Article

Social Psychological and Personality Science
2020, Vol. 11(6) 761-769

DOI: 10.1177/1948550619876638
journals.sagepub.com/home/spp
SAGE

Racial Biases in Officers' Decisions to Frisk Are Amplified for Black People Stopped Among Groups Leading to Similar Biases in Searches, Arrests, and Use of Force

Erin Cooley[1], Neil Hester[2], William Cipolli[1], Laura I. Rivera[3], Kaitlin Abrams[1], Jeremy Pagan[4], Samuel R. Sommers[4], and Keith Payne[5]

Abstract

Violent encounters between police and Black people have spurred debates about how race affects officer decision-making. We propose that racial disparities in police–civilian interactions are amplified when police interact with Black civilians who are encountered in groups. To test this possibility, we analyzed New York City stop and frisk data for over 2 million police stops. Results revealed that Black (vs. White) people were more likely to be frisked, searched, arrested, and have force used against them. Critically, these racial disparities were more pronounced for people stopped in groups (vs. alone): Being stopped in a group led to a 1.7% increase in racial disparities for frisks, a 1% increase for searches, a 0.3% increase for arrests, and a 1.7% increase for use of force. Moreover, these disparities held even when we controlled for a potential proxy of effective policing: discovery of illegal contraband. We conclude that groups amplify racial disparities in policing.

Keywords

prejudice/stereotyping, psychology and law, group processes

Numerous high-profile, fatal interactions between Black civilians and law enforcement have led to increased dialogue about the treatment of Black people by police (Garza, 2014; Sawyer & Gampa, 2018). Part of why these conversations are difficult is that we cannot know with certainty the degree to which racial biases influence any given officer–civilian encounter. However, analyses of aggregated data on police–civilian interactions suggest that, on average, racial disparities in policing occur. For example, Black and Hispanic/Latinx people are more likely to be stopped, frisked, and have force used against them than White people (Ferrandino, 2015; Gelman, Fagan, & Kiss, 2007; Goel, Raul, & Shroff, 2016), especially under suspicion of possessing illegal contraband (Ridgeway, 2007). At the same time, stops of Black and Hispanic/Latinx people are less likely than stops of White people to result in discovery of these incriminating items (Ayres & Borowsky, 2008; Dunn & Shames, 2019). Together, these patterns suggest that the observed racial disparities may be driven by racial biases rather than effective or accurate policing (Pierson et al., 2017).

Although analyses of field data cannot easily isolate the mechanisms behind differential treatment of Black people by police, experimental data suggest that racial stereotypes play an important role. These data suggest that there are strong cultural stereotypes that Black people are threatening and aggressive (Devine & Elliot, 1995; Eberhardt, Davies, Purdie-Vaughns, & Johnson, 2006; Richeson, Todd, Trawalter, & Baird, 2008; Sommers & Marotta, 2014) and that these stereotypes predict a greater likelihood of "shooting" unarmed Black (vs. White) people in laboratory tasks—at least among untrained civilians (Correll, Park, Judd, & Wittenbrink, 2002; Correll et al., 2007; Payne, 2001). Among sworn police officers, these racial biases in shooting decisions are attenuated, suggesting that training may help break the link between racial stereotypes and decisions to shoot (Correll et al., 2007). Likewise, recent data suggest that prior information, such as that provided by dispatch (e.g., reports that a suspect has a real gun,

[1] Colgate University, Hamilton, NY, USA
[2] McGill University, Montreal, Quebec, Canada
[3] University of Wisconsin–Madison, WI, USA
[4] Tufts University, Medford, MA, USA
[5] University of North Carolina at Chapel Hill, NC, USA

Corresponding Authors:
Erin Cooley, Colgate University, Hamilton, NY, USA; Neil Hester, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA.
Emails: ecooley@colgate.edu; nehester@live.unc.edu

when it is a toy), can also affect officers' decisions to shoot—raising the possibility that the ways that witnesses report a situation to dispatch (which may be mired with their own racial biases) have consequences for how officers respond (Johnson, Cesario, & Pleskac, 2018). Finally, individual attributes of Black people can also influence their experience of racial biases throughout the criminal justice system. Black individuals who are more phenotypically "Black" (Blair, Judd, & Chapleau, 2004; Levinson, Cai, & Young, 2010) or physically tall (Hester & Gray, 2018) are perceived as more threatening and, thus, tend to experience harsher treatment (e.g., more police use of force, longer sentencing). Together, these findings help explain the conditions under which Black individuals are most likely to experience disparate treatment when interacting with police.

Although isolating responses to Black versus White individuals is good practice in the lab (e.g., experimental control), in the real world, police often interact with *groups* of people, stopping and questioning multiple individuals at once. And, a variety of theoretical work suggests that the presence of groups may amplify racial biases (Cooley & Payne, 2017, 2018; Hamilton, Chen, Ko, Winczewski, Banerji, & Thurston, 2015). Thus, might Black people who are stopped among groups be particularly likely to experience racially disparate policing?

Race, Groups, and Threat

It is well known that groups are central to the *development* of racial biases (Allport, 1954; Hamilton et al., 2015). However, more recent research indicates that groups are also integral to the *application* of racial stereotypes. For example, Black groups are stereotyped as more "aggressive" and "untrustworthy" than Black individuals, and these effects emerge on both implicit (Cooley & Payne, 2017) and explicit measures (Cooley & Payne, 2018). Likewise, Asian groups are stereotyped as more "hardworking" than Asian individuals (Cooley & Payne, 2018). Such findings suggest that stereotypes about entire racial categories are applied more strongly to homogenous racial groups than individuals from those groups—perhaps because racial categories are themselves large groups.

Similarly, other research indicates that people perceive groups, regardless of their racial composition, as more threatening than individuals. For example, decades of data on the "interindividual intergroup discontinuity" suggest that the mere observance of groups, holding race constant, can lead to greater expectations of aggressive and untrustworthy behavior (Insko et al., 1987, 1998; Insko & Schopler, 1998; Wildschut, Pinter, Vevea, Insko, & Schopler, 2003). Similarly, evolutionary theory proposes that people experience "coalitional threat" such that an out-group member in a group is perceived as more threatening than any given individual from that group (Boyer, Firat, & van Leeuwen, 2015; Kurzban & Neuberg, 2005).

Building from these theories, we propose that two processes may lead police to perceive Black people in groups as particularly threatening. First, perception of Black people in groups of other Black people may amplify the activation of stereotypes and prejudice associated with Black people—including stereotypes of threat/aggression (Cooley & Payne, 2017, 2018). Second, there may be an interactive effect of threat stereotypes associated with Blackness (Devine & Elliot, 1995) and threat stereotypes associated with groups (Boyer et al., 2015; Wildschut et al., 2003) that leads Black people in groups of any race to seem particularly threatening. Together, these perspectives suggest that Black people stopped while in groups—regardless of the racial composition of those groups—may be particularly likely to elicit racial biases in officer decision-making.

In the present article, we analyze New York City Police Department stop and frisk data to test whether stops of Black people in groups are more likely to escalate to frisks, searches, arrests, and use of force than stops of Black individuals, White people in groups, or White individuals. Although a variety of work has documented overall disparities in how Black versus White people are treated under stop and frisk policy (Dunn & Shames, 2019; Ridgeway, 2007), this existing work has not examined whether being stopped in a group amplifies these disparities. Of note, groups (vs. individuals) may objectively present a greater threat to a single officer, providing a logical reason for escalation. However, such reasoning would not explain why Black people in groups would experience particularly extreme escalation as compared to White people in groups—unless Black people in groups are more likely to be engaged in illegal activity. To this point, we further test whether greater escalation of stops of Black people in groups is accounted for by evidence of effective policing (i.e., discovery of illegal contraband). Such findings would have important implications for enacting social policy, understanding structural racism, and improving officer training.

Method

New York Stop and Frisk Data

To test our hypotheses, we analyzed data from New York Police Department stop and frisk records from 2006 to 2012 (see https://on.nyc.gov/2B3r0t1). These data reflect stops of pedestrians, not traffic stops, and are traditionally called "stop and frisk" records even though stops do not always lead to frisking. During the analyzed time period, decisions to stop a suspect were considered legally justified when there was reasonable suspicion that a crime had been, was being, or was about to be committed (see Ridgeway, 2007). After a stop was made, frisking (i.e., pat down of the suspect's clothing for weapons) could follow if the officer believed that the suspect might be armed and dangerous. A frisk sometimes led to further outcomes including searching (i.e., which is a step beyond the pat down for weapons that frisking entails), arrests, and use of force (Busby, 2009). Our focus for the present analyses was not on the probability of the initial stop but rather on the escalation after the stop had occurred. We were interested in whether police treat Black people stopped in groups (vs. individually)

more severely—a prediction that is consistent with known links between being Black, being in a group, and perceived threat.

Although the data in this database were recorded from 2003 to 2017, we began analyses using data from 2006 because the reporting protocol changed in 2006 to include police sector (a key variable in our analyses). Likewise, we stopped compiling data after 2012 given that, in 2013, Judge Shira Scheindlin ruled that the stop and frisk policies of the New York Police Department violated the rights of minorities and allowed racial profiling (Goldstein, 2013). Thus, we reasoned that the nature of the stop and frisk program likely changed at that time. Although beyond the scope of the present work, future research could conduct a longitudinal analysis to examine how the racial disparities we test for here may (or may not) have changed since that ruling.

Finally, we focused on stops of non-Hispanic White and Black people so that we could examine racial disparities experienced by Black (vs. White) people, independent of disparities that may be experienced by Hispanic/Latinx people. Although much previous work has focused on threat stereotypes of Black men, Black women are also subject to threat stereotyping (Cooley, Winslow, Vojt, Shein, & Ho, 2018; Donovan & West, 2015; Thiem, Neel, Simpson, & Todd, 2019). Thus, we chose to retain stops of both male and female civilians. These restrictions left 2,432,105 stops (92.08% male, 7.38% female, and 0.54% other).

Summary of Analytic Plan

Because our analyses focused on real-world data rather than laboratory data, we considered strategies for addressing concerns about alternate explanatory variables that may covary with our constructs of theoretical interest. Many of these key alternate variables are geographically bound, such as land values, crime rates, and racial demographics. Fortunately, police jurisdictions are also divided geographically into "precincts" (e.g., the first precinct in Manhattan covers The World Trade Center, SoHo, Tribeca, and Wall Street), and each precinct is further divided into multiple sectors that are designed to correspond, as best as possible, to neighborhood boundaries (see https://www1.nyc.gov/site/nypd/bureaus/patrol/neighborhood-coordination-officers.page). Officers are then assigned to particular sectors so that they can develop familiarity with specific locations. Thus, one key tactic we used to address alternative explanations (e.g., regional crime rates) was to employ a mixed modeling approach, nesting all stops within police sectors (there were 849 sectors), nesting all sectors within police precincts (there were 76 precincts), and including a random intercept for sector (Gelman et al., 2007). Critically, nesting by sector allowed us to account for the fact that base rates of crime vary by location (Cesario, Johnson, & Terrill, 2019). And, because New York City is divided into over 800 different sectors in the stop and frisk data set, this provided geographic divisions that were fine-grained enough that they could accurately model a variety of geographically bound variables of interest in addition to crime rates, such as land value and racial demographics (Gelman et al., 2007). Furthermore, the high number of different sectors (much higher than the number of precincts) ensured wide variation on these dimensions, preventing range restriction.

In addition to a mixed-modeling approach, we included two control variables to further account for regional variation in crime base rates. First, we included officers' perception that the area of a stop was high-crime (recorded as "yes/no" for each individual stopped). Second, we included precinct-level felony rates. These rates were calculated by averaging felony counts per precinct across years 2006–2012 (source: https://www1.nyc.gov/site/nypd/stats/crime-statistics/historical.page) and then dividing those counts by precinct-level populations as indexed in the 2010 census. Because the census breaks down populations by neighborhood, rather than police precinct, population estimates for precincts were interpolated by mapping neighborhood data, as closely as possible, onto precinct coordinates (see http://johnkeefe.net/nyc-police-precinct-and-census-data). Ideally, we would have been able to further identify *race-specific* crime rates by precinct (see Cesario et al., 2019) but were not able to do so given the aggregated data available to us.

Finally, we conducted analyses to assess whether disparate policing of Black people stopped in groups was driven by accurate policing. To test this, we examined whether the greater likelihood of frisking, searching, arresting, and using force against Black people stopped in groups was driven by a greater likelihood of discovery of illegal contraband (i.e., weapons, drugs, or stolen property; Ridgeway, 2007) among Black people in groups (Ayres & Borowsky, 2008; Pierson et al., 2017).

Limitations

Before providing our results, we describe some limiting aspects of the data. Importantly, data in the stop and frisk data set are recorded at the level of each individual stopped. This means that if a given individual is recorded as stopped in a group, we cannot with certainty match that individual to other people who may have been stopped as a part of the same group. Thus, there is dependency in the data that we cannot accurately model. This limitation also means that we cannot know the racial composition nor size of the group in which an individual was stopped, nor can we know whether some or all of the people in that group were stopped or frisked by the police. As a result, even if data support our general hypothesis that Black people stopped in groups (vs. alone) will experience greater racial disparities in police treatment, we cannot conclusively know the mechanism.

However, as we described previously, multiple theories converge on the prediction that Black people in groups, regardless of the group's racial composition, will be perceived as particularly threatening and, thus, experience enhanced racial biases in policing as compared to Black individuals (Cooley & Payne, 2017, 2018; Wildschut et al., 2003). Furthermore, we expect that most of the stopped groups are either all/mostly White or all/mostly Black, given that work on racial homophily suggests that the composition of Americans' friend groups is more strongly determined by race than by any other social category

Table 1. Data Sample Illustrating the Structure of the Data.

Precinct	Sector	Race (Black)	Othpers (Yes)	High Crime (Yes)	Precinct Felony Rate	Frisked	Searched	Arrested	Force Used
1	1	1	1	1	.025	1	0	0	0
1	1	1	0	1	.025	1	1	1	0
1	2	0	0	1	.025	1	1	0	1
1	3	0	1	0	.025	0	0	0	0
5	9	1	0	0	.019	0	0	0	0
5	11	1	1	1	.019	1	0	0	0
5	11	1	1	1	.019	1	1	1	1
6	17	1	0	0	.027	1	0	0	0

Note. Stops are nested within sector; sectors are nested within precinct. Outcomes are binary.

(McPherson, Smith-Lovin, & Cook, 2001, pp. 420–422). Likewise, levels of racial segregation in NYC are quite high (Flores & Lobo, 2013; Lichter, Parisi, Taquino, 2015). This means that, although the unknown racial composition of groups might introduce some statistical noise, any patterns we do find likely reflect judgments of mostly homogeneous groups. Overall, despite these limitations, we reason that the present analyses fill a gap in the literature by extending existing theory and experiments into an important real-world setting.

Table 2. Black People Stopped In Groups Had the Highest Probability of Being Frisked.

Race	Group	Probability	*SE*	Lower	Upper
W	N	.437	.00792	.421	.452
B	N	.543	.00801	.527	.559
W	Y	.486	.00817	.470	.502
B	Y	.609	.00772	.594	.624

Note. W = White; B = Black; N = No; Y = Yes.

Results

To test our hypotheses, we used multilevel models that included race, group, and the key Race × Group interaction as predictors. Within each model, stops were nested within police sector, and sector was nested within police precinct, with a random intercept for sector.[1] To further account for variability in crime rates across stop locations, we included officers' perception that the area of the stop was high-crime and precinct-level felony rates as control variables. Table 1 provides clarification about the structure of our data.

The syntax for the basic multilevel model (without control variables) is as follows:

$$mod\_frisked\ 1 \leftarrow glmer(frisked \sim (1|pct/\text{sector}) + race \times othpers,\ + family = binomial(link = \text{“log}\,it\text{”}),\ data = data\_wb)$$

“*Mod_frisked1*” is the model, *glmer* is the function, *frisked* is the outcome, 1 is the fixed slope, *pct/sector* is the nesting structure, *race* × *othpers* is a factorial representation of the key predictors, *family* = *binomial* is the defined structure of the outcome, *link* = “*logit*” is the link function that allows for a linear model to provide a binomial outcome, and *data* = *data_wb* refers to a subset of our data that only includes White and Black stops. Supplementary Materials (accessed here: https://osf.io/dvpuk/?view_only=19a7548af29e426297b090d89a343798) provide the full set of R code, full model details, and the comparable results of several alternative models.

Next, we report the results separately for each of the following outcomes: probability of frisking, searching, arresting, and use of force. The first three outcomes are all directly provided as binary variables in the data set. The police force outcome, in contrast, documented whether force was used and, if so, which of nine different kinds of force (i.e., hands, suspect on ground, suspect against wall, weapon drawn, weapon pointed, baton, handcuffs, pepper spray, and other; see Supplementary Materials for descriptive statistics). Because we could not objectively rank these nine types of force to accurately reflect the severity of force used, we chose to treat “use of force” as a binary outcome to match the coding of decisions to frisk, search, and arrest. Thus, if any of the nine types of force were marked “yes” for a particular individual who was stopped, we then coded that individual as having force used against them. Log odds are reported as probabilities to improve interpretability.

Frisks

When predicting frisks, the model revealed an overall effect of race ($z = 89.59$, $p < .001$), such that the probability that Black people were frisked was .116 higher than the probability that White people were frisked, 95% confidence interval (CI) = [.113, .118]. There was also an overall effect of group ($z = 55.06$, $p < .001$), such that the probability that people stopped in groups were frisked was .059 higher than the probability that people stopped individually were frisked, 95% CI [.057, .061]. Finally, the model revealed the predicted Race × Group interaction ($z = 8.90$, $p < .001$), such that the effect of being Black (vs. White) on the probability of being frisked was larger for groups (.1238, $p < .001$) than for individuals (.1066, $p < .001$) by .0172, 95% CI [.0123, .0221]. See Table 2 for probabilities by condition. To summarize, being stopped in a group led to a 1.7% increase in racial disparities for frisks.

Table 3. Black People Stopped in Groups Had the Highest Probability of Being Searched.

Race	Group	Probability	*SE*	Lower	Upper
W	N	.0839	.00251	.0792	.0890
B	N	.0865	.00258	.0816	.0917
W	Y	.0838	.00265	.0788	.0892
B	Y	.0965	.00288	.0910	.1023

Note. W = White; B = Black; N = No; Y = Yes.

Table 4. Black People Stopped In Groups Had the Highest Probability of Being Arrested.

Race	Group	Probability	*SE*	Lower	Upper
W	N	.0589	.00214	.0548	.0632
B	N	.0620	.00224	.0577	.0665
W	Y	.0607	.00232	.0563	.0654
B	Y	.0666	.00242	.0620	.0715

Note. W = White; B = Black; N = No; Y = Yes.

Searches

When predicting searches, the model revealed an overall effect of race ($z = 10.39$, $p < .001$), such that the probability that Black people were searched was .008 higher than the probability that White people were searched, 95% CI [.006, .009]. There was also an overall effect of group ($z = 7.67$, $p < .001$), such that the probability that people stopped in groups were searched was .005 higher than people stopped individually, 95% CI [.0035, .006]. Finally, the model revealed the predicted Race × Group interaction ($z = 8.43$, $p < .001$), such that the effect of being Black (vs. White) on the probability of being searched was larger for groups (.0127, $p < .001$) than for individuals (.0026, $p < .001$) by .0101, 95% CI [.0074, .0128]. See Table 3 for probabilities by condition. To summarize, being stopped in a group led to a 1% increase in racial disparities for searches.

Arrests

When predicting arrests, the model revealed an overall effect of race ($z = 7.43$, $p < .001$), such that the probability that Black people were arrested was .0045 higher than the probability that White people were arrested, 95% CI [.0033, .0056]. There was also an overall effect of group ($z = 6.09$, $p < .001$), such that the probability that people stopped in groups were arrested was .0032 higher than people stopped individually, 95% CI [.0021, .0042]. Finally, the model revealed the predicted Race × Group interaction ($z = 2.64$, $p = .008$), such that the effect of being Black (vs. White) on the probability of arrest was larger for groups (.0059, $p < .001$) than for individuals (.0031, $p < .001$) by .0027, 95% CI [.0005, .0049]. See Table 4 for probabilities by condition. To summarize, being stopped in a group led to a 0.3% increase in racial disparities for arrests.

Table 5. Black People Stopped In Groups Had the Highest Probability of Having Force Used Against Them.

Race	Group	Probability	*SE*	Lower	Upper
W	N	.168	.00570	.157	.179
B	N	.203	.00664	.190	.216
W	Y	.216	.00710	.203	.230
B	Y	.268	.00806	.252	.284

Note. W = White; B = Black; N = No; Y = Yes.

Table 6. Summary of Findings, Including Absolute and Relative Racial Differences.

	Frisks	Searches	Arrests	Uses of Force
Base rates	53.1	8.3	5.8	21.4
Percentage values				
White, individual stops	43.7	8.4	5.9	16.8
Black, individual stops	54.3	8.7	6.2	20.2
White, group stops	48.6	8.4	6.1	21.6
Black, group stops	60.9	9.7	6.7	26.8
Absolute differences (percent)				
Black–White percentage, individual stops	10.7	0.3	0.3	3.4
Black–White percentage, group stops	12.4	1.3	0.6	5.1
Relative differences (ratio)				
Black-to-White ratio, individual stops	1.24	1.03	1.05	1.20
Black-to-White ratio, group stops	1.25	1.15	1.10	1.24

Use of Force

When predicting use of force, the model revealed an overall effect of race ($z = 29.26$, $p < .001$), such that the probability that Black people had force used against them was .0426 higher than the probability that White people had force used against them, 95% CI [.0398, .0455]. There was also an overall effect of group ($z = 34.23$, $p < .001$), such that the probability that people stopped in groups experienced use of force was .0564 higher than people stopped individually, 95% CI [.0532, .0596]. Finally, the model revealed the predicted Race × Group interaction ($z = 5.14$, $p < .001$), such that the effect of being Black (vs. White) on use of force was larger for groups (.0515, $p < .001$) than for individuals (.0346, $p < .001$) by .0169, 95% CI [.0130, .0208]. See Table 5 for probabilities by condition. To summarize, being stopped in a group led to a 1.7% increase in racial disparities for use of force.

Overall, these results suggest that racial disparities present in the NYPD's stop and frisk program are especially large for Black people stopped in groups as compared to Black people stopped alone. Moreover, these disparities replicated across multiple consequential outcomes from frisks to use of force and persisted even when controlling for two indices of overall crime: officer perceptions and precinct-level felony rates. Table 6 provides a summary of findings as they relate to base rates.

Table 7. Percentage of Stops Resulting in Discovery by Race and Group Status.

Outcome	Black	Black Individual	Black Group	White	White Individual	White Group
Percentage of stops resulting in contraband discovery	1.72	1.68	1.94	2.23	2.02	2.78

Table 8. Racial Disparities for Black People In Groups When Controlling for Discovery as Compared to Values When Not Controlling for Discovery.

Control Variable	Outcome	Probability (BG-WG) − (BI-WI)	*p* Value	Lower	Upper
w/"Discovery"	Frisks	.0076	.0001	.0033	.0120
	Searches	.0400	<.0001	.0326	.0474
	Arrests	.0355	<.0001	.0263	.0448
	Use of force	.0207	<.0001	.0151	.0262
w/o "Discovery"	Frisks	.0172	<.0001	.0123	.0221
	Searches	.0101	<.0001	.0074	.0128
	Arrests	.0027	<.0106	.0005	.0049
	Use of force	.0169	<.0001	.0130	.0208

Does Attention Toward Black Groups Reflect Effective Policing?

The findings above do not inherently reflect police bias; they might also reflect effective policing. For example, if Black people in groups are more likely to be engaged in drug use/dealing or violence, then they may experience more frisks, searches, arrests, and use of force—but because of officers' effective use of "reasonable suspicion" under the stop and frisk policy, rather than racial bias (Ridgeway, 2007). To test this possibility using the data available to us, we quantified effective policing as whether the stop ended in discovery of illegal contraband. We reasoned that discovery of such items may provide evidence of effective policing. By adding this variable as an additional predictor in our analyses, we could test whether enhanced racial disparities for Black people stopped in groups continued to emerge above and beyond accurate detection of threat. If so, this would suggest that the disparities experienced by Black people stopped in groups cannot be fully explained by effective policing.

Controlling for discovery. We reran the four models reported above, but, this time, we added "discovery" as an additional predictor (see Table 7). Discovery was coded as 1 for a given stop if illegal contraband was discovered (and 0 if not). Results of these models with discovery added remained consistent with our hypotheses (see Supplemental Materials for full model results). First, we will note that "discovery" was positively associated with all four outcomes: decisions to frisk ($z = 127.91$, $p < .001$), search ($z = 309.15$, $p < .001$), arrest ($z = 355.51$, $p < .001$), and use force ($z = 142.28$, $p < .001$).

Even with "discovery" included in the model, the predicted Race × Group interaction remained statistically significant for all outcomes: decisions to frisk ($z = 9.82$, $p < .001$), search ($z = 12.23$, $p < .001$), arrest ($z = 8.79$, $p < .001$), and use force ($z = 6.02$, $p < .001$). And, for each of these outcomes, the Race × Group interaction reflected that racial disparities were more pronounced for Black people stopped in groups than Black people stopped individually (see Table 8). In fact, the increase in racial disparities when traveling in groups was more pronounced, rather than less pronounced, for all later stages of escalation (i.e., searches, arrests, and use of force) when we added discovery as an additional control.

Together, these results suggest that the disproportionate escalation of stops of Black people in groups are not accounted for by accurate/effective policing—at least in terms of discovery of illegal contraband. Instead, results are consistent with our hypothesis that racial biases—in particular the stereotype that Black people, especially when in groups, pose a societal threat—may drive the observed disparities in escalation of stops from frisks to use of force.

Discussion

Several theories of group processes predict that Black people will experience increased prejudice and discrimination when in groups (Boyer et al., 2015; Cooley & Payne, 2017, 2018; Wildschut et al., 2003). Here, we provide further evidence for this prediction using an ecologically valid and important outcome—escalation of police–civilian interactions. Across over two million police stops, being in a group amplified racial disparities in police decisions to frisk, search, arrest, and use force against Black people. Such findings not only develop prior theory in an externally valid way but also highlight how decisions early in the stop and frisk process (disproportionate frisking experienced by Black people stopped in groups) may lead to disparate treatment throughout the encounter.

When considering the practical importance of these findings, the effect sizes of the Race × Group interactions were relatively small. However, these effects are stably estimated due to the large sample size, consistent with theory, and impactful when observed in a large-scale setting. We should also note that

the tendency for groups to amplify racial biases is smaller for searches and arrests than for frisks and uses of force, at least in terms of *absolute* differences. However, these differences are small partly because these outcomes are very uncommon in the first place. Although percentages reasonably approximate a linear function for middling values (i.e., between 20% and 80%), they are nonlinear at very high and very low values and require conversion to odds or ratios for substantive interpretation (e.g., a jump from 1% to 2% represents doubled odds; a jump from 50% to 51% represents a small increase in odds). When considered as ratios, group differences for these outcomes are more substantial, for example, 1.03 Black people are searched per White person for individual stops, but 1.15 Black people are searched per White person for group stops. The effect of group on relative differences in frisks, on the other hand, is actually quite small; but, because frisks are such a common outcome, the practical effect of this disparity is still meaningful.

Although the current data cannot conclusively address the possibility that the observed effects reflect effective policing, we find it compelling that even given the greater likelihood of Black people in groups being frisked and searched, discovery of illegal contraband does not explain these effects. In fact, if we use discovery of contraband as the criterion for effective policing, stops of Black civilians were less effective than were stops of White civilians—consistent with other research (Ayres & Borowsky, 2008; Pierson et al., 2017). Likewise, when we include discovery as an additional predictor in our models, larger racial disparities when traveling in a group versus alone only become more pronounced—at least at later stages in the escalation process (i.e., searches, arrests, and use of force). This suggests disparate police treatment of Black people stopped in groups may be driven by societal racial biases rather than Black groups being more likely to be engaged in illegal activity.

Critically, police are trained to diagnose levels of situational threat. When there is reason to believe that one's life is in danger, greater levels of force become legally justified. Although it may be reasonable/prudent for police to perceive greater threat in the presence of a crowd versus an individual, the present findings indicate that the greater threat associated with a group may be conflated with Blackness. Black people stopped in groups experienced uniquely harsh racial disparities as compared to Black people stopped individually. Thus, when training officers against the very real threat of interacting with groups, it may be helpful to simultaneously discuss the possibility that the presence of groups may make some people seem more threatening than others.

The nature of real-world data means that we cannot rule out all possible confounding variables. Likewise, the current data do not account for the number nor the race of group members other than the suspect, limiting our ability to conclusively identify the mechanism behind these effects. For example, one possibility is that differential treatment of Black people in groups is driven by groups amplifying the application of racial stereotypes (e.g., that Black = threat; Devine & Elliot, 1995). Such an explanation would be particularly compelling if we could assume that stops involving groups involved racially homogeneous groups (Cooley & Payne, 2017, 2018)—an assumption that may be reasonable given high levels of racial segregation in NYC (Flores & Lobo, 2013; Lichter et al., 2015) and high racial homophily in the United States (e.g., McPherson et al., 2001). However, if it is not reasonable to assume that people stopped in groups tend to be in racially homogeneous groups, then the observed findings may instead be the result of a compounding effect of the threat associated with groups (Wildschut et al., 2003) and the threat associated with Blackness (Devine & Elliot, 1995). Our suspicion is that both processes are likely at play in driving enhanced racial disparities for Black groups.

Conclusion

Overall, the present findings suggest that encountering police while in groups is associated with unique risks for Black people. Group entities face stronger stereotypes than individuals (Cooley & Payne, 2017, 2018; Hamilton et al., 2015) and both groups (Wildschut et al., 2003) and Black people (Devine & Elliot, 1995) are perceived as threatening. Here, we bolster these experimental findings by highlighting their real-world consequences: Black people in groups are disproportionately frisked by police, leading to racial disparities in more severe outcomes including use of force. Traveling in a group might offer a sense of security and protection; however, for Black people, traveling in groups may also bear hidden costs when it comes to interactions with police.

Authors' Notes

Erin Cooley and Neil Hester contributed equally to this work. The analyses reported in this manuscript were conducted with the help of the Data Science Collaboratory at Colgate University.

Declaration of Conflicting Interests

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

Funding

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

ORCID iD

Neil Hester https://orcid.org/0000-0003-3025-766X

Note

1. We did not include random slopes for statistical reasons, see Supplemental Materials.

References

Allport, G. (1954). *Handbook of social psychology*. New York, NY: Addison-Wesley.

Ayres, I., & Borowsky, J. (2008). *A study of racially disparate outcomes in the Los Angeles police department*. Los Angeles: ACLU of Southern California.

Blair, I. V., Judd, C. M., & Chapleau, K. M. (2004). The influence of Afrocentric facial features in criminal sentencing. *Psychological Science*, *15*, 674–679. doi:10.1037/0022-3514.87.6.763

Boyer, P., Firat, R., & van Leeuwen, F. (2015). Safety, threat, and stress in intergroup relations: A coalitional index model. *Perspectives on Psychological Science*, *10*, 434–450. doi:10.1177/1745691615583133

Busby, J. (2009, September 17). *Stop-and-frisk*. Retrieved July 12, 2018, from https://www.law.cornell.edu/wex/stop_and_frisk

Cesario, J., Johnson, D. J., & Terrill, W. (2019). Is there evidence of racial disparity in police use of deadly force? Analyses of officer-involved fatal shootings in 2015–2016. *Social Psychological and Personality Science*, *10*, 586–595. doi:10.1177/1948550618775108

Cooley, E., & Payne, B. K. (2017). Using groups to measure intergroup prejudice. *Personality and Social Psychology Bulletin*, *43*, 46–59. doi:10.1177/0146167216675331

Cooley, E., & Payne, B. K. (2018). A group is more than the average of its parts: Why existing stereotypes are applied more to the same individuals when viewed in groups than when viewed alone. *Group Processes and Intergroup Relations*, 1–15. doi:10.1177/1368430218756494

Cooley, E., Winslow, H., Vojt, A., Shein, J., & Ho, J. (2018). Bias at the intersection of identity: Conflicting social stereotypes of gender and race augment the perceived femininity and interpersonal warmth of smiling black women. *Journal of Experimental Social Psychology*, *74*, 43–49. doi:10.1016/j.jesp.2017.08.007

Correll, J., Park, B., Judd, C. M., & Wittenbrink, B. (2002). The police officer's dilemma: Using ethnicity to disambiguate potentially threatening individuals. *Journal of Personality and Social Psychology*, *83*, 1314–1329. doi:10.1037/0022-3514.83.6.1314

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*, 1006. doi:10.1037/0022-3514.92.6.1006

Devine, P. G., & Elliot, A. J. (1995). Are racial stereotypes really fading? The Princeton trilogy revisited. *Personality and Social Psychology Bulletin*, *21*, 1139–1150. doi:10.1177/01461672952111002

Donovan, R. A., & West, L. M. (2015). Stress and mental health: Moderating role of the strong Black woman stereotype. *Journal of Black Psychology*, *41*, 384–396. doi:10.1177/0095798414543014

Dunn, C., & Shames, M. (2019). *Stop-and-frisk in the de blasio era*. Retrieved April, 16, 2019, from https://www.nyclu.org/sites/default/files/field_documents/20190314_nyclu_stopfrisk_singles.pdf

Eberhardt, J. L., Davies, P. G., Purdie-Vaughns, V. J., & Johnson, S. L. (2006). Looking deathworthy: Perceived stereotypicality of black defendants predicts capital-sentencing outcomes. *Psychological Science*, *17*, 383–386. doi:10.1111/j.1467-9280.2006.01716.x

Ferrandino, J. (2015). Minority threat hypothesis and NYPD stop-and-frisk policy. *Criminal Justice Review*, *40*, 209–229. doi:10.1177/0734016814564989

Flores, R. J., & Lobo, A. P. (2013). The reassertion of a Black/non-Black color line: The rise in integrated neighborhoods without Blacks in New York City, 1970–2010. *Journal of Urban Affairs*, *35*, 255–282. doi:10.1111/j.1467-9906.2012.00646.x

Garza, A. (2014, October 7). *A herstory of the #blacklivesmatter movement*. Retrieved June 5, 2018, from http://www.thefeministwire.com/2014/10/blacklivesmatter-2/

Gelman, A., Fagan, J., & Kiss, A. (2007). An analysis of the New York City police department's "stop-and-frisk" policy in the context of claims of racial bias. *Journal of the American Statistical Association*, *102*, 813–823. doi:10.1198/016214506000001040

Goel, S., Rao, J. M., & Shroff, R. (2016). Precinct or prejudice? Understanding racial disparities in New York City's stop-and-frisk policy. *The Annals of Applied Statistics*, *10*, 365–394. doi:10.2139/ssrn.2572718

Goel, S., Rao, J. M., & Shroff, R. (2016). Precinct or prejudice? Understanding racial disparities in New York City's stop-and-frisk policy. *The Annals of Applied Statistics*, *10*, 365–394. doi:10.1214/15-AOAS897

Goldstein. (2013). Retrieved from https://www.nytimes.com/2013/08/13/nyregion/stop-and-frisk-practice-violated-rights-judge-rules.html

Hamilton, D. L., Chen, J. M., Ko, D. M., Winczewski, L., Banerji, I., & Thurston, J. A. (2015). Sowing the seeds of stereotypes: Spontaneous inferences about groups. *Journal of Personality and Social Psychology*, *109*, 569.

Hester, N., & Gray, K. (2018). For black men, being tall increases threat stereotyping and police stops. *Proceedings of the National Academy of Sciences*, *115*, 2711–2715. doi:10.1073/pnas.1714454115

Insko, C. A., Pinkley, R. L., Hoyle, R. H., Dalton, B., Hong, G., Slim, R., & Thibaut, J. (1987). Individual versus group discontinuity: The role of intergroup contact. *Journal of Experimental Social Psychology*, *23*, 250–267. doi:10.1016/0022-1031(87)90035-7

Insko, C. A., & Schopler, J. (1998). Differential distrust of groups and of individuals. In C. Sedikides, J. Schopler, & C. A. Insko (Eds.), *Intergroup cognition and intergroup behavior* (pp. 75–107). Hillsdale, NJ: Erlbaum.

Insko, C. A., Schopler, J., Pemberton, M. B., Wieselquist, J., McIlraith, S. A., Currey, D. P., & Gaertner, L. (1998). Long-term outcome maximization and the reduction of interindividual–intergroup discontinuity. *Journal of Personality and Social Psychology*, *75*, 695–710. doi:10.1037/0022-3514.75.3.695

Johnson, D. J., Cesario, J., & Pleskac, T. J. (2018). How prior information and police experience impact decisions to shoot. *Journal of Personality and Social Psychology*, *115*, 601. doi:10.1037/pspa0000130

Kurzban, R., & Neuberg, S. (2005). Managing ingroup and outgroup relationships. In D. M. Buss (Ed.), *The handbook of evolutionary psychology* (pp. 653–675). Hoboken, NJ: John Wiley.

Levinson, J. D., Cai, H., & Young, D. (2010). Guilty by implicit racial bias: The guilty/not guilty implicit association test. *Ohio State Journal of Criminal Law*, *8*, 187–208.

Lichter, D. T., Parisi, D., & Taquino, M. C. (2015). Toward a new macro-segregation? Decomposing segregation within and between

metropolitan cities and suburbs. *American Sociological Review*, *80*, 843–873.

McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, *27*, 415–444. doi:10.1146/annurev.soc.27.1.41

Payne, B. K. (2001). Prejudice and perception: The role of automatic and controlled processes in perceiving a weapon. *Journal of Personality and Social Psychology*, *81*, 181–192. doi:10.1037/0022-3514.81.2.181

Pierson, E., Simoiu, C., Overgoor, J., Corbett-Davies, S., Ramachandran, V., Phillips, C., & Goel, S. (2017). A large-scale analysis of racial disparities in police stops across the United States. *Proceedings of the Twenty-First International Conference on Artificial Intelligence and Statistics*. arXiv preprint arXiv:1706.05678.

Richeson, J. A., Todd, A. R., Trawalter, S., & Baird, A. A. (2008). Eye-gaze direction modulates race-related amygdala activity. *Group Processes & Intergroup Relations*, *11*, 233–246. doi:10.1177/1368430207088040

Ridgeway, G. (2007). *Analysis of racial disparities in the New York police department's stop, question, and frisk practices*. Santa Monica, CA: Rand Corporation.

Sawyer, J., & Gampa, A. (2018). Implicit and explicit racial attitudes changed during black lives matter. *Personality and Social Psychology Bulletin*, *44*, 1039–1059. doi:10.1177/0146167218757454

Sommers, S. R., & Marotta, S. A. (2014). Racial disparities in legal outcomes: On policing, charging decisions, and criminal trial proceedings. *Policy Insights from the Behavioral and Brain Sciences*, *1*, 103–111. doi:10.1177/2372732214548431

Thiem, K. C., Neel, R., Simpson, A. J., & Todd, A. R. (2019). Are black women and girls associated with danger? Implicit racial bias at the intersection of target age and gender. *Personality and Social Psychology Bulletin*, 1–13. doi:10.1177/0146167219829182

Wildschut, T., Pinter, B., Vevea, J. L., Insko, C. A., & Schopler, J. (2003). Beyond the group mind: A quantitative review of the interindividual-intergroup discontinuity effect. *Psychological Bulletin*, *129*, 698–722. doi:10.1037/0033-2909.129.5.698

Author Biographies

Erin Cooley is an assistant professor at Colgate University. Her research examines the cognitive, affective, and physiological processes that result from, and contribute to, social inequality.

Neil Hester is a postdoctoral researcher at McGill University. He studies how different social categories such as race, gender, and age intersect to produce unique patterns of stereotyping and discrimination.

William Cipolli is an assistant professor at Colgate University. His research is focused on collaborative statistical research and Bayesian nonparametric inference via Polya trees.

Laura I. Rivera is a lab manager for the prejudice and intergroup relations lab at UW-Madison.

Kaitlin Abrams is a graduate of Colgate University's Psychological and Brain Sciences Department.

Jeremy Pagan is a graduate student at Tufts University.

Samuel R. Sommers is a professor at Tufts University. He is a social psychologist who studies, stereotyping, diversity, and the intersection of psychology and law.

Keith Payne is a professor at the University of North Carolina at Chapel Hill. He studies unintended bias and psychological effects of inequality. More information about research can be found at http://bkpayne.web.unc.edu/

Handling Editor: Joseph Cesario