

Police Practice and Research

An International Journal

ISSN: (Print) (Online) Journal homepage: <https://www.tandfonline.com/loi/gppr20>

Examining the effects of the killing of George Floyd by police in the United States on attitudes of Black Londoners: a replication

Amy Nivette, Christof Nägel & Emily Gilbert

To cite this article: Amy Nivette, Christof Nägel & Emily Gilbert (2023) Examining the effects of the killing of George Floyd by police in the United States on attitudes of Black Londoners: a replication, *Police Practice and Research*, 24:5, 575-596, DOI: [10.1080/15614263.2023.2185241](https://doi.org/10.1080/15614263.2023.2185241)

To link to this article: <https://doi.org/10.1080/15614263.2023.2185241>



© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



[View supplementary material](#)



Published online: 02 Mar 2023.



[Submit your article to this journal](#)



Article views: 1831



[View related articles](#)





[View Crossmark data](#)



Citing articles: 1 [View citing articles](#)

Examining the effects of the killing of George Floyd by police in the United States on attitudes of Black Londoners: a replication

Amy Nivette ^a, Christof Nägel ^c and Emily Gilbert^b

^aDepartment of Sociology, Utrecht University, Utrecht, the Netherlands; ^bMayor's Office for Policing and Crime, London, UK; ^cInstitute of Sociology and Social Psychology, University of Cologne, Cologne, Germany

ABSTRACT

High-profile incidents of police misconduct can have serious consequences for public trust in the police. A recent study in the *British Journal of Political Science* found that Eric Garner's death in NYC led to more negative attitudes towards the police in London among Black residents compared to White and Asian residents. The current study aimed to replicate this transnational effect by assessing the impact of George Floyd's death on Londoners' perceptions of police. Using the same data and methodological approach, we did not replicate the immediate effect on Black Londoners' attitudes. We did find that attitudes across ethnic groups became more negative when using a wider temporal bandwidth. However, we discovered violations to the excludability assumption, meaning we cannot be certain that the effect is solely due to the murder of George Floyd, or at least partly due to different dynamics, like the effects of the COVID-19 pandemic and the accompanying policies. This means that while it is possible that police killings in other contexts play a role in shaping attitudes towards local police, these effects are difficult to disentangle from other global and local factors.

ARTICLE HISTORY

Received 25 July 2022
Accepted 23 February 2023

KEYWORDS

Protest policing; attitudes towards police; George Floyd; ethnicity; replication

Previous research has shown that high-profile incidents of police misconduct can reduce trust and confidence in the police among the wider public (Kochel, 2019; Nägel & Lutter, 2021; Reny & Newman, 2021). However, the effects of these 'vicarious' experiences can be heterogeneous and even null depending on the institutional and societal context, sub-population, and specific circumstances of the event (Muñoz et al., 2020; Nägel & Nivette, 2022). There is relatively little research on to what extent incidents of police violence in one country resonate across the globe, influencing individual attitudes towards local police.

The study 'Phantom Pains' in the *British Journal of Political Science* aimed to address this research gap by assessing the impact of the death of Eric Garner in the United States on Londoners' attitudes towards the London Metropolitan Police in the United Kingdom (Laniyonu, 2021). His study takes advantage of the overlap between fieldwork conducted in London for the Mayor's Office for Policing and Crime [MOPAC] Public Attitude Survey and the event, utilizing the interview date as an instrumental variable to assign respondents to treatment (post-event) or control (pre-event). This type of quasi-experimental design, known as unexpected event during survey design [UESD], assumes that because the event is 'unexpected' assignment to treatment and control is as-good-as-random, conditional on a set of assumptions. Laniyonu (2021) analysed treatment effects using regression discontinuity design [RDD], and found that Black Londoners

had significantly lower support for the Metropolitan Police following the death of Eric Garner. White and South Asian respondents showed little to no change following the event. Laniyonu argued that this is evidence that police violence can influence attitudes towards local police in other contexts. Specifically, he concludes that '[p]olice violence against Black persons in the US likely resonates with the experiences of Black persons in London (and elsewhere), and political actions taken in those contexts are aimed toward addressing systemic inequality there' (Laniyonu, 2021, p. 13). However, as Laniyonu (2021) rightly states, the generalizability of UESD studies is often limited due to the focus on a single case and context.

The current study aims to replicate the effect of Eric Garner's death in New York City on attitudes towards the police in London. Specifically, we conduct a replication of another high-profile event (i.e., the death of George Floyd in 2020) on attitudes towards the police using the same data from the MOPAC Public Attitude Survey in London. In addition, we evaluate the robustness and reliability of results by assessing the ignorability and excludability assumptions, as well as alternative design strategies and specifications, as recommended in the UESD literature (Muñoz et al., 2020; Nägel & Nivette, 2022). In doing so, we plan to implement the full range of 'good-practice-recommendations' discussed in the literature, thus providing a comprehensive and rigorous understanding of the hypothesized effect. The overarching goal of this replication is to critically assess the rigorousness and generalizability of causal claims regarding the transnational impact of high profile police incidents on attitudes using public opinion surveys.

The effect of police killings on attitudes towards the police

Procedural justice theory proposes that the quality of treatment and decision-making during interactions with the police can shape individual perceptions of trust and legitimacy (Mazerolle et al., 2013; Reisig et al., 2018; Solomon & Chenane, 2021; Tyler & Huo, 2002). Typically, these studies focus on direct or interpersonal interactions between an individual and a police officer. However, there is growing evidence to suggest that indirect or 'vicarious' interactions with police, such as through (social) media, can have substantive effects on attitudes towards the police (Choi, 2021; Graziano et al., 2010; Graziano, 2019; Intravia et al., 2020; Nägel & Nivette, 2022; Weitzer, 2002). In addition, these effects can be relatively greater among ethnic minority compared to majority groups, as minority group members are more likely to experience negative police contact and mistreatment (Braga et al., 2019). For example, when Michael Brown, an African American, was shot and killed by a White police officer in Ferguson, Missouri (USA), Kochel (2019) found that confidence and trustworthiness of police declined only among Black residents compared to non-Black residents. She argues that this effect was likely amplified by social identity, with African Americans 'relating more strongly to his circumstances or internalizing the experience thinking "that could have been me"' (Kochel, 2019, p. 394).

Laniyonu (2021) draws on this body of research to argue how and when incidents of police misconduct influence attitudes towards the police outside of the national context in which the incident occurred. Specifically, he outlines three explanations for why high-profile incidents in the United States could have an effect on attitudes among Afro-descendants in other contexts: 1) Afro-descendant persons maintain a shared cultural affinity and attachment, and identify with experiences of other Black individuals (Narayan, 2019; Patterson & Kelley, 2000), 2) Black residents of other countries are also more likely to experience aggressive policing tactics, such that when incidents occur in the United States, this can resonate with Black residents in particular (Bradford et al., 2009; Joseph-Salisbury et al., 2020; Millings, 2013; Vomfell & Stewart, 2021), and 3) Black residents may perceive police violence as rooted in shared experiences of racism and marginalization (Goldsmith & McLaughlin, 2022; Waters, 2014).

At the same time, Laniyonu provides several reasons why incidents in the United States may *not* have an effect on attitudes among Black residents in other countries. For example, while there may be shared identity and feelings across countries, the strength of the collective identification may

differ across contexts. Police-community relations and experiences with the police may be relatively more positive among Black residents in other countries, such as Canada and the United Kingdom, compared to the United States (Laniyonu, 2021; cf. Wortley & Owusu-Bempah, 2011). The lethality of encounters with the police also differs substantially between countries. The rate of fatal police violence is higher in the US compared to England and Wales, with estimated rates of 3.418 per 1 million (2020) and 0.286 per 1 million (2019) respectively (Hirschfield, 2023). In addition, police organizations in the UK, including the Metropolitan Police, released a statement on June 3rd condemning the death of George Floyd.¹ This statement may have served as a signal to the public condemning racial injustice and reaffirming commitment to policing by consent. There are also important within-group differences in Afro-descendant attitudes towards the police, whereby Afro-Caribbean Britons have more negative perceptions of police compared to African Britons (Bradford, 2011).

In 'Phantom Pains', Laniyonu (2021) focuses on a high-profile incident of police violence in which Eric Garner died when New York City police officers restrained him lying face down on the sidewalk. On 17 July 2014, officers approached and attempted to arrest Garner for selling loose cigarettes, pinning him to the ground and placing him in a chokehold (prohibited by the NYPD). Garner stated several times that he could not breathe while pinned down before losing consciousness. He was pronounced dead about one hour later at a local hospital. Video footage captured by a bystander was released on social media soon after the event, motivating protests and the growth of the Black Lives Matter movement worldwide (Lyle & Esmail, 2016).

Based on procedural justice research and the notion of shared identity and experiences, Laniyonu (2021, p. 6) hypothesizes that 'the police killing of Eric Garner will reduce Black Londoners' evaluations of police.' Similarly, he hypothesizes that the killing of Garner will have no effect on White and South Asian Londoners' evaluations of police. Laniyonu argues, however, that the incident will only affect certain types of attitudes, specifically perceptions of fairness and satisfaction, but not effectiveness and community engagement. His third hypothesis therefore states that 'task-specific evaluations of police behavior and police work – such as effectiveness in fighting crime or engaging with the community – will be unaffected by police killings in the United States' (p. 7).

The current study

Laniyonu takes a clear approach to evaluating these hypotheses, including conducting multiple tests of relevant causal assumptions (e.g., examining non-response, pre-existing trends, simultaneous events, covariate adjustment, and examining alternative bandwidths). He shows that the results are mostly robust to these tests and alternative specifications. However, conclusions drawn from UESD studies are somewhat restricted in that they focus on a single case, limiting external validity and generalization of effects (Muñoz et al., 2020). Furthermore, recent literature on quasi-experimental designs such as UESD have emphasized the importance of checking all relevant assumptions, establishing adequate power, and evaluating the robustness of results across different model specifications (e.g., RDD and OLS, see Muñoz et al., 2020; Nägel & Nivette, 2022). For these reasons, the close evaluation of assumptions, specifications, and different case studies is necessary to determine to what extent the effect detected after an incident is robust and generalizable.

Our current study therefore aims to assess the robustness and replicability of the effect of police killings in the United States on attitudes towards the police among Black Londoners. We take Laniyonu's paper as a model for our replication, while incorporating additional tests of assumptions and alternative specifications as recommended by recent UESD literature (Muñoz et al., 2020; Nägel & Nivette, 2022). Specifically, we examine the effect of the death of George Floyd, an African American man who was killed after a White police officer knelt on his neck during a police stop on suspicion of using a counterfeit bill (BBC News, 2020b). Floyd was handcuffed, lying face down on the street, complaining of breathing difficulties. After several minutes, with the officer's knee still on

his neck, Floyd became silent, and another officer could not detect a pulse. About an hour after Floyd was taken to the hospital, he was pronounced dead. The death of George Floyd sparked rapid mobilization and large-scale protests across the United States (Reny & Newman, 2021). The circumstances of Garner and Floyd's death are relatively similar, however the cases differed in the scale of media coverage following the incident. The size, scale, and broad (social) media coverage of the protests within the United States and internationally (Dunivin et al., 2022; van Haperen et al., 2022), made the George Floyd incident more likely to exert broad opinion-mobilizing effects among the public compared to relatively smaller-scale incidents and protests (Reny & Newman, 2021). Indeed, Reny and Newman (2021) show that the death of George Floyd and subsequent protests lead to less favorable views of the police across ethnic groups in the United States, but particularly among individuals with low prejudice and those identifying as politically liberal (i.e., Democrats).

Following Laniyonu (2021, pp. 6–7), we assess two hypotheses that aim to replicate the effects of the killing of George Floyd on Black Londoners' attitudes towards the police.

- H1: The police killing of George Floyd will have an effect on Black Londoners' overall evaluations (i.e., public satisfaction) and perceptions of police fairness, in comparison to White and Asian Londoners.
- H2: The police killing of George Floyd will have no effect on Black Londoners' task-specific evaluations of police behavior and work, such as police effectiveness in fighting crime or engaging with the community.

Methods

This paper has the goal to replicate the results of Laniyonu (2021) using the recent case of the death of George Floyd in 2020. The current study replicates the original UESD design, which combines observational data with natural experiments (i.e., police killings) to assign respondents to treatment (post-event) and control (pre-event). Laniyonu's study uses the MOPAC Public Attitude Survey [PAS] collected in 2014. In order to replicate the results, we use the MOPAC Public Attitude Survey [PAS] collected in 2020. The 2020 PAS contains most (but not all) of the same items used to measure the independent and dependent variables, and so we are able to replicate the original design and analysis. A list of variables included in the 2014 and 2020 PAS data is available in Appendix A. The replication was pre-registered on the Open Science Framework on 15 June 2022 (<https://osf.io/tjr7a/>). Any deviations from the protocol are noted below.

The code used to conduct this analysis is available on OSF (see link above). The data are available from the third author on request.

Data

The MOPAC Public Attitude Survey [PAS] is a continuous survey that conducts face-to-face interviews with a random sample of respondents across 32 London boroughs (excluding the City of London). The survey aims to complete 100 interviews in each borough per quarter, resulting in about 12,800 interviews per year. The survey covers a variety of topics surrounding perceptions of police, crime and safety, and people's experiences with crime and anti-social behavior. Prior to April 2020 and the COVID-19 pandemic, surveys were administered through Computer-Aided Personal Interviewing (CAPI), with interviewers visiting respondents in their homes. An address-based sampling approach was used, with residential addresses randomly selected from the Postcode Address File (PAF) and then a single respondent within a household randomly selected at the point of contact with the household. Following the onset of the COVID-19 pandemic, interviews were conducted via telephone in order to comply with local restrictions. Given the UK does not have a telephone sampling frame, a dual approach was used to select telephone numbers to form the

sample. Firstly, a Random Digit Dialing approach was used. London-based landline telephone prefixes were auto-filled, and then random numbers were inserted in order to complete the telephone number. Secondly, a list of mobile telephone numbers known to belong to London residents who had consented to sharing their telephone number in this way was purchased from a third-party supplier. The third-party company is a commercial supplier of telephone numbers for market research and other purposes.

For both methods of data collection, the sample achieved is broadly representative of the London population across a range of demographics. Post-hoc weights are then created for each annual sample to ensure the sample reflects the London population in terms of gender, age, ethnicity, housing tenure and working status. A total of 12,736 interviews were conducted for fiscal year 2020–2021.

Power considerations

PAS is an ongoing survey that occurs year-round, with new waves beginning each fiscal year (April). The event occurred in May of 2020. For the replication analyses, we will use data collected for the fiscal year in which the event took place (i.e., 2020–2021). This ensures that the contents of the survey remain the same throughout the time period covering the event. A total of 1928 interviews were conducted prior to the event ($n = 10,808$ post-event). We conducted a power sensitivity analysis using G*Power to determine the minimum size of the effect that can be reliably detected using the given sample (Perugini et al., 2018). Using a two-tailed test with an alpha of 0.05 and power level of 0.80, we would be able to detect an effect size of $d = 0.069$. This suggests that we will be able to detect reasonably small effects if using the full bandwidth.

Variables

We selected variables for replication according to the original design, as reported by Laniyonu (2021), as well as additional variables to examine robustness of different scale constructions and related outcomes. Some items are only available in 2014 and not 2020 as a result of changes to the questionnaires, and so we are not able to use them in the replication. As a result, we had to deviate from the original study and pre-registration. We note these deviations below. The relevant variables for the replication include the following items from the 2020 questionnaire.

Treatment variable

The binary treatment indicator D_i will be constructed as follows:

$$D_i \begin{cases} = 0 & \text{If respondent } i \text{ is interviewed before the death of George Floyd} \\ 1 & \text{If respondent } i \text{ is interviewed after the death of George Floyd} \end{cases}$$

Dependent variables

In the original analysis, Laniyonu included four dependent variables: public satisfaction with police, police fairness, police effectiveness, and community engagement. The 2020 PAS did not contain the items on police effectiveness, and so we were unable to include this outcome in the analyses. In the original study, Laniyonu reconstructed the scales so that the indices range from -5 (very negative evaluation) to $+5$ (very positive evaluation), with 0 reflecting neutral response values. We deviate from this approach and create a simple mean scale where outcomes range from 1 (negative

evaluation) to 5 (positive evaluation). Scales were reverse-coded when necessary so that higher values equals more positive evaluations.

Public satisfaction with the police is measured using two items that ask respondents ‘Taking everything into account, how good a job do you think the police in London as a whole are doing?’ And ‘how good a job do you think the police in your area are doing?’ Note that the original analysis uses only the first item, whereas we include the second because it was available and would improve reliability.² Responses range from 1 ‘very poor’ to 5 ‘excellent’ (Cronbach’s alpha = 0.77).

Police fairness was originally measured using four items, however the 2020 PAS contained only two of these four items. These asked respondents to what extent they agree with statements about police in their area (their area is defined by 15 minutes’ walk from their home). The statements include: ‘they [the police] can be relied on to be there when you need them,’ and ‘the police in your area treat everyone fairly regardless of who they are.’ Respondents could indicate their agreement on a five-point Likert scale from 1 ‘strongly disagree’ to 5 ‘strongly agree’ (Cronbach’s alpha = 0.71). The items that are not included ask whether the police are ‘helpful’ and ‘friendly and approachable.’

Community engagement is measured using three items. Again, we were not able to include all of the items from the 2014 survey, as they were not fielded in the 2020 survey. These items asked respondents to what extent they agree that the police ‘are dealing with the things that matter to people in this community,’ ‘listen to the concerns of local people,’ and ‘can be relied upon to be there when you need them.’ Responses are measured on a 5-point Likert scale ranging from 1 ‘strongly disagree’ to 5 ‘strongly agree’ (Cronbach’s alpha = 0.81). We were not able to include the original item asking whether respondents agreed that the police ‘understand issues facing the local community.’

Grouping variable: ethnicity

In the original paper, ethnicity was recoded to create three groups: White, Black, and South Asian. Respondents who reported ‘mixed’ ethnicity (i.e., White and Black Caribbean, White and Black African) were recoded and included in the ‘Black’ ethnicity category.

In the 2020/2021 data, six broader categories of ethnic identity are available: Asian, Black, Mixed ethnicity, White British, White Other, and Other ethnicity.³ We are not able to decompose the ‘Mixed’ category in the available data due to privacy reasons, so we exclude this category as well as ‘Other ethnicity’. Note we are also not able to specify ‘South Asian’ respondents from the broader ‘Asian’ category. We will combine the White British and White Other categories to create a broader ‘White’ ethnicity category. As such, for the current replication, we use the following three ethnic categories: White, Black, and Asian.

Covariates

A number of covariates are included to conduct balance checks. In the original paper, Laniyonu includes: whether the respondent was stopped by the police, whether the respondent was searched or arrested, whether the respondent contacted the police, the respondent’s employment status, age, and borough of residence. However, several of these covariates were not available in the 2020 PAS: whether the respondent was stopped, searched or arrested by police or has contacted the police. Our study therefore deviates slightly from the pre-registration and original study, as we are only able to include employment status, age, and borough of residence. In line with the pre-registration, we include two additional covariates in the replication: gender and whether the respondent has been a victim of crime or antisocial behaviour in the past 12 months.

Robustness and placebo checks

As an extension to the original study, we include several alternative outcomes that are in line with theoretical expectations (Sunshine & Tyler, 2003), as well as items relevant for placebo checks. The

alternative outcomes include items that ask respondents to what extent they agree that ‘the Metropolitan Police Service is an organisation I can trust,’ whether the respondent feels ‘an obligation to obey the law at all times,’ and feels ‘an obligation to follow police orders.’ Items used for placebo checks ask to what extent the respondent agrees that the Central Government, National Health Service (NHS), and media companies are organizations that they can trust. All responses are measured on a 5-point Likert scale ranging from 1 ‘strongly disagree’ to 5 ‘strongly agree.’ In our pre-registration, we also planned to use the item ‘The police have the same sense of right and wrong as I do,’ however, it was ultimately not available in the dataset.

Analytical approach

In order to replicate the results of the original paper, we make use of the UESD identification strategy, which relies on the timing of the interview as an instrumental variable to assign respondents to control (pre-event) or treatment (post-event) groups. Because the primary objective of the proposed study is a replication of a previously published case study (Laniyonu, 2021), our design and analysis techniques resemble those used in that paper. This means we apply a regression discontinuity design which is a quasi-experimental method to analyze observational data. As part of the design, a so-called running variable is defined, which, in this case, is the day of the interview. The name-giving discontinuity results from also defining a cut-point at which a change in the level and trend of the dependent variable(s) is assumed. This cut-point is 26 May 2020, the day after George Floyd was killed by police officers in Minneapolis (likewise 17 July 2014 for Eric Garner’s death). George Floyd was killed on May 25, around 8:00 AM, which corresponds to 4:00 PM London time. Since it might have taken some time for the video of the incident to spread, it is very unlikely to assume that respondents interviewed on May 25 would have been exposed to the event. Accordingly, we use May 26 as the cut-point. However, we also run robustness checks where we exclude respondents interviewed on May 25, since we cannot be entirely sure that they would belong to the control group. Respondents who were interviewed just before that day should theoretically be comparable to respondents who were interviewed just after that day, given that the interview timing is essentially as-good-as-random. Hence, all other things equal, any difference in the respective outcome (attitudinal measures of police) should be attributable to the event.

The number of days before and after the cut point (i.e., the bandwidth) is determined by a completely data-driven algorithm that approximates the smallest Mean Square Error (MSE) of the local average treatment effect (LATE). We apply both the MSE method used by Laniyonu, as well as another commonly applied procedure (Imbens & Kalyanaraman, 2012). Specifically, we replicate the ‘uniform’ kernel method to estimate the LATE that Laniyonu employed and extend this by the more commonly applied ‘triangular’ kernel approach.⁴ We run models with and without covariates and with and without borough level clustered standard errors. While we use the ‘rdrobust’ package utilized by Laniyonu, we also run similar models with another software package, namely the ‘rdd’ package provided by Dimmery (2016). To explore the alleged heterogeneity of the treatment effect reported in Laniyonu, this analysis is conducted for the different ethnic groups included in the data. While we closely mirror the analysis in the Laniyonu paper, we also apply all robustness checks that are discussed in the relevant literature (Muñoz et al., 2020; Nägel & Nivette, 2022).

Specifically, we extend the RDD identification strategy by a more conventional difference-in-differences (DiD) specification in which we create the treatment variable from the product term of the respective ethnic groups and the binary event indicator. The regressions themselves are estimated by way of ordinary least squares. The hypotheses outlined above are tested by recoding the interaction of the binary event indicator with the respective ethnic groups (i.e., Black compared to all other included groups). While we base our results mainly on a model including the whole fieldwork period in order to increase statistical power, we also re-run the analysis with smaller, weekly bandwidths.

We make inferences about the effect of the death of George Floyd on perceptions of police based on p-values. We conclude that the analysis supports our hypothesis if the p-value is smaller than 0.05 using a two-tailed test. More specifically, we use heteroscedasticity-robust Huber-White standard errors (HC0) for all statistical inferences using the ‘lm_robust’ command from the estimator package. In order to closely replicate Laneyonu’s approach, we will also apply HC2 standard errors.

Assumptions

The validity of conclusions drawn from our design hinges on two important causal inference assumptions. (1) *Excludability* implies that the timing of the interview affects the outcome through no other channel than the event itself, i.e., the exclusion restriction in instrumental variable analysis (see Labrecque & Swanson, 2018; Murray, 2006; Stock & Watson, 2007). This assumption can be violated when collateral or simultaneous events occurred unrelated time trends are present or the timing of the event itself is endogenous. We follow the literature (Muñoz et al., 2020) and apply the following robustness checks: Examination of pre-existing time trends as well as falsification tests on other units (placebo specifications for the previous year of data collection) and on other outcomes (trust in government, NHS, media) will be assessed in each case study. (2) The *ignorability* assumption addresses whether the chance of being assigned to either the control or the treatment group is as-good-as-random (Bor et al., 2014; Gangl, 2010; Legewie, 2013; Muñoz et al., 2020). Possible violations include imbalances on observables, reachability, attrition, noncompliance, heterogeneous effects. We will conduct balance tests, choose multiple bandwidth selections, adjust for covariates, and analyze non-response patterns and placebo treatments. In the pre-registration, we stated that we would conduct weekly bandwidths extending four weeks (i.e., 14, 21, 28, 35 days). Here we extend the bandwidth selection to include all possible weekly bandwidths following the event. This allows us to assess the widest possible range of effects. We provide the results of all these assumption checks in the appendix and present a brief discussion of the results in the paper.

Missing data

We use listwise deletion to restrict observations to respondents with complete information available for all covariates used in the models. We, however, do not delete observations a priori based on missing values. This way, the maximum number of observations will differ between baseline models including just the binary treatment indicator and full models including the treatment variable and covariates.

Results

Descriptive results

Table 1 summarizes the descriptive statistics of the main variables used in this analysis. The column ‘mean Δ ’ provides the calculated difference between the group as well as a statistical test of significance (t-test) to analyze potential imbalances between groups. While all three outcome variables demonstrate significant negative differences between the treatment and the control group, the covariate distributions between respondents interviewed before and after the George Floyd incident remains largely balanced. However, the post-treatment group contains a larger number of respondents in the 35 to 44 age category, fewer respondents in the 45 to 54 age category, fewer respondents not working and more respondents who report to be unemployed. We base our results on models that include all covariates that are presented here as well as Borough level fixed effects.

Table 1. Descriptive statistics and balance checks.

Control group						Treatment group						
Statistic	N	Mean	St. Dev.	Min	Max	Statistic	N	Mean	St. Dev.	Min	Max	Mean Δ
Satisfaction	1,774	3.651	0.78	1	5	Satisfaction	9,594	3.455	0.857	1	5	-0.196 ***
Fairness	1,747	4.222	0.814	1	5	Fairness	9,586	3.946	0.897	1	5	-0.276 ***
Engagement	1,661	3.855	0.866	1	5	Engagement	9,085	3.639	0.914	1	5	-0.216 ***
Male	1,949	0.444	0.497	0	1	Male	10,787	0.439	0.496	0	1	-0.005
Female	1,949	0.556	0.497	0	1	Female	10,787	0.559	0.496	0	1	0.003
Other	1,949	0.001	0.023	0	1	Other	10,787	0.001	0.037	0	1	0
16 to 24	1,948	0.093	0.29	0	1	16 to 24	10,787	0.078	0.268	0	1	-0.015
25 to 34	1,948	0.226	0.418	0	1	25 to 34	10,787	0.214	0.41	0	1	-0.012
35 to 44	1,948	0.199	0.399	0	1	35 to 44	10,787	0.234	0.424	0	1	0.035 ***
45 to 54	1,948	0.146	0.354	0	1	45 to 54	10,787	0.128	0.334	0	1	-0.018 *
55 to 64	1,948	0.135	0.342	0	1	55 to 64	10,787	0.141	0.348	0	1	0.006
65+	1,948	0.201	0.401	0	1	65+	10,787	0.205	0.403	0	1	0.004
Victim	1,945	0.194	0.396	0	1	Victim	10,774	0.184	0.387	0	1	-0.01
Fulltime	1,919	0.439	0.496	0	1	Fulltime	10,536	0.453	0.498	0	1	0.014
Parttime	1,919	0.146	0.354	0	1	Parttime	10,536	0.147	0.354	0	1	0.001
Not working	1,919	0.105	0.307	0	1	Not working	10,536	0.07	0.256	0	1	-0.035 ***
Houseperson	1,919	0.031	0.174	0	1	Houseperson	10,536	0.026	0.16	0	1	-0.005
Retired	1,919	0.184	0.388	0	1	Retired	10,536	0.2	0.4	0	1	0.016
Unemployed	1,919	0.032	0.175	0	1	Unemployed	10,536	0.043	0.202	0	1	0.011 **
Student	1,919	0.043	0.203	0	1	Student	10,536	0.037	0.189	0	1	-0.006
Other employment	1,919	0.018	0.134	0	1	Other employment	10,536	0.024	0.154	0	1	0.006

The differences in the outcome variables can also be observed in a graphical way in [Figure 1](#). The three outcome evolution plots visualize the results of linear models that estimate the respective outcome variables regressed on the interview date. The dashed line represents 25 May 2020; the day George Floyd was killed. There are visible but small discontinuities in the three outcome variables 'Satisfaction,' 'Fairness,' and 'Engagement.'

The histogram below the outcome evolution plots represents the daily number of interviews. As can be seen in [Figure 1](#), the number of interviews did not change in any way before or after the incident. Accordingly, it is unlikely that the results are due to non-response patterns.

Main analysis

Regression discontinuity design

In line with our pre-registration, we first attempted to study the effect of the George Floyd incident on the attitudes of Black Londoners in a similar way to how Laniyonu investigated the effect of Eric Garner's death. Accordingly, we split the dataset by the three respective ethnic groups; Black, White, and Asian and performed a regression discontinuity analysis for each subsample with the exact same model specification according to Laniyonu's original analysis. As outlined in the pre-registration, we used 26 May 2020, as the cut-point since it might have taken some time for the video of the incident to spread, and it is very unlikely to assume that respondents interviewed on May 25 would have been exposed to the event. Results are summarized in [Figure 2](#).

The bandwidths for satisfaction, fairness, and engagement were estimated at 18.24, 14.76, and 17.45 days, respectively. [Figure 2](#) shows that, while the estimates are predominantly negative, none of the models show a significant difference in attitudes between the before and after comparison groups (see Appendix B for the full results). This means that we do not detect an effect on any attitudinal outcomes within the first 2 to 2.5 weeks following the incident.

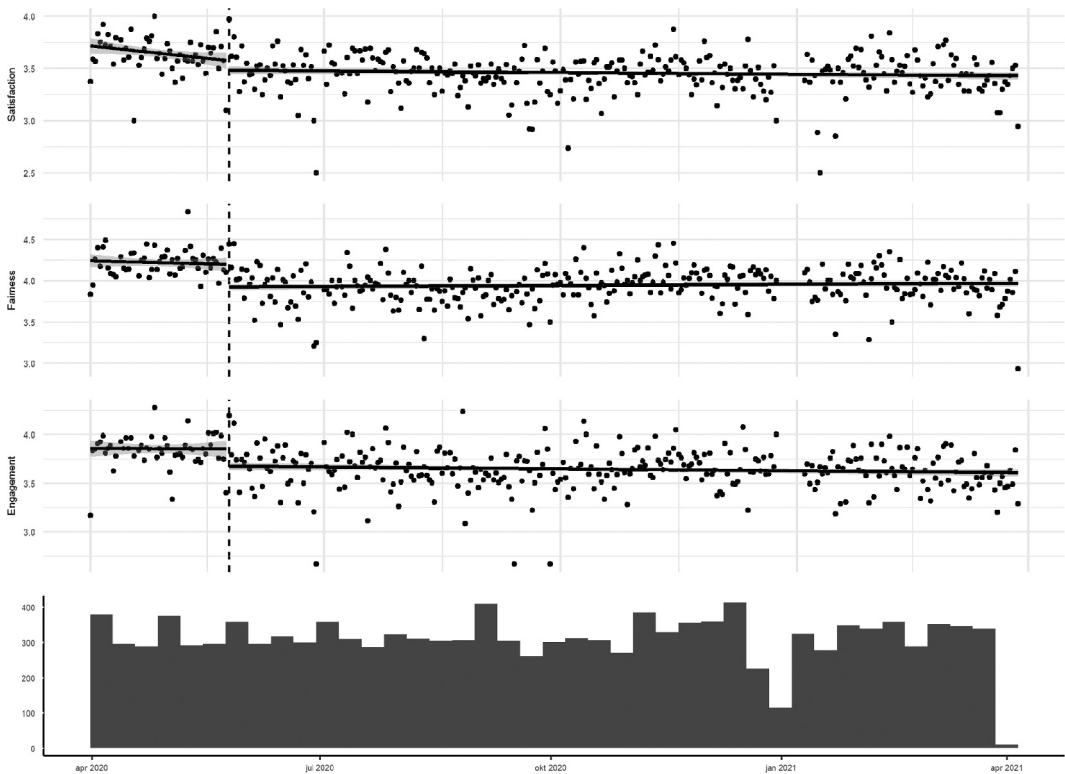


Figure 1. Outcome evolution before and after the George Floyd incident. Histograms represent daily number of interviews. Dashed line is May 25, 2020.

Simple before/after design

We extend the RDD identification strategy by a more conventional before/after specification that resembles a difference-in-differences (DiD) design, in which we create the treatment variable from the product term of the respective ethnic groups and the binary event indicator. The regressions themselves are estimated by way of ordinary least squares (OLS). The hypotheses outlined above can be tested by creating an interaction term between the binary event indicator and the respective ethnic groups (e.g., Black compared to all other included groups). Variables that were used to create the interaction were mean-centered, so that the respective main effect can be interpreted as is. In contrast to the RD models, this approach uses the entire bandwidth before and after the event (42 weeks post-event).

The results in [Figure 3](#) indicate a negative main effect on all three outcomes. The interaction terms, however, are only significant for White respondents, showing a small positive effect. This indicates that the negative treatment effect was less pronounced for White respondents compared to Black and Asian respondents. In order to demonstrate this more clearly, we re-modelled the interaction to estimate the effect of non-white respondents (i.e., Black and Asian respondents) compared to White respondents. [Figure 4](#) hence shows that the negative effect on all three outcomes was slightly more pronounced for non-white respondents as compared to White respondents.

Returning to our hypotheses, these findings imply that there is little evidence that attitudes changed directly after the incident (i.e., within the 2–2.5 weeks bandwidth estimated in the RDD), since the RD estimates are almost always non-significant. There is also little evidence that Black respondents' attitudes changed more negatively compared to White and Asian respondents in response to the event. While the only significant point estimates were found in the White Londoner

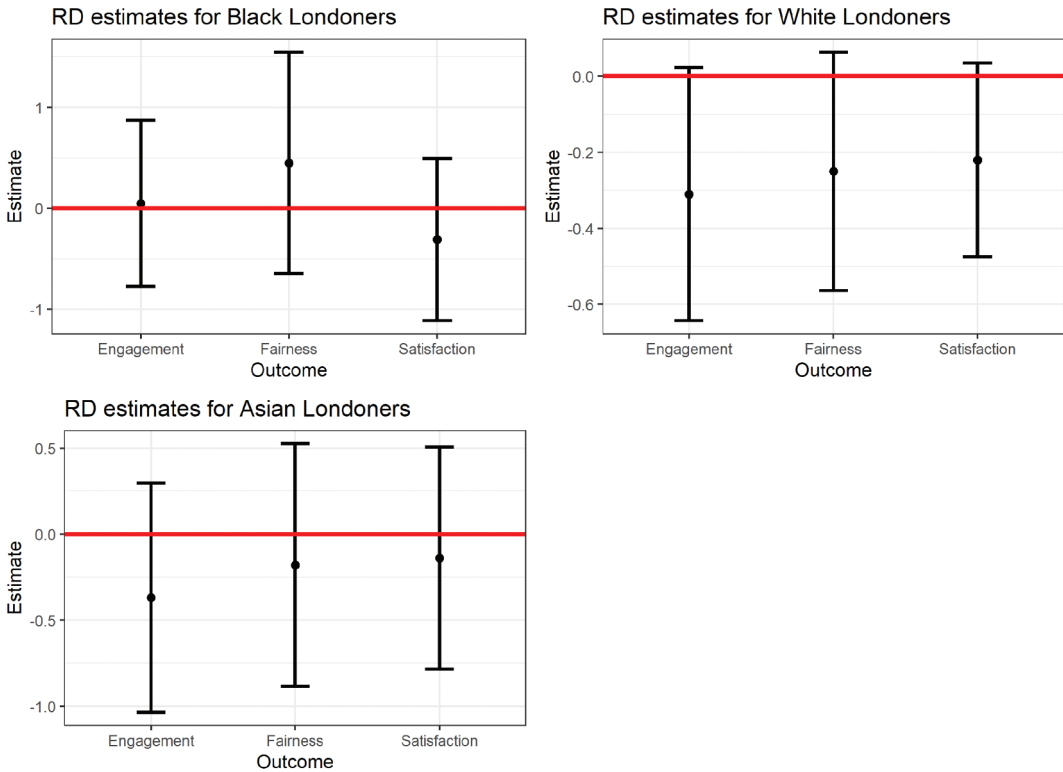


Figure 2. RD estimates and 95% confidence intervals of the effect of George Floyd's death on attitudes towards police (uniform kernel, HC2 standard errors, clustered at the borough level).

sample, we attribute this finding rather to a power issue in the other subsamples as opposed to evidence for a heterogeneous treatment effect. The simple before/after comparison on the other hand, which includes the entire post-event bandwidth (42 weeks), indicates a negative effect on all three outcomes. However, there is only very limited heterogeneity in the effect sizes between ethnic groups. Thus, we do not find support for H1, since the killing of George Floyd by police did not have an effect on Black Londoners' overall evaluations (i.e., public satisfaction) and perceptions of police fairness, in comparison to White and Asian Londoners. Instead, the negative effect is ubiquitous among ethnic groups, but seemingly more pronounced for non-white respondents compared to White respondents.

In H2 we expected that the police killing of George Floyd will have no effect on Black Londoners' task-specific evaluations of police behavior and work, such as police effectiveness in fighting crime or engaging with the community. While we could not study police effectiveness since the items were not included, we found evidence that engagement was similarly affected by the incident as the other included outcome variables. Accordingly, we do not find support for this hypothesis.

Specification changes

In line with our pre-registration, we considered a number of specification changes to test the robustness of the results. These tests include: setting the cut-point to May 25 and May 27 respectively instead of May 26, (no changes to previous results except for a significant negative effect on satisfaction in the White subsample when using May 27 as cut-point, see Figures C1 and C2 in Appendix C), changing the kernel to estimate the bandwidth from 'uniform' to 'triangular' (the effect is negative and significant for engagement and satisfaction in the White Londoner sample and

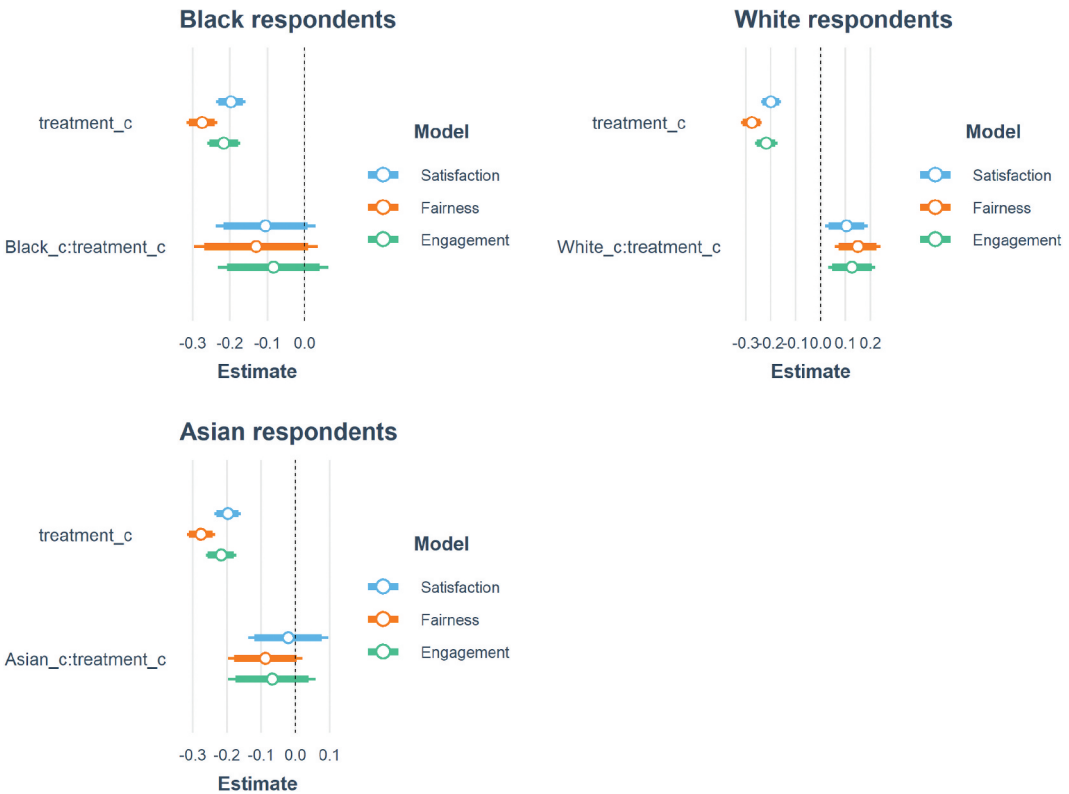


Figure 3. DiD estimates for all three ethnic groups. All covariates and borough fixed effects are included.

otherwise non-significant, Figure C3), and changing the software package to the ‘rdd’ package (Dimmery, 2016), which also did not alter the main results. The different RD analyses point to null or mixed results (see Figure C4). Almost all point estimates are negative but non-significant. When estimating the effect with a triangular kernel the effect becomes significant for the satisfaction and engagement outcome in the White Londoners sample. This points to the limited power when using a RD approach to identify the effect since both the triangular kernel and the White Londoner sample increase the effective sample size of the analysis. The issue of power limitations has been raised by other researchers when applying the RD approach in a political science context (Stommes et al., 2021).

Concerning the more conventional before/after specification, we pre-registered to re-run the analysis with smaller, weekly bandwidths. We decided to use all possible bandwidth choices, which resulted in 42 models for each outcome, since the fieldwork period lasted for up to 42 weeks after the George Floyd incident. We always used all respondents interviewed before the incident as the control group, and iteratively added one week to the ‘treatment’ group, starting with one week after the incident. Figure 5 is an attempt to visualize these $42 \times 3 = 126$ models. The picture that emerges from Figure 5 is quite clear: all estimates are significantly negative for all three outcomes when using a time window of around 2–3 weeks after the event. This points either to limited power of models using a shorter treatment window, or the notion that the negative effect of the George Floyd killing took some time to develop.

Placebo and robustness checks

We now turn to assess the robustness of the results according to the causal inference assumptions discussed in the literature (Muñoz et al., 2020). All results of the robustness and placebo checks are

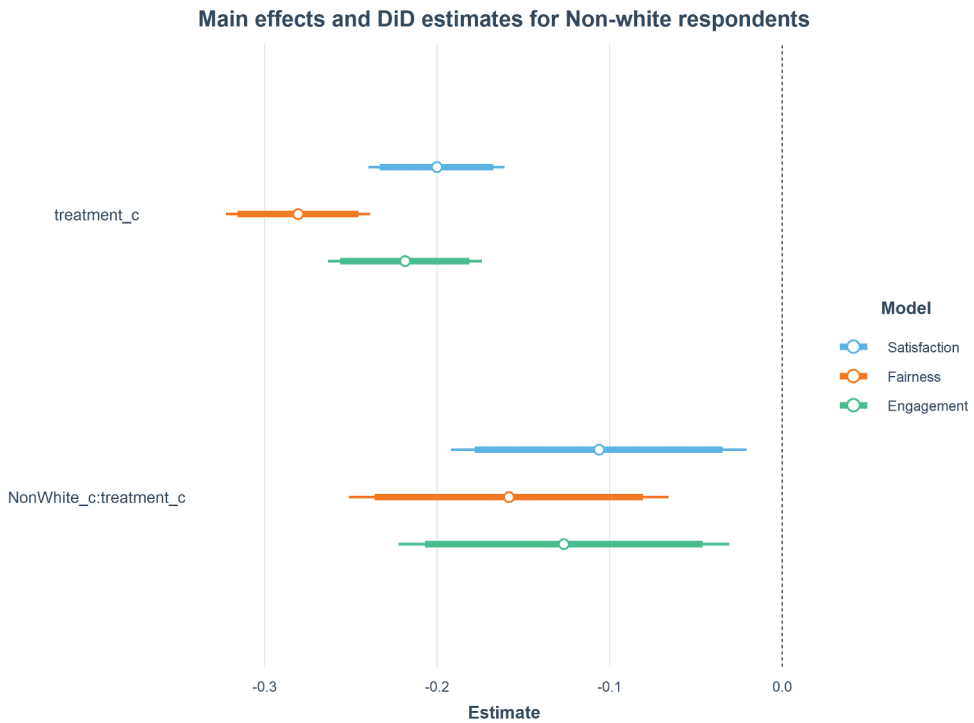


Figure 4. DiD of non-white respondents (i.e., Black and Asian respondents) compared to white respondents.

included in Appendix D. Covariate adjustment, balance checks for the complete sample, and the sensitivity of the results to the bandwidth choice have already been assessed within the previous analyses. These did not point to any violations. To further assess the validity of the ignorability assumption, we re-tested the balance for the Black subsample, inspected potential pre-existing time trends, and assessed whether there is a placebo effect when using the median of the control group as the cut-off point. Except for a slight imbalance (there less people who report to be not working in the post intervention sample), the results suggest that being assigned to the pre- or post-intervention group would not be determined by anything else than chance, i.e., according to our tests, the ignorability assumption is met.

To further investigate potential violations that relate to excludability, we checked whether there are effects of the treatment on unrelated items. Indeed, we find significant placebo effects on ‘trust in the central government,’ ‘trust in the National Health Services,’ and (although only significant for $p < 0.1$) ‘trust in media companies,’ which imply that the excludability assumption is violated (see Figure D4). The turbulent time in which the George Floyd incident occurred, namely the COVID-19 pandemic, might be one explanation as to why we see these placebo effects. Accordingly, our estimates are difficult to distinguish from potential spillover effects of collateral events since we cannot be entirely sure whether the effects we are identifying are solely due to our proposed treatment, i.e., the murder of George Floyd, or at least partly due to different dynamics, like the effects of the COVID-19 pandemic and the accompanying policies.

We also investigated whether we see similar effects with different survey rounds of the MOPAC Public Attitude Survey. Here, we used May 25 as the cut-point for the respective five years, from 2015 to 2019, to study whether a seasonal trend might be responsible for driving the results. As can be seen in Figure 6, there is only one significant placebo result when using 25 May 2018, as the cut-point for the fairness item. Generally, these results provide evidence that that dynamics in the outcome variable are probably not due to seasonal trends, which might bias the results.

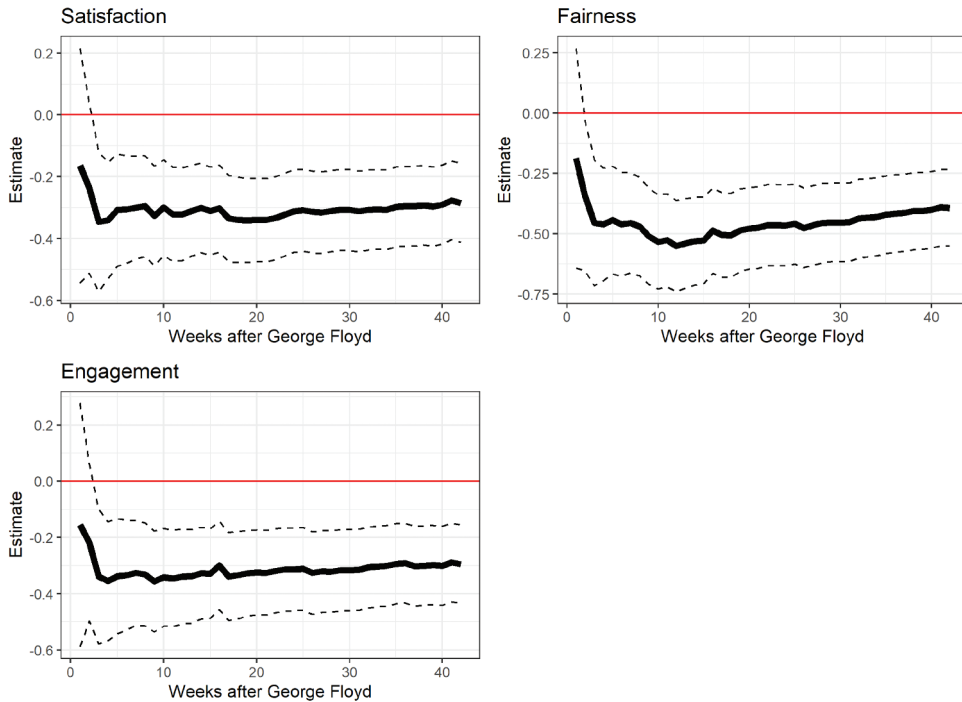


Figure 5. Point estimates and 95% confidence intervals (computed from HC0 robust standard errors) for models adding one week at a time to the ‘treatment’ period for the Black subsample. The control group uses the complete sample before May 25, 2020.

We also assessed the robustness of the results to using different outcome variables related to trust and police legitimacy. These include ‘The Metropolitan Police Service is an organisation that I can trust,’ ‘I feel an obligation to obey the law at all times,’ and ‘I feel an obligation to follow police orders.’ For all of these outcomes, we found strong negative effects (see Figure D5).

Finally, as additional analyses, we evaluate alternative cut points based on subsequent responses to the event by the Metropolitan Police and subsequent protests. This serves to examine whether Londoners were responding to the event itself, or the actions of local police. Specifically, we use RDD with the exact same specifications as in the main models (see above) to assess four additional cut points that reflect different moments in which local events and police may have shifted attitudes: 3 June (a joint statement condemning the death of George Floyd was released), 5 June (a senior Met police officer stated that the protests were unlawful and urge people not to gather in large groups), 13 June (counter-protests occur, including incidents of violence), 14 June (Metropolitan police statement condemning the violent protests and seeking a ban on further protests). The results are reported in Figure D6 in the Appendix. The results show no consistent evidence that any of these subsequent events might have had a causal impact on any of our outcomes. We do, however, find two significant effects in the Asian subgroup on both Satisfaction and Fairness with June 13 as cut point ($p < .05$). These two effects are not robust to shifting the cut point one day ahead which is why we consider these chance-findings due to running many tests. Additionally, these findings only appear in the Asian subsample and not in the Black or White subsample. It should be noted that the RDD is bound to operate with small sample sizes (in this case $n = 158$) that are close to the cut point. Non-linearities can easily be mistaken for discontinuities, which is why it is important to consider different model specifications, such as sensitivity to the chosen bandwidth and modelling of the pre- and post-slopes which might be non-linear. These two estimates are not robust to these specification changes; thus we consider them spurious.

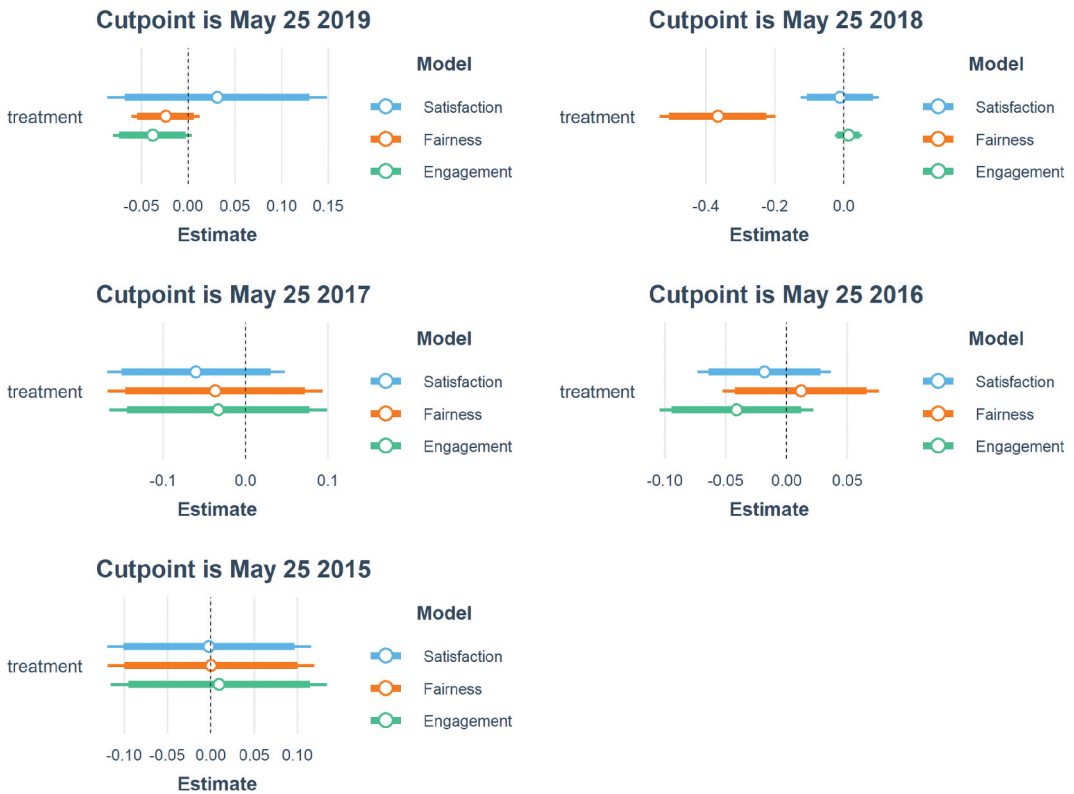


Figure 6. Placebo tests using May 25 as a cut-point in different survey years (Black subsample).

Discussion

Replication provides a tool to evaluate the generalizability and robustness of individual studies, and to what extent researchers and policymakers can be confident in certain findings (McNeeley & Warner, 2015). The current study aimed to replicate the finding that police killings in the United States can influence attitudes towards local police in other contexts. The original study found that the death of Eric Garner in the US in 2014 caused significant negative changes in evaluations of police among Black Londoners (Laniyonu, 2021). Following the original study, we used the same data source (MOPAC Public Attitude Survey), with many of the same items, and same analytical approach to examine the effect of the killing of George Floyd on Black Londoners' attitudes in 2020. Contrary to the original study, we find no immediate effects of the incident on Londoners' attitudes, regardless of ethnicity. Taking the whole time period of the survey into account, we find that attitudes were overall more negative following the death of George Floyd. There was some small heterogeneity in these effects between ethnic groups, but not in the way that was hypothesized in the original study. Black respondents did not have significantly more negative attitudes following the event compared to Asian and White respondents. Black and Asian respondents had small but significantly more negative effects compared to White respondents. One important caveat is that while we were able to establish that the ignorability assumption has not been violated, the excludability assumption was violated. This means that we cannot be certain that the death of George Floyd can exclusively account for these changes in attitudes towards police. We discuss two main implications for understanding how high-profile incidents contribute to public perceptions of police in different contexts.

First, while we were not able to reproduce the immediate effect of the incident on Black Londoners' attitudes, as demonstrated in the original study, we find evidence that the death of George Floyd was associated with overall more negative attitudes towards local police in London within a longer timeframe. This is in line with studies that show high-profile incidents of police violence can have substantive impacts on trust and police legitimacy (Kochel, 2019; Nägel & Lutter, 2021; Nägel & Nivette, 2022; Reny & Newman, 2021). However, these effects were not heterogeneous in the way that was expected based on previous research. Specifically, while Black Londoners showed more negative evaluations post-event compared to White Londoners, there was no significant difference between Black and Asian Londoners' evaluations of police following the event. This may be because the vast scope of the global media response to the murder of George Floyd led to greater exposure of the incident across population groups (Barrie, 2020; *The Economist*, 2020). The slightly larger effect among non-Whites suggests that the event resonated with Asian residents to a similar degree as with Black residents. This may be because of feelings of shared identity as ethnic minorities who perceive to be subject to institutional racism and aggressive policing tactics (Millings, 2013).

One can argue that media and social media scrutiny of police was more prevalent in 2020 compared to July 2014. This heightened scrutiny by the public could account for the more general effects observed following the death of George Floyd compared to Eric Garner. The death of Garner contributed to the re-ignition of the Black Lives Matter movement, which gained significant momentum following the police killing of Michael Brown in Ferguson just a few weeks after Garner's death. These high-profile incidents and subsequent protests across the US may be seen as a 'turning point' in public scrutiny of police (Capellan et al., 2020), igniting grassroots social movements calling for racial justice and police reform (Carney, 2016; Edrington & Lee, 2018). Analyses of the number of tweets including the keyword 'Black Lives Matter' show a substantial spike following the death of Michael Brown, and the volume of BLM tweets remained generally high in the years following (Giorgi et al., 2022). While BLM activism on Twitter was present prior to the death of Garner (most notably surrounding the acquittal of George Zimmerman),⁵ the discourse of racial injustice and police violence on social media increased substantially and remained consistently present following Garner's and Brown's deaths. However, if higher scrutiny accounts for these effects, one would expect that we would observe more 'shocks' in public disapproval following other fatal shootings of Black individuals prior to George Floyd. In an analysis of attitudinal trends before and after the death of Floyd, Reny and Newman (2021) showed that unfavorable perceptions of police were generally declining in the period prior to the incident, which included a large number of fatal shootings (see Figure C.1, Supplemental Appendix, pg. 16). Variations in the volume and framing of media coverage, as well as the presence (or absence) of large-scale protests across different incidents may explain this heterogeneity in effects. More research is needed to understand how and why certain incidents are more likely to first gain widespread media coverage and second spark protests that mobilize public opinion.

However, it is interesting to note that the shifts in public opinion were not immediate (or detectable), as shown in the null results from regression discontinuity models. While news of the incident spread rapidly through traditional and social media outlets, large-scale public mobilization did not occur in London until a few days later on 30–31 May (Goldsmith & McLaughlin, 2022). Still, the RDD bandwidth included responses up to 2.5 weeks after the event. It is therefore likely that the subsequent protests and sustained global attention to the Black Lives Matter movement drove changes in attitudes, rather than the incident alone. With ongoing protests comes greater visibility and media coverage, and grassroots mobilization can increase public awareness of racial issues and 'activate' shifts in public opinion (Lee, 2002; Reny & Newman, 2021). As anti-racist protests continued in the weeks following George Floyd's death, police were also criticized for their handling of the protests and accused of using unlawful tactics against protesters (BBC News, 2020a; Smoke, 2020). Changes in public opinion following police violence in transnational contexts may therefore

depend on whether and to what extent the incident sparks similar sustained mass mobilization, and how the police respond to protest activity.

Second, an important caveat to the above findings is that placebo tests indicated that the excludability assumption was violated. This means that while it is plausible that the death of George Floyd contributed in some way to the change in attitudes, we cannot be certain that other collateral events did not (also or entirely) play a role in this shift. The most likely collateral event at that time was the beginning of the COVID-19 pandemic and surrounding policy responses. Research suggests that the COVID-19 pandemic brought structural and racial injustice into the foreground in the United Kingdom, as the pandemic and subsequent social and financial consequences have disproportionately affected minority groups (Burgess et al., 2021; Harris et al., 2021; Katikireddi et al., 2021). Following the implementation of measures to mitigate the spread of the virus (e.g., social distancing, lockdowns, curfews), police were given expanded powers to enforce these new measures (Harris et al., 2021). While the onset of the pandemic saw ‘rally’ effects, leading to more positive perceptions of police and other institutions, the longer-term effects of crisis policing can deteriorate public support and cooperation (Perry et al., 2021; Sibley et al., 2020). In addition, research on policing and the pandemic in the US and UK has shown racial disparities in COVID-policing (Harris et al., 2022; Kajeepeta et al., 2022), with BLM UK declaring that racism and COVID-19 were interrelated ‘deadly pandemics’ (Goldsmith & McLaughlin, 2022). The effect of the death of George Floyd and subsequent protests is therefore difficult to disentangle from ongoing effects of broader institutional racism and racialized policing brought into focus by the COVID pandemic.

It is possible that other police killings at the time may be collateral events. In order to evaluate this, we take advantage of the thorough analysis conducted by Reny and Newman (2021), who examined the effect of the death of George Floyd on attitudes towards police in the US. In their robustness and mechanism checks, they examined the trends in public opinion prior to the death of Floyd that coincides with other police killings of unarmed Black individuals. In their analysis, they find that perceptions of police did not change in response to these other killings (see Figure C.1 in the Supplemental Appendix). To our knowledge, there were no police killings in the United Kingdom near the time of the event (May-June 2020). We are therefore reasonably confident that other high profile police killings did not influence the results of the current study.

In the original paper, Lanionu (2021, p. 14) concludes that ‘efforts by the Metropolitan Police to improve Black Londoners’ attitudes toward the police are affected by events beyond their control.’ To some extent, our results can be interpreted as support for this conclusion, as the decline does coincide with a high-profile police killing in another country and an ongoing global pandemic, which are two events that are beyond the control of local police. However, we hesitate to come to the same conclusion for two reasons. First, the presence of several collateral events means that we cannot be entirely sure which factor, or combination of factors, contributed to these changes. Second, it is possible that attitudes changed as a result of how local police responded to these events. Most notably, the expanded police powers and enforcement of COVID restrictions may have played an important role in contributing to declining satisfaction with the police during this period (Ghaemmaghami et al., 2022; Perry et al., 2021). Indeed, the slight delay in effect suggests that attitudes may be responding to how the local police *reacted* to the protests instead of (or in addition to) the event and protests themselves.

Nevertheless, research suggests that these events mobilize opinion because they resonate with *local* historical and lived experiences of structural and institutional inequality. This implies that any reforms must address both embedded structural inequalities and improve accountability, transparency, and reconciliation in policing (Bell, 2017; Goldsmith & McLaughlin, 2022). For example, Bell (2017) argues that democratization of the police is one important step in reducing inequalities and building community trust. In particular, openly providing information on interactions and body worn cameras can help improve transparency, and ultimately build trust: ‘If one suspects that most police interactions go the way they should, data and transparency can potentially be a boon to

solidarity between officers and communities. Data can perhaps put what police actually do most of the time in clearer perspective' (Bell, 2017, pg. 2144). Minimizing the likelihood of disproportionate and deadly use of force is also imperative to prevent these incidents from occurring in the first place. An evaluation of police use of deadly force in the United States found that agencies that required officers to file a report for any incident in which they pointed their gun but did not shoot had significantly lower rates of fatal shootings by police compared to agencies that did not require this (Jennings & Rubado, 2017). The authors argue that this requirement may have led officers to use more caution when using their weapon, and signal an institutional commitment to minimizing unnecessary use of force. Training in de-escalation tactics, especially involving unarmed and erratic citizens, may also help to reduce unnecessary use of force incidents (Engel et al., 2022). However, more research is needed to understand the processes in which police use force, how the use of force is distributed across encounters, and identify the sources of racial disparities in the use of force (Bennell et al., 2021).

Limitations and conclusion

We highlight several limitations to this replication study. First, while we were able to use the same ongoing survey to replicate the measures and test the effect, not all items that were fielded in 2014 were available in the 2020 PAS. Notably, we were not able to test the null effect of the event on the task-specific evaluation of police effectiveness as Laniyonu hypothesized. We suspect that, given the wider placebo effects we detected, that we may have also found negative overall effects stemming from experiences with COVID policing. However, this is speculation, and more research is needed to understand how personal and vicarious experiences of injustice differentially impact affective and instrumental evaluations of police. Second, another notable difference from the original is the composition of our ethnic group categories. Due to privacy issues, we were not able to disaggregate the 'mixed' category, which may include respondents with Afro-Caribbean and African backgrounds. Without these individuals in the analysis, lose power and may have underestimated the size of the treatment effect. Third, while we have discussed the implications of the COVID pandemic on estimating causal effects, it is still important to note that the pandemic serves as an important difference between the original context and our replication. In addition, researchers have noted that the wider opinion-mobilizing effects of the death of George Floyd were in part the result of a cumulation of frustration and amplification following multiple deaths of Black Americans by police (Reny & Newman, 2021). This makes the comparison between Eric Garner and George Floyd similar in circumstance but embedded in different social and political contexts. These differences in context may in part explain why we were not able to replicate the results of the original study.

Nevertheless, replication serves as one important step in the accumulation of knowledge about how vicarious experiences influence national and international public opinion. A single study or replication tells only part of the story, and so does a single replication. Systematic reviews and meta-analyses are needed to summarize effects and evaluate conditions that lead to heterogeneity (Lösel, 2018). Although the current study did not replicate the original results, we found that it is plausible that police killings in other contexts, alongside other global and local factors, can influence attitudes towards local police. However, the presence of collateral events and violation of the exclusion restriction means that we cannot be certain which factor(s) contributed to overall changes in attitudes towards police in London. Future research should continue examine when and how these incidents lead to widespread media coverage, spark protests, and influence attitudes towards local police across the globe.

Notes

1. See: <https://news.npcc.police.uk/releases/uk-police-stand-with-those-appalled-by-george-floyd-death>.
2. These two items are generally referred to as measures of 'confidence', however Laniyonu (2021) refers to the item as 'satisfaction'. We use the term 'satisfaction' instead of confidence in order to maintain consistency with the original paper. In addition, we ran models using only the single item in line with Laniyonu's analysis, and results were unchanged.
3. The White and White Other background includes English, Welsh, Scottish, Northern Irish, British, Irish, Gypsy or Irish Traveller, and Any other White background. The Asian or Asian British category includes Indian, Pakistan, Bangladeshi, Chinese, and Any other Asian background. The Black category includes African, Caribbean, and any other Black/African/Caribbean background. The Mixed category includes White and Black Caribbean, White and Black African, White and Asian, and Any other mixed/multiple ethnic backgrounds. The Other ethnic group includes Arab and Any other ethnic group.
4. The triangular kernel method uses a local linear regression (LOESS) and gives observations closer to the cut point more weight. The uniform kernel on the other hand weights all observation equally.
5. George Zimmerman fatally shot Trayvon Martin, an unarmed 17 year old African American student, on February 26, 2012. Zimmerman claimed he acted in self-defense. On July 13, 2013, Zimmerman was acquitted of murder and manslaughter charges.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

The work was supported by the Nederlandse Organisatie voor Wetenschappelijk Onderzoek [VI.Vidi.191.135].

Notes on contributors

Amy Nivette is Associate Professor at Utrecht University. Her research interests include the determinants of institutional legitimacy and attitudes toward the law over the life course, with a particular focus on what social, experiential, and individual factors shape attitudes towards the police. She is also a Research Fellow at the Netherlands Institute for the Study of Crime and Law Enforcement.

Christof Nägel is a postdoctoral researcher at the University of Cologne. His research and teaching areas include causes and consequences of police misconduct, as well as methods of causal inference in criminological research.

Emily Gilbert is a Principal Research Officer at the London Mayor's Office for Policing And Crime (MOPAC). She gained her PhD in Survey Methodology from the University of Essex. Her research interests include attitude measurement, data quality, and the use of new technology and innovative methods for data collection, as well as public perceptions of the police.

ORCID

Amy Nivette  <http://orcid.org/0000-0003-0597-3648>

Christof Nägel  <http://orcid.org/0000-0003-2161-539X>

References

- Barrie, C. (2020). Searching racism after George Floyd. *Socius: Sociological Research for a Dynamic World*, 6, 237802312097150. <https://doi.org/10.1177/2378023120971507>
- BBC News. (2020a, June 7). George Floyd: London anti-racism protests leave 27 officers hurt. *BBC News*. <https://www.bbc.com/news/uk-england-london-52954899>
- BBC News. (2020b, July 16). George Floyd: What happened in the final moments of his life. *BBC News*. <https://www.bbc.com/news/world-us-canada-52861726>
- Bell, M. C. (2017). Police reform and the dismantling of legal estrangement. *The Yale Law Journal*, 126(7), 2054–2150.
- Bennell, C., Alpert, G., Andersen, J. P., Arpaia, J., Huhta, J., Kahn, K. B., Khanizadeh, A., McCarthy, M., McLean, K., Mitchell, R. J., Nieuwenhuys, A., Palmer, A., & White, M. D. (2021). Advancing police use of force research and

- practice: Urgent issues and prospects. *Legal and Criminological Psychology*, 26(2), 121–144. <https://doi.org/10.1111/lcrp.12191>
- Bor, J., Moscoe, E., Mutevedzi, P., Newell, M. -L., & Bärnighausen, T. (2014). Regression discontinuity designs in epidemiology: Causal inference without randomized trials. *Epidemiology*, 25(5), 729–737. <https://doi.org/10.1097/EDE.0000000000000138>
- Bradford, B. (2011). Convergence, Not Divergence?: Trends and trajectories in public contact and confidence in the police. *The British Journal of Criminology*, 51(1), 179–200. <https://doi.org/10.1093/bjc/azq078>
- Bradford, B., Jackson, J., & Stanko, E. A. (2009). Contact and confidence: Revisiting the impact of public encounters with the police. *Policing and Society*, 19(1), 20–46. <https://doi.org/10.1080/10439460802457594>
- Braga, A. A., Brunson, R. K., & Drakulich, K. M. (2019). Race, place, and effective policing. *Annual Review of Sociology*, 45(1), 535–555. <https://doi.org/10.1146/annurev-soc-073018-022541>
- Burgess, R. A., Kanu, N., Matthews, T., Mukotekwa, O., Smith-Gul, A., Yusuf, I., Lamptey, I., Macauley, N., Wilson, R., Pirisola, M., & Gul, M. (2021). Exploring experiences and impact of the COVID-19 pandemic on young racially minoritised people in the United Kingdom: A qualitative study. *PLoS One*, 17(5), e0266504. <https://doi.org/10.1371/journal.pone.0266504>
- Capellan, J. A., Lautenschlager, R., & Silva, J. R. (2020). Deconstructing the Ferguson effect: A multilevel mediation analysis of public scrutiny, de-policing, and crime. *Journal of Crime and Justice*, 43(2), 125–144. <https://doi.org/10.1080/0735648X.2019.1652921>
- Carney, N. (2016). All lives matter, but so does race: Black lives matter and the evolving role of social media. *Humanity & Society*, 40(2), 180–199. <https://doi.org/10.1177/0160597616643868>
- Choi, J. (2021). Asymmetry in media effects on perceptions of police: An analysis using a within-subjects design experiment. *Police Practice & Research*, 22(1), 557–573. <https://doi.org/10.1080/15614263.2020.1749624>
- Dimmery, D. (2016). <https://cran.r-project.org/web/packages/rdd/index.html>
- Dunivin, Z. O., Yan, H. Y., Ince, J., & Rojas, F. (2022). Black lives matter protests shift public discourse. *Proceedings of the National Academy of Sciences*, 119(10), e2117320119. <https://doi.org/10.1073/pnas.2117320119>
- Edrington, C. L., & Lee, N. (2018). Tweeting a social movement: Black lives matter and its use of twitter to share information, build community, and promote action. *Journal of Public Interest Communications*, 2(2), 289. <https://doi.org/10.32473/jpic.v2.i2.p289>
- Engel, R. S., Corsaro, N., Isaza, G. T., & McManus, H. D. (2022). Assessing the impact of de-escalation training on police behavior: Reducing police use of force in the Louisville, KY metro police department. *Criminology & Public Policy*, 21(2), 199–233. <https://doi.org/10.1111/1745-9133.12574>
- Gangl, M. (2010). Causal Inference in Sociological Research. *Annual Review of Sociology*, 36(1), 21–47. <https://doi.org/10.1146/annurev.soc.012809.102702>
- Ghaemmaghani, A., Inkpen, R., Charman, S., Ilett, C., Bennett, S., Smith, P., & Newiss, G. (2022). Responding to the public during a pandemic: perceptions of ‘satisfactory’ and ‘unsatisfactory’ policing. *Policing: A Journal of Policy and Practice*, 15(4), 2310–2328. <https://doi.org/10.1093/police/paab058>
- Giorgi, S., Guntuku, S. C., Himelein-Wachowiak, M., Kwarteng, A., Hwang, S., Rahman, M., & Curtis, B. (2022). *Twitter corpus of the #blacklivesmatter movement and counter protests: 2013 to 2021*. arXiv:2009.00596. arXiv. <https://doi.org/10.1609/icwsm.v16i1.19373>
- Goldsmith, A., & McLaughlin, E. (2022). Policing’s new vulnerability re-envisioning local accountability in an era of global outrage. *The British Journal of Criminology*, 62(3), 716–733. <https://doi.org/10.1093/bjc/azab073>
- Graziano, L. M. (2019). News media and perceptions of police: A state-of-the-art-review. *Policing: An International Journal*, 42(2), 209–225. <https://doi.org/10.1108/PIJPSM-11-2017-0134>
- Graziano, L. M., Schuck, A., & Martin, C. (2010). Police misconduct, media coverage, and public perceptions of racial profiling: An experiment. *Justice Quarterly*, 27(1), 52–76. <https://doi.org/10.1080/07418820902763046>
- Harris, S., Joseph-Salisbury, R., Williams, P., & White, L. (2021). *A collision of crises: Racism, policing, and the COVID-19 pandemic*. Runnymede Trust, Centre on the Dynamics of Ethnicity. May. <https://www.runnymede-trust.org/publications/collision-of-crises-racism-policing-and-the-covid-19-pandemic>
- Harris, S., Joseph-Salisbury, R., Williams, P., & White, L. (2022). Notes on policing, racism and the Covid-19 pandemic in the UK. *Race & Class*, 63(3), 92–102. <https://doi.org/10.1177/03063968211063436>
- Hirschfield, P. J. (2023). Exceptionally Lethal: American police killings in a comparative perspective. *Annual Review of Criminology*, 6(1). [annurev-criminol-030421-040247. https://doi.org/10.1146/annurev-criminol-030421-040247.](https://doi.org/10.1146/annurev-criminol-030421-040247)
- Imbens, G., & Kalyanaraman, K. (2012). Optimal bandwidth choice for the regression discontinuity estimator. *The Review of Economic Studies*, 79(3), 933–959. <https://doi.org/10.1093/restud/rdr043>
- Intravia, J., Thompson, A. J., & Pickett, J. T. (2020). Net legitimacy: Internet and social media exposure and attitudes toward the police. *Sociological Spectrum*, 40(1), 58–80. <https://doi.org/10.1080/02732173.2020.1720554>
- Jennings, J. T., & Rubado, M. E. (2017). Preventing the use of deadly force: The relationship between police agency policies and rates of officer-involved gun deaths. *Public Administration Review*, 77(2), 217–226. <https://doi.org/10.1111/puar.12738>

- Joseph-Salisbury, R., Connelly, L., & Wangari-Jones, P. (2020). "The UK is not innocent": Black lives matter, policing and abolition in the UK. *Equality, Diversity and Inclusion: An International Journal*, 40(1), 21–28. <https://doi.org/10.1108/EDI-06-2020-0170>
- Kajepeta, S., Bruzelius, E., Ho, J. Z., & Prins, S. J. (2022). Policing the pandemic: Estimating spatial and racialized inequities in New York City police enforcement of COVID-19 mandates. *Critical Public Health*, 32(1), 56–67. <https://doi.org/10.1080/09581596.2021.1987387>
- Katikireddi, S. V., Lal, S., Carrol, E. D., Niedzwiedz, C. L., Khunti, K., Dundas, R., Diderichsen, F., & Barr, B. (2021). Unequal impact of the COVID-19 crisis on minority ethnic groups: A framework for understanding and addressing inequalities. *Journal of Epidemiology and Community Health*, 75(10), 970–974. <https://doi.org/10.1136/jech-2020-216061>
- Kochel, T. R. (2019). Explaining racial differences in ferguson's impact on local residents' trust and perceived legitimacy: Policy implications for police. *Criminal Justice Policy Review*, 30(3), 374–405. <https://doi.org/10.1177/0887403416684923>
- Labrecque, J., & Swanson, S. A. (2018). Understanding the assumptions underlying instrumental variable analyses: A brief review of falsification strategies and related tools. *Current Epidemiology Reports*, 5(3), 214–220. <https://doi.org/10.1007/s40471-018-0152-1>
- Laniyonu, A. (2021). Phantom pains: The effect of police killings of black Americans on black British attitudes. *British Journal of Political Science*, 52(4), 1–17. <https://doi.org/10.1017/S0007123421000557>
- Lee, T. (2002). *Mobilizing public opinion: Black insurgency and racial attitudes in the civil rights era*. University of Chicago Press.
- Legewie, J. (2013). Terrorist events and attitudes toward Immigrants: A natural experiment. *The American Journal of Sociology*, 118(5), 1199–1245. <https://doi.org/10.1086/669605>
- Lösel, F. (2018). Evidence comes by replication, but needs differentiation: The reproducibility issue in science and its relevance for criminology. *Journal of Experimental Criminology*, 14(3), 257–278. <https://doi.org/10.1007/s11292-017-9297-z>
- Lyle, P., & Esmail, A. M. (2016). Sworn to protect: Police brutality – a dilemma for america's police. *Race, Gender & Class*, 23(3–4), 155–185.
- Mazerolle, L., Antrobus, E., Bennett, S., & Tyler, T. R. (2013). Shaping citizen perceptions of police legitimacy: A randomized field trial of procedural justice: Shaping citizen perceptions of police. *Criminology*, 51(1), 33–63. <https://doi.org/10.1111/j.1745-9125.2012.00289.x>
- McNeeley, S., & Warner, J. J. (2015). Replication in criminology: A necessary practice. *European Journal of Criminology*, 12(5), 581–597. <https://doi.org/10.1177/1477370815578197>
- Millings, M. (2013). Policing British Asian Identities: The enduring role of the police in Young British Asian Men's situated negotiation of identity and belonging. *The British Journal of Criminology*, 53(6), 1075–1092. <https://doi.org/10.1093/bjc/azt041>
- Muñoz, J., Falcó-Gimeno, A., & Hernández, E. (2020). Unexpected event during survey design: Promise and pitfalls for causal inference. *Political Analysis*, 28(2), 186–206. <https://doi.org/10.1017/pan.2019.27>
- Murray, M. P. (2006). Avoiding invalid instruments and coping with weak instruments. *Journal of Economic Perspectives*, 20(4), 111–132. <https://doi.org/10.1257/jep.20.4.111>
- Nägel, C., & Lutter, M. (2021). The 2017 French riots and trust in the police: A quasi-experimental approach. *European Journal of Criminology*, 20(1), 22. <https://doi.org/10.1177/1477370821998974>
- Nägel, C., & Nivette, A. E. (2022). Unexpected events during survey design and trust in the police: A systematic review. *Journal of Experimental Criminology*. <https://doi.org/10.1007/s11292-022-09508-y>
- Narayan, J. (2019). British black power: The anti-imperialism of political blackness and the problem of nativist socialism. *The Sociological Review*, 67(5), 945–967. <https://doi.org/10.1177/0038026119845550>
- Patterson, T. R., & Kelley, R. D. G. (2000). Unfinished migrations: Reflections on the African diaspora and the making of the modern world. *African Studies Review*, 43(1), 11. <https://doi.org/10.2307/524719>
- Perry, G., Jonathan-Zamir, T., & Factor, R. (2021). The long-term effects of policing the COVID-19 pandemic: Public attitudes toward the police in the 'new normal'. *Policing: A Journal of Policy and Practice*, 14(4), 1073–1091. <https://doi.org/10.1093/police/paab070>
- Perugini, M., Gallucci, M., & Costantini, G. (2018). A practical primer to power analysis for simple experimental designs. *International Review of Social Psychology*, 31(1), 20. <https://doi.org/10.5334/irsp.181>
- Reisig, M. D., Mays, R. D., & Telep, C. W. (2018). The effects of procedural injustice during police–citizen encounters: A factorial vignette study. *Journal of Experimental Criminology*, 14(1), 49–58. <https://doi.org/10.1007/s11292-017-9307-1>
- Reny, T. T., & Newman, B. J. (2021). The opinion-mobilizing effect of social protest against police violence: Evidence from the 2020 George Floyd Protests. *The American Political Science Review*, 115(4), 1–9. <https://doi.org/10.1017/S0003055421000460>
- Sibley, C. G., Greaves, L. M., Satherley, N., Wilson, M. S., Overall, N. C., Lee, C. H. J., Milojev, P., Bulbulia, J., Osborne, D., Milfont, T. L., Houkamau, C. A., Duck, I. M., Vickers-Jones, R., & Barlow, F. K. (2020). Effects of the

- COVID-19 pandemic and nationwide lockdown on trust, attitudes toward government, and well-being. *The American Psychologist*, 75(5), 618–630. <https://doi.org/10.1037/amp0000662>
- Smoke, B. C. (2020, June 9). Police accused of unlawful tactics at London BLM Protests. *Vice News*. <https://www.vice.com/en/article/pkyznz/police-accused-of-unlawful-tactics-at-london-blm-protests>
- Solomon, S. J., & Chenane, J. L. (2021). Testing the fair process heuristic in a traffic stop context: Evidence from a factorial study with video vignettes. *The British Journal of Criminology*, 61(4), 1086–1106. <https://doi.org/10.1093/bjc/azaa096>
- Stock, J. H., & Watson, M. W. (2007). *Introduction to econometrics* (2. ed., Pearson international ed.). Pearson Addison-Wesley.
- Stommes, D., Aronow, P. M., & Sävje, F. (2021). On the reliability of published findings using the regression discontinuity design in political science. *ArXiv:2109.14526 [Stat]*. <http://arxiv.org/abs/2109.14526>
- Sunshine, J., & Tyler, T. R. (2003). The role of procedural justice and legitimacy in shaping public support for policing. *Law Society Review*, 37(3), 513–548. <https://doi.org/10.1111/1540-5893.3703002>
- The Economist. (2020, June 9). George Floyd is remembered around the world. *The Economist*. <https://www.economist.com/graphic-detail/2020/06/09/george-floyd-is-remembered-around-the-world>
- Tyler, T. R., & Huo, Y. J. (2002). *Trust in the Law: Encouraging public cooperation with the police and courts*. Russell Sage Foundation.
- van Haperen, S., Uitermark, J., & Nicholls, W. (2022). The Swarm versus the grassroots: Places and networks of supporters and opponents of black lives matter on twitter. *Social Movement Studies*, 22(2), 1–19. <https://doi.org/10.1080/14742837.2022.2031954>
- Vomfell, L., & Stewart, N. (2021). Officer bias, over-patrolling and ethnic disparities in stop and search. *Nature Human Behaviour*, 5(5), 566–575. <https://doi.org/10.1038/s41562-020-01029-w>
- Waters, R. (2014). African Canadian anti-discrimination activism and the transnational civil rights movement, 1945–1965. *Journal of the Canadian Historical Association*, 24(2), 386–424. <https://doi.org/10.7202/1025083ar>
- Weitzer, R. (2002). Incidents of police misconduct and public opinion. *Journal of Criminal Justice*, 30(5), 397–408. [https://doi.org/10.1016/S0047-2352\(02\)00150-2](https://doi.org/10.1016/S0047-2352(02)00150-2)
- Wortley, S., & Owusu-Bempah, A. (2011). The usual suspects: Police stop and search practices in Canada. *Policing and Society*, 21(4), 395–407. <https://doi.org/10.1080/10439463.2011.610198>