Finally, 74% of the 466 dining encounters analyzed by Lynn et al. (2012) included a White server. In short, the small number of Black servers that were directly or indirectly represented in these studies limit the inferences that can be drawn about customers' proclivities to tip Black servers, in particular, less than their White counterparts.
19. Controlling for servers' sex/race and customers' sex, Black respondents in Study 1 (b = −1.40, p < .000), Study 2 (b = −1.29, p < .000), and Study 3 (b = −1.72, p =.002) were observed to tip significantly less than White respondents.
20. If we assume that the results from these previous studies were valid reflections of the empirical reality then our experiments were more than sufficiently powered to detect such an effect. Of the two studies centered on restaurant servers, Brewster and Lynn (2014) observed a smaller effect, with a mean difference in tip percent of 1.50 percentage points between Black servers and White servers after controlling for service behaviors and service quality. If this is in fact a reasonable estimate of the effect size of servers' race on tipping then the statistical power would be 93%, 99%, and 76% for Studies 1, 2, and 3, respectively. For Study 2, in particular, the power indicates a vanishingly small likelihood that an actual effect equivalent to that observed by Brewster and Lynn (2014) would not have been observed if, in fact, it existed. If the actual effect was presumed to be larger, as suggested by Lynn et al.'s (2008) results, these power estimates would be even stronger. In short, the likelihood of a Type 1 error in the two earlier studies far exceeds the likelihood of three Type 2 errors in the current survey experiments. Our inability to replicate these earlier results underscores the various problems with underpowered studies including the heightened risk of overestimating effect size (error of magnitude, see Gelman & Carlin, 2014; Forstmeier & Schielzeth, 2011; Schimmack, 2020) and falsely rejecting null hypotheses. When this risk is coupled with publication biases (see Gerber & Malhotra, 2008), the outcome is that such results are often times not reproducible.
21. https://www.washingtonpost.com/food/2020/03/05/olive-garden-manager-black-server/
22. https://www.wave3.com/2018/09/29/racist-note-left-instead-tip-kentucky-applebees/
23. https://local12.com/news/local/black-nky-server-says-he-was-left-without-a-tip-all-lives-matter-message-instead-cincinnati
24. Black New York City strippers have alleged (#NYCStripperStrike) that they have a "harder time getting scheduled on nights with high earning potential and gaining access to VIP spaces." (see https://www.vice.com/en_us/article/bj7je8/meet-the-new-york-strippers-organizing-against-racism-in-the-industry).
25. https://www.washingtonpost.com/news/voraciously/wp/2020/09/30/parent-company-of-olive-garden-violates-the-civil-rights-act-with-its-tipping-policies-activists-say/?outputType=amp
26. For instance, it is estimated that 58% of tipped workers in full-service casual dining establishments (e.g., Applebee's, Olive Garden, Denny's, etc.) are White. In contrast, 78% of those working in fine dining establishments are White (see https://onefairwage.com/wp-content/uploads/2020/07/LegacyOfSlavery_4.pdf).

References

Alexander v. Casino Queen, Inc. 975, 739 F.3d972 (7th Cir. 2014).

Aruguete, M. S., Huynh, H., Browne, B. L., Jurs, B., Flint, E., & McCutcheon, L. E. (2019). How serious is the "carelessness" problem on Mechanical Turk? *International Journal of Social Research Methodology*, *22*(5), 441–449. https://doi.org/10.1080/13645579.2018.1563966

Ayres, I., Vars, F. E., & Zakariya, N. (2005). To insure prejudice: Racial disparities in taxicab tipping essay. *Yale Law Journal*, *114*(7), 1613–1674.

Banks, G. C., Woznyj, H. M., Kepes, S., Batchelor, J. H., & McDaniel, M. A. (2018). A meta-analytic review of tipping compensation practices: An agency theory perspective. *Personnel Psychology*, *71*(3), 457–478. https://doi.org/10.1111/peps.12261

Bendick, M., Jr., Rodriguez, R. E., & Jayaraman, S. (2010). Employment discrimination in upscale restaurants: Evidence from matched pair testing. *The Social Science Journal*, *47*(4), 802–818. https://doi.org/10.1016/j.soscij.2010.04.001

Berinsky, A. J., Huber, G. A., & Lenz, G. S. (2012). Evaluating online labor markets for experimental research: Amazon.com's Mechanical Turk. *Political Analysis*, *20*(3), 351–368. https://doi.org/10.1093/pan/mpr057

Bertrand, M., & Mullainathan, S. (2004). Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. *American Economic Review*, *94*(4), 991–1013. https://doi.org/10.1257/0002828042002561

Brauer, J. R., Day, J. C., & Hammond, B. M. (2019). Do employers "walk the talk" after all? An illustration of methods for assessing signals in underpowered designs. *Sociological Methods & Research*. Advance online publication. https://doi.org/10.1177/0049124119826158

Brewster, Z. W. (2015). Perceptions of intergroup tipping differences, discriminatory service, and tip earnings among restaurant servers. *International Journal of Hospitality Management*, *46*, 15–25. https://doi.org/10.1016/j.ijhm.2015.01.001

Brewster, Z. W., & Brauer, J. R. (2017). Tableside justice: Racial differences in retributive reactions to dissatisfaction. *Sociology of Race and Ethnicity*, *3*(3), 383–397.

Brewster, Z. W., & Lynn, M. (2014). Black–white earnings gap among restaurant servers: A replication, extension, and exploration of consumer racial discrimination in tipping. *Sociological Inquiry*, *84*(4), 545–569. https://doi.org/10.1111/soin.12056

Brewster, Z. W., & Nowak, G. R. (2019). Racial prejudices, racialized workplaces, and restaurant servers' hyperbolic perceptions of Black–White tipping differences. *Cornell Hospitality Quarterly*, *60*(2), 159–173.

Carman, T. (2020, September 30). Parent company of Olive Garden violates the Civil Rights Act with its tipping policies, activists say. *The Washington Post*. https://www.washingtonpost.com/news/voraciously/wp/2020/09/30/parent-company-of-olive-garden-violates-the-civil-rights-act-with-its-tipping-policies-activists-say/

Casler, K., Bickel, L., & Hackett, E. (2013). Separate but equal? A comparison of participants and data gathered via Amazon's MTurk, social media, and face-to-face behavioral testing. *Computers in Human Behavior*, *29*(6), 2156–2160.

Conlin, M., Lynn, M., & O'Donoghue, T. (2003). The norm of restaurant tipping. *Journal of Economic Behavior & Organization*, *52*(3), 297–321.

Forstmeier, W., & Schielzeth, H. (2011). Cryptic multiple hypotheses testing in linear models: Overestimated effect sizes and the winner's curse. *Behavioral Ecology and Sociobiology*, *65*, 47–55.

Gaddis, S. (2017). How Black are Lakisha and Jamal? Racial perceptions from names used in correspondence audit studies. *Sociological Science*, *4*, 469–489. https://doi.org/10.15195/v4.a19

Gatta, M. L. (2002). *Juggling food and feelings: Emotional balance in the workplace*. Lexington Books.

Gelman, A., & Carlin, J. (2014). Beyond power calculations: Assessing type S (Sign) and type M (Magnitude) errors. *Perspectives on Psychological Science*, *9*(6), 641–651.

Gerber, A. S., & Malhotra, N. (2008). Publication bias in empirical sociological research: Do arbitrary significance levels distort published results? *Sociological Methods & Research*, *37*(1), 3–30.

Hainmueller, J., Hangartner, D., & Yamamoto, T. (2015). Validating vignette and conjoint survey experiments against real-world behavior. *Proceedings of the National Academy of Sciences*, *112*(8), 2395–2400.

Hauser, D. J., & Schwarz, N. (2016). Attentive Turkers: MTurk participants perform better on online attention checks than do subject pool participants. *Behavior Research Methods*, *48*(1), 400–407. https://doi.org/10.3758/s13428-015-0578-z

Jayaraman, S. (2013). *Behind the kitchen door*. ILR Press.

Jayaraman, S. (2016). *Forked: A new standard for American dining*. Oxford University Press.

Kim, P. B., Nemeschansky, B., & Brandt, L. (2017). An exploratory study of determinants for restaurant servers' actual tip earnings: Individual characteristics and work conditions. *Journal of Foodservice Business Research*, *20*(1), 15–33.

Kline, J. (2016). Fifteen percent or less: A title VII Analysis of racial discrimination in restaurant tipping notes. *Iowa Law Review*, *101*(4), 1651–1680.

Lynn, M. (2006). Race differences in restaurant tipping. *Journal of Foodservice Business Research*, *9*(4), 99–113. https://doi.org/10.1300/j369v09n04_07

Lynn, M. (2017). Should US restaurants abandon tipping? A review of the issues and evidence. *Psychosociological Issues in Human Resource Management*, *5*(1), 120–159.

Lynn, M., Jabbour, P., & Kim, W. G. (2012). Who uses tips as a reward for service and when? An examination of potential moderators of the service–tipping relationship. *Journal of Economic Psychology*, *33*(1), 90–103. https://doi.org/10.1016/j.joep.2011.09.009

Lynn, M., Kwortnik, R. J., & Sturman, M. C. (2011). Voluntary tipping and the selective attraction and retention of service workers in the USA: An application of the ASA model. *The International Journal of Human Resource Management*, *22*(9), 1887–1901. https://doi.org/10.1080/09585192.2011.573968

Lynn, M., & McCall, M. (2009). Techniques for increasing servers' tips: How generalizable are they? *Cornell Hospitality Quarterly*, *50*(2), 198–208. https://doi.org/10.1177/1938965509334221

Lynn, M., & McCall, M. (2016). *Beyond gratitude and gratuity: A meta-analytic review of the predictors of restaurant tipping* [Electronic version]. http://scholarship.sha.cornell.edu/workingpapers/21

Lynn, M., & Simons, T. (2000). Predictors of male and female servers' average tip earnings. *Journal of Applied Social Psychology*, *30*(2), 241–252. https://doi.org/10.1111/j.1559-1816.2000.tb02314.x

Lynn, M., Sturman, M., Ganley, C., Adams, E., Douglas, M., & McNeil, J. (2008). Consumer racial discrimination in tipping: A replication and extension. *Journal of Applied Social*

Psychology, *38*(4), 1045–1060. https://doi.org/10.1111/j.1559-1816.2008.00338.x

Murphy, K. R., Herr, B. M., Lockhart, M. C., & Maguire, E. (1986). Evaluating the performance of paper people. *Journal of Applied Psychology*, *71*(4), 654–661.

Nelson, R. (2020). Substantive pay equality: Tips, commissions, and how to remedy the pay disparities they inflict. *SSRN Electronic Journal*. Advance online publication. https://doi.org/10.2139/ssrn.3583157

One Fair Wage. (2020). *A persistent legacy of slavery*. http://onefairwage.site/wp-content/uploads/2020/11/LegacyOfSlavery_5.pdf

Pager, D., Bonikowski, B., & Western, B. (2009). Discrimination in a low-wage labor market: A field experiment. *American Sociological Review*, *74*(5), 777–799. https://doi.org/10.1177/000312240907400505

Parrett, M. (2011). Customer discrimination in restaurants: Dining frequency matters. *Journal of Labor Research*, *32*(2), 87–112. https://doi.org/10.1007/s12122-011-9107-8

Press, A. N. (2020, October 2). It's time to end the subminimum wage for tipped workers. *Jacobin*. https://jacobinmag.com/2020/10/tipped-restaurant-workers-waiters-coronavirus-subminimum-wage

Rouse, S. V. (2015). A reliability analysis of Mechanical Turk data. *Computers in Human Behavior*, *43*, 304–307. https://doi.org/10.1016/j.chb.2014.11.004

Schimmack, U. (2020, December 17). *A quick fix for the replication crisis in psychology*. Replicability-Index. https://replicationindex.com/2020/12/17/a-quick-fix-for-the-replication-crisis-in-psychology/

Seiter, J. S., & Weger, H., Jr. (2010). The effect of generalized compliments, sex of server, and size of dining party on tipping behavior in restaurants. *Journal of Applied Social Psychology*, *40*(1), 1–12. https://doi.org/10.1111/j.1559-1816.2009.00560.x

Smith, S. M., Roster, C. A., Golden, L. L., & Albaum, G. S. (2016). A multi-group analysis of online survey respondent data quality: Comparing a regular USA consumer panel to MTurk samples. *Journal of Business Research*, *69*(8), 3139–3148. https://doi.org/10.1016/j.jbusres.2015.12.002

Wang, L.-i. (2014). At the tipping point: Race and gender discrimination in a common economic transaction. *Virginia Journal of Social Policy & the Law*, *21*(1), 101–166.

Wilson, E. R. (2019). Tip work: Examining the relational dynamics of tipping beyond the service counter. *Symbolic Interaction*, *42*(4), 669–690. https://doi.org/10.1002/symb.413

Author Biographies

Zachary W. Brewster is an Associate Professor in the Department of Sociology at Wayne State University.

Kenneth Gourlay is an MA student in the Department of Sociology at Wayne State University.

Gerald Roman Nowak III is a PhD student in the Department of Sociology at Michigan State University.