Consistency of Supervisory Interpretations of Stop-Search Justification in London: A Vignette Assessment Analysis

Andy Brittain[1] · Christopher Sims[2] · Vincent Harinam[2] · Heather Strang[2]

Abstract

Research Question Do police supervisors reach different conclusions about the legality of a decision to stop and search in survey vignettes using similar facts with the ethnicity of subjects described as either black or white?

Data We utilize a vignette survey design, presenting 15 real-world stop and search examples from within the South Basic Command Unit policing area to 118 frontline uniformed supervisors (Sergeants and Inspectors) in the Metropolitan Police Service of London.

Methods We introduce a randomised characteristic assignment of the ethnicity of the subject featured in the vignette to compare officer decision-making when the suspect is black or white. Using both Likert scale and free-text responses, a combination of descriptive statistics, inferential methods, and text mining is applied to the survey data.

Findings We found no substantive difference in the justification of stop and search powers by officers between white suspects and black suspects. We did find substantial variability in supervisor assessments of whether vignettes provided sufficient legal grounds for conducting a stop and search.

Conclusions Within the limits of the methodology, we conclude that there is no racial disparity in perceptions of legal thresholds for conducting searches across a range of circumstances. We must also conclude that there is a substantial range of opinion regarding different specific circumstances that are understood to provide a threshold of sufficient evidence to search.

Keywords Stop & search · Police supervision · Racial disparity in policing · Police legal powers

✉ Heather Strang
  hs404@cam.ac.uk

1   Metropolitan Police Service, London, UK

2   University of Cambridge, Cambridge, UK

🙌 Springer

Introduction

Police forces in the UK, and in particular the Metropolitan Police Service (MPS), are facing an acute crisis related to the dissipation of public trust and confidence. More generally, a catalogue of actions by police within the UK and across the world have created a challenging environment for police officers to operate within. The death of George Floyd in May 2020 in the USA and the murder of Sarah Everard in March 2021 in the UK are two such watershed moments that have hurt public confidence in law enforcement. In particular, Sarah Everard's death sparked significant public discussions regarding both violence against women and girls and the actions of law enforcement personnel in reducing these cases. Moreover, the culture of the MPS was called into question with the investigation in 2022 by the Independent Office for Police Conduct, Operation Hotton (Independent Office for Police Conduct (IOPC), 2022) of officers from the MPS Charing Cross Police Station that uncovered evidence of discrimination, misogyny, harassment, and bullying.

More recently, the MPS has been subject to a review into the standards of behaviour and internal culture, led by Baroness Casey. The contents of the report found the MPS to be institutionally racist, homophobic, and sexist (Casey, 2023). Importantly, the report makes specific mention of the MPS' policing of London's black community, stating "the black community are more likely to be stopped and searched, handcuffed, batoned and tasered, leading to generational mistrust within this particular community." More generally, the report highlights the connection between stop and search and legitimacy in policing.

Similarly, a report by the Children's Commissioner Dame Rachel de Souza (BBC News, 2023) raises concerns in relation to the strip searching of children, with 2847 children strip searched in England and Wales between 2018 and mid-2022. This report found that black children were six times more likely to be searched when compared with the overall child population.

The use of stop and search remains an enduring issue in the UK. This is particularly due to ongoing concerns about racial profiling and discrimination. Importantly, stop and search remains a primary tactic in the Metropolitan Police Service's plan to tackle violent crime in London. Within the South Basic Command Unit (SN BCU), the power to stop and search was used on 22,505 occasions in the calendar year of 2021. This creates a significant resourcing challenge due to volume of the policing interactions that not only consume a considerable amount of operational time but may also impact local trust and confidence in policing if not undertaken both lawfully and proportionately.

This study examines the decision-making of police supervisors in relation to stop and search to identify the presence or absence of discriminatory behaviour among these key operational officers. More specifically, we utilize a vignette survey design, presenting 15 real-world stop and search examples from within the South Basic Command Unit policing area to 118 frontline uniformed supervisors (Sergeants and Inspectors) in the MPS. These examples were obtained from the MPS' criminal intelligence database (CRIMINT). Importantly, we introduce a

randomised characteristic assignment with relation to the ethnicity of the subject featured in the vignette to test the impact on officer decision-making when the suspect is black or white. Given the collection of both Likert scale and free text responses, a combination of descriptive statistics, inferential methods, and text mining is utilized. More specially, we employ summary statistics, t tests, descriptive text analysis, and sentiment analysis.

Research Question

This research seeks to identify variability in officer decision-making and looks to identify key characteristics that may have an impact on that decision-making. Moreover, this research will help to shape and develop the following key areas that are critical to policing: support the development of bespoke training in relation to stop and search, promote organisational inclusivity, build a more representative police force, and increase trust and confidence in policing. More generally, the results of this study will help to inform MPS policy on the use and capacity of stop and search as a crime reduction and intelligence gathering tool. Indeed, if the results suggest that officers are indeed relying on stereotypes and categories when making decisions about stop and search then additional bespoke training may be necessary to help officers recognise and overcome these conscious and unconscious biases. Additionally, policy changes could be implemented to ensure that stop and search procedures are conducted fairly and without discrimination.

Prior Research

While many police officers consider stop and search to be a necessary tactic to keep communities safe, others believe that it undermines the legitimacy of the police. As a result, this remains subject to intense scrutiny from the media, politicians, and the wider communities. Furthermore, while this paper will not comment on the effectiveness of stop and search, it is important to note that prior research (Miller et al., 2000) indicated that this tactic does not have a significant impact on overall crime rates. Nevertheless, stop and search does contribute to arrests for crimes that are detectable by searches and serves as a key intelligence gathering tool. There is also a notable difference in how police organizations employ the tactic, with some forces relying more heavily than others on stop and searches to make arrests.

Legitimacy in Policing

Police legitimacy is a concept that is well established in the UK and fundamental to "policing by consent", which is crucial in a democratic society (Her Majesty's Inspectorate of Constabulary and Fire and Rescue Services (HMICFRS), 2017). Research to identify the drivers of police legitimacy have identified four key building blocks: procedural justice, distributive justice, effectiveness, and lawfulness (Bottoms & Tankebe, 2017). The use of stop and search touches on all four of these

aspects. Procedural justice, defined as "the fairness of the process employed to reach specific outcomes or decisions", includes the quality of both decision-making and the treatment of citizens (Bottoms & Tankebe, 2017). To achieve procedural justice, police organizations need to ensure that they recognize the rights of the individual, along with acknowledgement of their humanity and in doing so that individual feels valued. The second building block, distributive justice, emphasizes the importance of how individuals are treated by the police. This includes the over and/or under-policing of specific communities or groups of people. The third and fourth building blocks, effectiveness and lawfulness, relate to the public's expectation of efficacious policing and the requirement of officers to act within legislative boundaries, respectively.

In a study examining the role of procedural justice and legitimacy in shaping public support for policing, Sunshine and Tyler (2003) utilised two surveys of New Yorkers to gain an understanding of how they judged the legitimacy of policing. The findings suggest that the fairness of the practices and tactics used by the police is the key driver of legitimacy. Interestingly, both white and non-white respondents indicated that public perceptions of fair police practices affected overall views of police legitimacy. A recent study (Murray et al., 2021) conducted in several cities across England and Wales found that stop and search not only eroded trust among adolescents but also resulted in increased criminal activity in specific cases.

Effects of Stop and Search

Extensive research has been conducted on both the effects and disparities of stop and search on both individuals and communities. Research utilising data from 'Stop Watch' and qualitative interviews with children searched demonstrated that the use of this power on children can result in severe emotional distress for the individuals involved and contribute to their engagement in criminal activities later on in life (Flacks, 2018). Another study (Hargreaves, 2018) examining the use of stop and search on the British Muslim community found statistical data to support claims that this practice was being applied in a disproportionate manner, particularly when compared to white Christian groups.

Legal Framework

In conducting research on stop and search, it is crucial to consider the factors that influence the decision-making of officers as well as the cultural and environmental context in which they work. The legal framework for stop and search requires officers to have a reasonable suspicion to utilise their legislative powers. However, "suspicion" is subjective and is affected by officers' own judgement and operational context. While officers are trained in the legislation, many are unable to describe what reasonable suspicion meant in practice, with clear differences between officers on what would amount to reasonable suspicion (Quinton, 2011). This research found that when a Sect. 60 Criminal Justice Act authority was implemented which did not require reasonable suspicion (because a senior officer believed serious violence may take place in this location),

an officer's motivation to search became less focused on finding a weapon and more "reactive to the individuals' behaviour towards officers". This included threatening or abusive behaviour towards the officers or challenges to the officer's authority. When considering operational settings, there were clear differences in police practice based upon the locality the officer operated in. For example, some police areas in London had a distinct performance culture while others were more focused on community policing. Quinton (2011) concluded that decisions to initiate encounters were based on broad generalisations and stereotypes based on memberships of defined social categories.

Cultural Dimensions

Police culture is comprised of two distinct working environments. There is an organisational environment which consists of interactions with supervisors and senior officers, and an operational environment that includes interactions with the public and is rife with the constant risk of physical danger. These two working environments are said to create considerable stress and anxiety for officers leading to a somewhat unique police culture (Terrill et al., 2003). Cockcroft (2007) highlights three potential descriptions of police culture:

1. A layer of informal occupational norms and values operating under the apparently rigid hierarchical structure of police organisations.
2. Accepted practices, rules, and principles of conduct that are situationally applied, and generalized rationales and beliefs.
3. A patterned set of understandings which help to cope with and adjust to the pressures and tensions which confront the police.

The extent to which officers adopt this culture will be evaluated through a vignette-based study that will measure their willingness to tolerate varying levels of suspicious circumstances. Using these vignettes, we seek to gain an understanding of the police culture within an MPS BCU, which will support the organisation with the development of tailored and specific awareness training.

Despite a significant body of research on the effects of stop and search on local communities, there has been limited critical analysis of the actual decision-making process of officers in relation to specific community groups. This research will contribute to the existing body of knowledge on stop and search and provide an evidence base for consideration by senior police leaders within the MPS. The research provides insights into the decision-making processes of police supervisors when conducting stop and search on members of the black community and recommend ways to improve the use of this policing tactic.

Data

The location of this study is the MPS South Basic Command Unit (BCU), which consists of the London Boroughs of Sutton, Croydon, and Bromley. This the fourth largest BCU in the MPS by officer count, with a population of just under one million, 52% of whom are members of minority communities (Office of National Statistics, 2021). In 2022, the South BCU undertook 16,950 stop and searches, which represented 9% of all stop and searches in the MPS. Importantly, drugs, weapons, and stolen property constituted 60.3%, 17.3%, and 13.5% of reasons for stop and searches, respectively. Furthermore, 72.3% of all stop and searches resulted in no further action being taken, with only 13.7% resulting in an arrest.

Study Design

This study utilised an online survey tool to provide vignette stories to test the variability in decision-making by supervisory sergeants and inspectors operating in the South BCU. The vignette approach was chosen as it offers a standardised approach to survey questions ensuring all participants are given the same situation and information thereby reducing variation in responses. These surveys involve presenting participants with a hypothetical scenario (vignette) which describes a situation or event, and asks the participant to provide their opinions, attitudes, or behaviours based on that set of circumstances. In short, vignettes also provide a degree of operational realism, making it easier for the participants to relate to the situation presented and make decisions in an environment that they are comfortable operating within. This serves to improve the overall quality of the data obtained. The use of vignettes also allows researchers a greater control of the variables being studied. In effect, this approach allows for the controlled isolation of specific factors and their corresponding impact on participant responses. In this case, we used black or white suspects in the same vignette description.

Focus on Supervisors

The data were generated from frontline supervisory ranks of sergeants and inspectors. Given the low average age and length of service of police constables in this police force (estimated to be under 2 or 3 years), we decided to focus on supervisors. The officers sought for the surveys had encountered many more situations that could justify stop and search than probationary and recent recruits, and would have made more decisions as to whether a stop was appropriate. Moreover, their responses indicate the content of what they might train younger colleagues to look for in deciding whether a stop meets the required legal thresholds.

Vignettes

To create an operational reality in the vignettes, the MPS Stop and Search database was searched to identify suitable real incidents to test the supervisors with. All

incidents consisted of searches that had resulted in a negative outcome, that lacked strong recorded grounds, and had been supervised by a sergeant or inspector. These incidents were redacted to ensure no personal data remained and then incorporated into the online survey tool. It should be noted that the vignettes are relatively simplistic in design. However, that reflects the quality of the information provided by operational officers when completing a record of the stop and search. It was also felt that the vignettes needed to be simple to ensure that the survey could be completed within a reasonable timeframe without survey fatigue affecting the outcome of the respondents' answers.

By ensuring respondent anonymity through an online survey tool, we believe the study encouraged responses that were open and honest reflections of the officers' thought processes. To this extent, the use of real-world stop and search scenarios ensured that the vignette was recognised by respondents as a real situation as opposed to a contrived one. The only factual difference in the individual vignettes was the random assignment of either a black or white suspect within the vignette itself, ensuring everyone was assessed against the same circumstances.

Including questions on officer's rank, years of service, sex, and ethnicity, the survey examined officer decision-making by having them explain their utilisation (or lack thereof) of stop and search powers based on the information available within 15 vignettes (see Appendix 1 for vignette list). Their decision-making was scored though a scale of -5 to $+5$, with -5 indicating an unwillingness to search against a score of $+5$ indicating a strong willingness to search. The survey also asked the participants to provide written feedback articulating their thought processes in free text boxes. Crucially, the online survey tool was comprised of fifteen selected vignettes and programmed to randomly change the ethnicity of the suspect (e.g. black or white). Officers would not receive the same vignette twice with different ethnicities.

Participant Selection

Given this study's explicit focus on testing the decision-making of police supervisors who oversee officers undertaking stop and searches, lists of every supervisory officer with the rank of sergeant and inspector within the South BCU were generated with the aid of the MPS human resources department. Of the 154 eligible officers, 118 participated in this study. The work emails for each officer where then entered on the online survey tool which then provided a randomised candidate number for every officer. The survey tool then tracked the officers' returns and allowed for follow-up emails to be sent to those that failed to respond. At no point in this study did the survey tool allow a direct link to be drawn between individual officer and their generated candidate number, ensuring absolute anonymity. Importantly, all officers were guaranteed of their anonymity and told that completion of the survey was optional.

Methods

We leverage a combination of descriptive statistics, inferential methods, and text mining. We use both Microsoft Excel and R to conduct these analyses. As it relates to descriptive statistics, we employ simple summary statistics to determine the demographical characteristics of the officers involved in the study. We utilize t-tests to measure the Likert scale responses (-5 to $+5$) for each vignette, comparing responses when the suspect was black against the responses when the suspect was white.

Text Mining

Text mining is *a* relatively new technological development and is defined as the act of "processing a collection of documents, or corpus, in which documents are converted into structured data, such that each document is described using a set of features called concepts to provide a holistic perspective of textual and non-textual information" (Mikroyannidis & Theodoulidis, 2006, 45). More generally, text mining allows for the automatic analysis of large amounts of qualitative data. Given that this study collected free text responses for each vignette, traditional qualitative research approaches such as content analysis, while appropriate, were deemed inadequate. Text mining represented the most viable methodological option as we sought to measure the frequency of word use and the sentiment of officer responses against the ethnicity of the suspect in the vignette.

Text mining necessitates a series of pre-processing procedures. For this study, data pre-processing consisted of tokenization, filtering, and stemming. More specifically, the textual data were cleaned by removing punctuation, special characters, digits and uniform resource locator links. Tokenization, the process of reducing words into pieces of information called tokens, was conducted in order to identify meaningful keywords. Next, all stop-words were removed from the corpus. These are words such as 'and' or 'the' that do not carry information. Next, word stemming is conducted—a process of transforming words into their roots. All these data were then converted to a corpus which is a large, structured set of texts. From here, these data were converted into a structured format from which analyses can be conducted. Finally, a vectorspace model was used to capture the relevant features for each document within the data.

Text Analysis

We leverage descriptive text analysis to identify the frequency and distribution of words used by the respondents. We also calculate the power law distribution of word use using a cumulative distribution function. It should be noted that there are limitations to this analysis. Indeed, descriptive text analysis may

not necessarily consider the context in which specific words are used. This may lead to the misinterpretation or oversimplification of the significance of specific words.

Sentiment Analysis

We utilize sentiment analysis to gauge the tone of officers' responses when the suspect in the vignette is black or white. Sentiment analysis is the practice of applying natural language processing (NLP) and text analysis to identify subjective information within written text (Hussein, 2018). The main goal of sentiment analysis is to automatically classify the sentiment of a piece of text into positive, negative, or neutral classification. To achieve this, NLP-based algorithms use various techniques such as lexicon-based analysis. Lexicon-based analysis involves assigning scores to words based on their positive or negative connotations. We use the "qdap" package in R to provide polarity scores ranging from -1, indicating negative sentiment, to $+1$, indicating positive sentiment. Importantly, sentiment analysis possesses two broad weaknesses. First, it may fail to recognize the ambiguity of language as some words can having different meanings and connotations depending on the tone, cultural background, and context in which it is used. Second, it may fail to account for sarcasm and irony which can produce incorrect scores.

Findings

This study has a response rate of 77%, with 118 of the 154 supervisory officers completing the survey. Table 1 presents a breakdown of key characteristics for the respondents. As it apparent, 67.8% (80) of officers were male, with 81.4% of all respondents identifying as British (white). Furthermore, 76.3% (90), were police sergeants, with 40.7% between the ages of 40 and 49,.

Scaled Response Results

Table 2 presents the t-test results for each vignette, comparing officers' responses when suspects were black or white. It is important to note that a small p value does not necessarily imply that the differences in officers' decision to stop and search black or white suspects is large or meaningful, or that the alternative hypothesis is true. It simply means that the data offered strong evidence against a null hypothesis. Additionally, the interpretation of p-values should always be made in the context of the study design, the sample size, and the potential sources of bias. As noted above, the mean score averages individual responses ranging from -5 (very *unlikely* to justify a stop-and-search to $+5$ (highly *likely* likely to justify a search).

Based on the results of the t-test, it appears that officers are, in general, neither more or less likely to stop and search black or white suspects. Indeed, the race of the suspect is seemingly inconsequential in an officer's decision to use their power to stop and search. This can be gleaned from the mean scores for black and whites

Table 1 Characteristics of respondents

	N	%
Gender		
Female	24	20.3%
Male	80	67.8%
Prefer not to say	14	11.9%
Ethnicity		
Other Asian background	1	0.8%
Other Ethnic Group	2	1.7%
Mixed background	2	1.7%
British (White)	96	81.4%
Caribbean	1	0.8%
Prefer not to say	11	9.3%
Other	4	3.4%
White and Asian	1	0.8%
Rank		
A/inspector	3	2.5%
Inspector	25	21.2%
Police sergeant	90	76.3%
Age group		
21–29	2	1.7%
30–39	37	31.4%
40–49	48	40.7%
50 or older	20	16.9%
Prefer not to say	11	9.3%
Years of service		
0–5 years	1	0.8%
6–10 years	19	16.1%
11–15 years	13	11.0%
16–20 years	46	39.0%
21–25 years	26	22.0%
26 and over	13	11.0%
BCU		
Bromley	35	29.7%
Croydon	40	33.9%
Pan BCU	20	16.9%
Sutton	23	19.5%
Role		
Emergency response policing	52	44.1%
Headquarters	11	9.3%
Local investigation	5	4.2%
Neighbourhood policing	40	33.9%
Other	8	6.8%
Public protection	2	1.7%

Table 2 *T* test results of all vignettes

Question/race	N	Mean	SD	t	t-crit	DF	p
Question 1	118	0.81	2.94	1.13	1.98	116	0.26
Black	61	0.51	2.91				
White	57	1.12	2.98				
Question 2	118	−2.18	3.05	−0.74	1.98	116	0.46
Black	52	−1.94	3.23				
White	66	−2.36	2.91				
Question 3	118	3.25	2.20	−0.06	1.98	116	0.95
Black	56	3.27	2.28				
White	62	3.24	2.15				
Question 4	118	1.88	2.98	0.46	1.98	116	0.65
Black	64	1.77	2.88				
White	54	2.02	3.12				
Question 5	118	2.26	2.56	1.15	1.98	116	0.25
Black	57	1.98	2.92				
White	61	2.52	2.16				
Question 6	118	2.08	2.76	0.8	1.98	116	0.42
Black	60	1.88	2.91				
White	58	2.29	2.61				
Question 7	118	3.09	2.12	1.47	1.98	116	0.14
Black	63	2.83	2.30				
White	55	3.40	1.87				
Question 8	118	3.52	1.77	0.95	1.98	116	0.34
Black	66	3.38	1.93				
White	52	3.69	1.55				
Question 9	118	1.38	2.84	2.21	1.98	116	**0.03**
Black	50	0.72	3.00				
White	68	1.87	2.63				
Question 10	118	0.38	2.89	1.17	1.98	116	0.25
Black	61	0.08	3.00				
White	57	0.70	2.76				
Question 11	118	−0.62	3.10	1.97	1.98	116	**0.05**
Black	62	−1.15	3.09				
White	56	−0.04	3.03				
Question 12	118	3.47	1.98	1.27	1.98	116	0.21
Black	56	3.23	1.69				
White	62	3.69	2.19				
Question 13	118	1.79	2.71	2.33	1.98	116	**0.02**
Black	54	1.17	2.63				
White	64	2.31	2.68				
Question 14	118	−0.48	2.89	−0.6	1.98	116	0.55
Black	59	−0.32	2.65				
White	59	−0.64	3.13				
Question 15	118	1.53	2.81	−0.66	1.98	116	0.51

Table 2 (continued)

Question/race	N	Mean	SD	t	t-crit	DF	p
Black	58	1.71	2.96				
White	60	1.37	2.67				
All	1770	1.48	3.11	2.34	1.96	1768	**0.02**
Black	879	1.30	3.12				
White	891	1.65	3.09				

across all vignettes. Nevertheless, there are three vignettes where there exists a statistically significant difference between officer responses between black and white suspects. These are vignettes 9, 11, and 12 where black suspects were less likely than white suspects to be stopped and searched by officers. Moreover, the aggregated averages for black and white suspects indicate that white suspects are, overall, more likely than black suspects to be stopped and searched by officers evaluating the vignettes.

Descriptive Text Analysis

Analysis of the free text responses of participants demonstrated a power law distribution, with approximately 1% of words being using 26% of the time (see Fig. 1). Alternatively, 20% of all words were used 80% of the time, abiding by the popular 80/20 principle. For context, there were 2344 unique words and 19,749 total words used by the 118 survey respondents.

Word frequency refers to how often a particular word is used in a language or text. Figures 2a–c present the top 20 words used by respondents for all vignettes, vignettes where the suspect is white, and vignettes where the suspect is black, respectively. Upon examination, the words used across all three groups, particularly
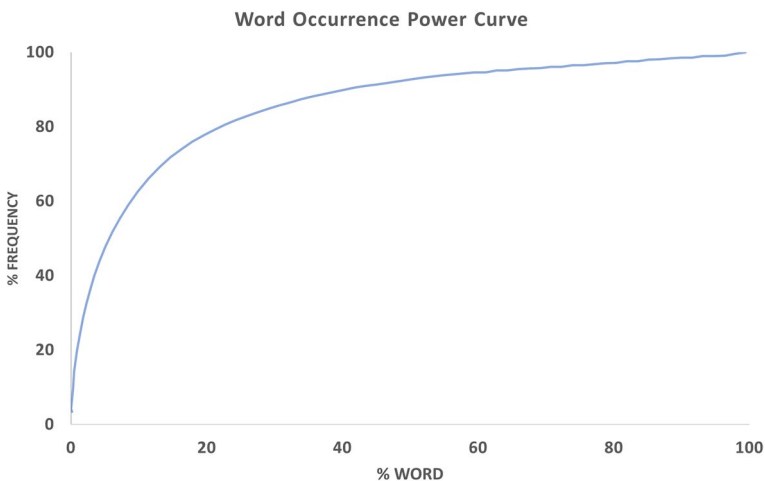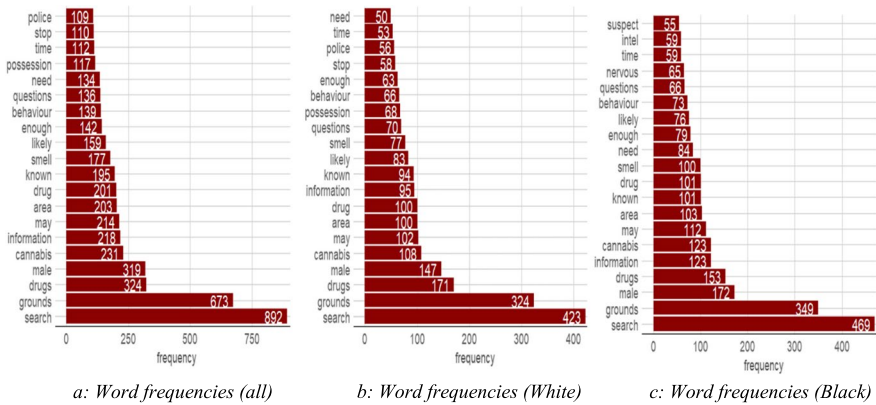

Fig. 1 Word occurrence power curve

a: Word frequencies (all)   b: Word frequencies (White)   c: Word frequencies (Black)

Fig. 2 **a** Word frequencies (all). **b** Word frequencies (White). **c** Word frequencies (Black)

when the suspect is black or white, are the same or similar. Indeed, there is seemingly no substantive difference in the free text responses officers when the suspect is black or white. Moreover, it is clear from these frequently used words that the respondents were searching for more information. This is evidenced by the use of the words: "search", "grounds", "male", "drug", and "information", which are held in common across all three groups. In effect, the respondents were investigating the circumstances within the vignette, wanting to know and gather more information in order to justify the use of stop and search powers. Furthermore, it is also clear that these words are non-malicious in nature and there is little significant difference between the black and white suspect groups. These frequently used words are indicative of professional conduct by the respondents to the survey.

Sentiment Analysis

Table 3 presents the distribution of polarity scores for officer responses for black suspects, white suspects, and all suspects. Based on these results, there is no substantive difference between the average polarity scores of officer responses when the suspect is black or white. This is further substantiated by the closeness of the standard deviations of both groups as the majority of sentiment scores fall close to the mean (see Fig. 3a–c). Importantly, the score of -0.1 for white suspects and -0.11 for black suspects, while negative, is relatively close to 0, indicating value neutrality

Table 3 Distribution of polarity scores

Corpus	Total sentences	Average polarity	Std dev polarity	Std mean polarity
All	1770	-0.10	0.28	-0.364
White	891	-0.10	0.279	-0.342
Black	879	-0.11	0.28	-0.386

a: Sentiment polarity (all)   b: Sentiment polarity (white)   c: Sentiment polarity (black)
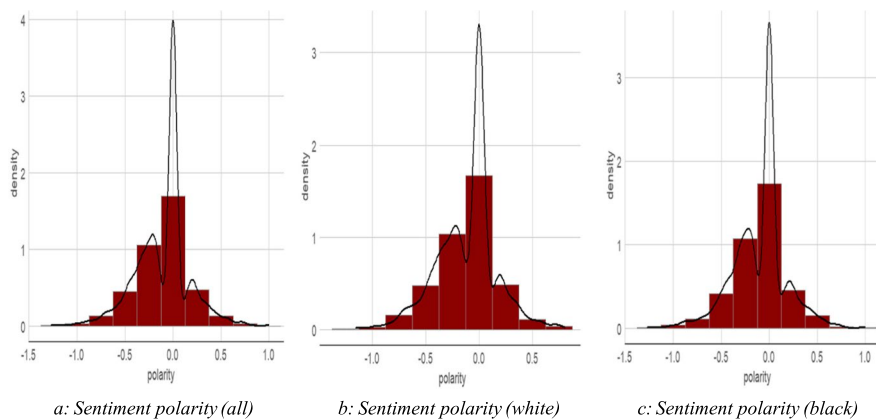
Fig. 3 **a** Sentiment polarity (all). **b** Sentiment polarity (white). **c** Sentiment polarity (black)

where the verbiage of officers was neither positive nor negative. Overall, sentiment analysis corroborates previous finding where the words used by officers convey professional and sober search for evidence to legally justify the decision to stop and search.

Discussion and Conclusion

This study has demonstrated that, given a consistent set of circumstances in relation to stop and search scenarios, operational supervisors may apply different interpretations and weighing to the information provided. This, in turn, leads to variability in both the decision to search and their strength of conviction in relation to the use of that power. Nevertheless, when examined across the *t* tests, descriptive text analysis, and sentiment analysis, this research paints a clear picture regarding the use of stop and search as it relates to the race of the suspect. It finds there is no substantive difference in the justification of stop and search powers between white suspects and black suspects.

In general, the study has identified a lack of disproportionality when comparing decisions made by officers in relation to black and white suspects. While this is a significant outcome which seemingly goes against popular narrative regarding law enforcement's treatment of minorities communities, it may be tempered by a social desirability effect (McCambridge et al., 2014) causing supervisors to select what they thought was the ''right'' answer. We must, therefore, consider these issues and their potential to influence officer decision-making when interpreting these data. Indeed, as the survey was designed by a senior MPS officer, the respondents, all MPS officers, would have known that their responses, while still anonymous, would be closely examined. We must also acknowledge that these officers operate against the backdrop of significant organisational criticism for its treatment of individuals from ethnic minority backgrounds which has resulted in the implementation of policies and processes that hold officers to account when evidence of discriminatory

behaviour is presented. This coupled with a negative police culture of mistrust may have influenced the responses given by officers. However, every effort was made to ensure the survey responses remained anonymous and officers were personally reassured of this throughout the survey period.

It is important to note that sample size is a critical aspect of any survey as it determines the reliability and validity of the results. This survey has focused on the South BCU and, as such, its findings should not be used to inform judgements in relation to the entire MPS. As of February 2023, the MPS has grown to 34,350 officers with 810 inspectors and 3,235 sergeants. The survey originally recruited 154 supervisors which equates to 3.8% of these ranks across the MPS. Moreover, this survey's response rate would equate to 2.9% of all first and second-line supervisors across the entire MPS. Therefore, future research should be run with a much larger sample size, perhaps focused on both Front-Line Policing (FLP) units along with a number of Pan-London operational commands. This approach would include multiple BCUs that police very different areas and communities in London along with the Territorial Support Group (TSG) and Violent Crime Task Force (VCTF) which operate across all 32 London Boroughs.

To further enhance this research, there is an option to utilise virtual reality (VR) technology as this would provide a unique, immersive, and realistic experience for the respondents. The benefit of a VR approach is the creation of real-world scenarios, providing a vivid representation of the vignette simulating the real-world environment, pressures, and behaviours of the suspect. This approach would allow the respondent to interact with measurable elements in a more immersive way. The impact of experiencing the vignette in a real-world environment versus reading a story are significant and should not be underestimated. There are also some considerations to this approach that need to be balanced against the benefits. First, the use of VR requires specialist and costly equipment along with additional production costs in relation to the material used. Second, with any hi-tech equipment there is always the risk of equipment failure leading to respondent frustration and skewed data. Third, the use of VR may create bias if it is not representative of the respondent's physical environment, particularly in terms of the locations used and the suspect's characteristics. Overall, the use of VR in surveys creates an exciting development of this key research type, but it should be used judiciously with due consideration to the limitations.

The decision-making process of police officers is a critical aspect of law enforcement. It is imperative that officers make sound, effective, and lawful decisions to ensure public safety and maintain order. Poor decision-making can lead to erosion of trust and confidence in policing. However, there is often a significant variability in decision-making among police supervisors who, when presented with similar situations, can make very different decisions which can lead to inconsistencies in the application of laws and regulations. While this research demonstrates direct evidence of this variability in decision making it also, and perhaps more importantly in relation to trust and confidence, shows there is no evidence of a negative bias towards members of the black community when compared to the white community.

While the lack of disproportionality is a worthwhile finding, it must be considered against the backdrop of the survey size, the simplistic nature of the vignettes,

and the Hawthorne effect. However, the survey was randomised and anonymised, with officers being personally assured of confidentiality and as such, it is hoped that these findings are representative of their thought processes and do provide some reassurance to the South BCU's local communities. To further test these findings and to provide an even greater understanding, additional research should consider utilising a randomised controlled trial (RCT) approach that not only considers officers response to vignettes in a survey but also cross-compares actual stop and search data along with reviews of stop and search interactions on body worn video. While it is possible that these findings may be at odds with the lived experience of many in specific communities, this research should, nevertheless, encourage other researchers to design future studies that enhance our understanding of police decision-making and stop and search.

Appendix 1

Vignettes

Q	Questions text
1	You see a white male, approx. 18 years of age, on a pedal cycle on the pavement, with hood up and face covered. When stopped by you he immediately puts his hand into his pocket. The area is well known for gang violence and robbery. Please use the scale below with −5 being least likely to search and +5 being most likely to search
2	You see a black male, approx. 18 years of age, on a pedal cycle on the pavement, with hood up and face covered. When stopped by you, he immediately puts his hand into his pocket. The area is well known for gang violence and robbery. Please use the scale below with −5 being least likely to search and +5 being most likely to search
3	You are called to Co-op food shop, to a drunk 30-year-old, white male trying to fight staff. Upon arrival, you see the subject sitting on a bench drinking outside the store, he matches the description of the suspect. Please use the scale below with −5 being least likely to search and +5 being most likely to search
4	You are called to Co-op food shop, to a drunk 30-year-old, black male trying to fight staff. Upon arrival, you see the subject sitting on a bench drinking outside the store, and he matches the description of the suspect. Please use the scale below with −5 being least likely to search and +5 being most likely to search
5	You see a 17-year-old white male walking along the high street. On seeing police, his attention was drawn to the police vehicle so much that it draws your attention, on driving past male checks over his shoulder on several occasions; after a long time, he then turns another time to check on the location of the police vehicle. Upon stopping and speaking to male, he is extremely nervous, stating he does smoke cannabis, and he also smells of cannabis. Please use the scale below with −5 being least likely to search and +5 being most likely to search
6	You see a 17-year-old black male walking along the high street. On seeing police, his attention was drawn to the police vehicle so much that it draws your attention, on driving past male checks over his shoulder on several occasions, after a long time he then turns another time to check on the location of the police vehicle. Upon stopping and speaking to male he is extremely nervous, stating he does smoke cannabis, he also smells of cannabis. Please use the scale below with −5 being least likely to search and +5 being most likely to search
7	You are called to the O'Neill's bar on the High street, in relation to a fight between three males. Suspect descriptions are provided. On your arrival, the subject a 21-year-old white male is seen standing in an aggressive manner squaring up with another male. Officers have to move the subject away. On being moved, a hard object could be seen in the pocket of the subject. Please use the scale below with −5 being least likely to search and +5 being most likely to search

Q	Questions text
8	You are called to the O'Neill's bar on the High street, in relation to a fight between three males. Suspect descriptions are provided. On your arrival, the subject a 21-year-old black male is seen standing in an aggressive manner squaring up with another male. Officers have to move the subject away. On being moved, a hard object could be seen in the pocket of the subject. Please use the scale below with − 5 being least likely to search and + 5 being most likely to search
9	You see a 37-year-old white male walking around the local town centre, he is a well-known class A drug user who frequents the area. When stopped and spoken to, he admits to having a crack pipe in his possession and admitted to having smoked crack cocaine last week. Please use the scale below with − 5 being least likely to search and + 5 being most likely to search
10	You see a 37-year-old black male walking around the local town centre, he is a well-known class A drug user who frequents the area. When stopped and spoken to, he admits to having a crack pipe in his possession and admitted to having smoked crack cocaine last week. Please use the scale below with − 5 being least likely to search and + 5 being most likely to search
11	You are on mobile patrol when you notice a male and a female, both white and approx. 30 years of age, chatting in a doorway late at night. They were stumbling around and appeared under the influence of something. When you are speaking to them, you are able smell cannabis, and they fail to give a reason for their presence at the location. Please use the scale below with − 5 being least likely to search and + 5 being most likely to search
12	You are on mobile patrol when you notice a male and a female, both black and approx. 30 years of age, chatting in a doorway late at night. They were stumbling around and appeared under the influence of something. When you are speaking to them you are able smell cannabis, and they fail to give a reason for their presence at the location. Please use the scale below with − 5 being least likely to search and + 5 being most likely to search
13	You see a white male, approx. 28 years of age driving a motor vehicle, you signal for him to stop. However, he takes a considerable amount of time to stop and is seen to be rummaging in the centre console of his vehicle whilst slowing. The vehicle is pulled over in a drug supply hot spot. The driver is visibly shaking and appears to be nervous. Please use the scale below with − 5 being least likely to search and + 5 being most likely to search
14	You see a black male, approx. 28 years of age driving a motor vehicle, you signal for him to stop. However, he takes a considerable amount of time to stop and is seen to be rummaging in the centre console of his vehicle whilst slowing. The vehicle is pulled over in a drug supply hot spot. The driver is visibly shaking and appears to be nervous. Please use the scale below with − 5 being least likely to search and + 5 being most likely to search
15	A white 18-year-old male, is seen loitering by an alleyway in the company with a male known to carry knives. The area is known for knife crime, recent gang tensions and stabbings. The male appears to be nervous and evasive to simple questions and then became extremely agitated. Please use the scale below with − 5 being least likely to search and + 5 being most likely to search
16	A black 18-year-old male is seen loitering by an alleyway in the company with a male known to carry knives. The area is known for knife crime, recent gang tensions, and stabbings. The male appears to be nervous and evasive to simple questions and then became extremely agitated. Please use the scale below with − 5 being least likely to search and + 5 being most likely to search
17	You see a 16-year-old white male, hanging around a convenience shop, late at night in an area well known for drug use and supply. Subject meets with another male very briefly. When you stop him, he refuses to provide a reason for his presence at the location. Please use the scale below with − 5 being least likely to search and + 5 being most likely to search
18	You see a 16-year-old black male, hanging around a convenience shop, late at night in an area well known for drug use and supply. Subject meets with another male very briefly. When you stop him, he refuses to provide a reason for his presence at the location. Please use the scale below with − 5 being least likely to search and + 5 being most likely to search

Q	Questions text
19	You see a white, 23-year-old male sitting in a secluded area in the grounds of a church, behind a bush on a step. This area is regularly frequented by heroin users. Subject was asked to account for why he was there, he simply replied that he was counting his money. Subject's clothing was untidy and unclean. Please use the scale below with −5 being least likely to search and +5 being most likely to search
20	You see a black, 23-year-old male sitting in a secluded area in the grounds of a church, behind a bush on a step. This area is regularly frequented by heroin users. Subject was asked to account for why he was there, he simply replied that he was counting his money. Subject's clothing was untidy and unclean. Please use the scale below with −5 being least likely to search and +5 being most likely to search
21	You see a 38-year-old, white male standing in the road, apparently trying to beg for money from a driver in a stationary car stopped at the lights. He has the appearance of drug user, his clothes are tatty and he is generally dishevelled, when approached his eyes appear glazed and pupils dilated. He is rambling in a nonsensical way and when asked what he is doing he says he needs money. Please use the scale below with −5 being least likely to search and +5 being most likely to search
22	You see a 38-year-old, black male standing in the road, apparently trying to beg for money from a driver in a stationary car stopped at the lights. He has the appearance of drug user, his clothes are tatty and he is generally dishevelled, when approached his eyes appear glazed and pupils dilated. He is rambling in a nonsensical way and when asked what he is doing he says he needs money. Please use the scale below with −5 being least likely to search and +5 being most likely to search
23	You are patrolling a location that is regularly reported for ASB, namely Cannabis smoking, late into the evenings. You see three (3) young white, teenage females, sitting in the dark in the middle of the children's play area. On approach, there is a strong smell of Cannabis & the females are uncooperative. Please use the scale below with −5 being least likely to search and +5 being most likely to search
24	You are patrolling a location that is regularly reported for ASB, namely Cannabis smoking, late into the evenings. You see three (3) young black, teenage females, sitting in the dark in the middle of the children's play area. On approach, there is a strong smell of Cannabis & the females are uncooperative. Please use the scale below with −5 being least likely to search and +5 being most likely to search
25	Whilst on patrol around robbery hotspots on the local housing estate, based on recent events and current intelligence. You see a white male, approx. 15 years of age. His age and general clothing match the description of the suspects in recent robberies. He is evasive when spoken to. Please use the scale below with −5 being least likely to search and +5 being most likely to search
26	Whilst on patrol around robbery hotspots on the local housing estate, based on recent events and current intelligence. You see a black male, approx. 15 years of age. His age and general clothing match the description of the suspects in recent robberies. He is evasive when spoken to. Please use the scale below with −5 being least likely to search and +5 being most likely to search
27	You see a 20-year-old white male drive at speed through a temporary red light. The area is well known for use of drugs and weapons. You stop the vehicle and speak to the male. His leg is shaking and on speaking to him he appears to be nervous. Please use the scale below with −5 being least likely to search and +5 being most likely to search
28	You see a 20-year-old black male drive at speed through a temporary red light. The area is well known for use of drugs and weapons. You stop the vehicle and speak to the male. His leg is shaking and on speaking to him he appears to be nervous. Please use the scale below with −5 being least likely to search and +5 being most likely to search
29	You see a 28-year-old white male driving a vehicle which is driving erratically along the A23. As you get closer you recognise the driver as a known gang nominal and concerned in the supply of drugs. When the vehicle is stopped the driver is a bit resistant to police and is argumentative. Please use the scale below with −5 being least likely to search and +5 being most likely to search
30	You see a 28-year-old black male driving a vehicle which is driving erratically along the A23. As you get closer, you recognise the driver as a known gang nominal and concerned in the supply of drugs. When the vehicle is stopped the driver is a bit resistant to police and is argumentative. Please use the scale below with −5 being least likely to search and +5 being most likely to search

Appendix 2

Officer responses

Table 4 Vignette responses on likelihood of a stop and search being justified

	Score									
	−5	−4	−3	−2	−1	+1	+2	+3	+4	+5
Question 1	8 (6.8%)	3 (2.5%)	10 (8.5%)	11 (9.3%)	10 (8.5%)	15 (12.7%)	20 (16.9%)	19 16.1(%)	13 (11%)	9 (7.6%)
Black	5 (8.2%)	2 (3.3%)	4 (6.6%)	7 (11.5%)	6 (9.8%)	7 (11.5%)	10 (16.4%)	14 (23%)	3 (4.9%)	3 (4.9%)
White	3 (5.3%)	1 (1.8%)	6 (10.5%)	4 (7%)	4 (5%)	8 (14%)	10 (17.5%)	5 (8.8%)	10 (17.5%)	6 (10.5%)
Question 2	43 (36.4%)	11 (9.3%)	16 (13.6%)	5 (4.2%)	10 (8.5%)	19 (16.1%)	3 (2.5%)	4 (3.4%)	2 (1.7%)	5 (4.2%)
Black	17 (32.7%)	6 (11.5%)	8 (15.4%)	1 (1.9%)	3 (5.8%)	11 (21.2%)	0 (%)	2 (3.8%)	0 (%)	4 (7.7%)
White	26 (39.4%)	5 (7.6%)	8 (12.1%)	4 (6.1%)	7 (10.6%)	8 (12.1%)	3 (4.5%)	2 (3.0%)	2 (3.0%)	1 (1.5%)
Question 3	1 (0.8%)	1 (0.8%)	3 (2.5%)	2 (1.7%)	5 (4.2%)	6 (5.1%)	7 (5.9%)	22 (18.6%)	30 (25.4%)	41 (34.71%)
Black	1 (1.8%)	1 (1.8%)	0 (0%)	1 (1.8%)	3 (5.4%)	2 (3.6%)	3 (5.4%)	11 (19.6%)	14 (25.0%)	20 (35.7%)
White	0 (0%)	0 (0%)	3 (4.8%)	1 (1.6%)	2 (3.2%)	4 (6.5%)	4 (6.5%)	11 (17.7%)	16 (25.8%)	21 (33.9%)
Question 4	4 (3.4%)	5 (4.2%)	5 (4.2%)	9 (7.6%)	4 (3.4%)	17 (14.4%)	16 (13.6%)	12 (10.2%)	16 (13.6%)	30 (25.4%)
Black	1 (1.6%)	4 (6.3%)	3 (4.7%)	3 (4.7%)	3 (4.7%)	14 (21.9%)	8 (12.5%)	4 (6.3%)	10 (15.6%)	14 (21.9%)
White	3 (5.6%)	1 (1.9%)	2 (3.7%)	6 (11.1%)	1 (1.9%)	3 (5.6%)	8 (14.8%)	8 (14.8%)	6 (11.1%)	16 (29.6%)
Question 5	2 (1.7%)	2 (1.7%)	3 (2.5%)	6 (5.1%)	12 (10.2%)	6 (5.1%)	17 (14.4%)	26 (22%)	20 (16.9%)	24 (20.3%)
Black	2 (3.5%)	2 (3.5%)	2 (3.5%)	3 (5.3%)	6 (10.5%)	2 (3.5%)	6 (10.5%)	12 (21.1%)	11 (19.3%)	11 (19.3%)
White	0 (0%)	0 (0%)	1 (1.6%)	3 (4.9%)	6 (9.8%)	4 (6.6%)	11 (18.0%)	14 (23.0%)	9 (14.8%)	13 (21.3%)

	Score									
	−5	−4	−3	−2	−1	+1	+2	+3	+4	+5
Question 6	2 (1.7%)	3 (2.5%)	6 (5.1%)	5 (4.2%)	13 (11.0%)	9 (7.6%)	12 (10.2%)	20 (16.9%)	24 (20.3%)	24 (20.3%)
Black	2 (3.3%)	2 (3.3%)	1 (1.7%)	5 (8.3%)	7 (11.7%)	4 (6.7%)	4 (6.7%)	12 (20.0%)	12 (20.0%)	11 (18.3%)
White	0 (0%)	1 (1.7%)	5 (8.6%)	0 (0%)	6 (10.3%)	5 (8.6%)	8 (13.8%)	8 (13.8%)	12 (20.7%)	13 (22.4%)
Question 7	2 (1.7%)	1 (0.8%)	2 (1.7%)	4 (3.3%)	9 (7.5%)	15 (12.5%)	25 (20.8%)	25 (20.8%)	35 (29.2%)	2 (1.7%)
Black	1 (1.6%)	1 (1.6%)	2 (3.1%)	2 (3.1%)	7 (10.9%)	8 (12.5%)	12 (18.8%)	14 (21.9%)	16 (25.0%)	1 (1.6%)
White	1 (1.8%)	0 (0%)	0 (0%)	2 (3.6%)	2 (3.6%)	7 (12.5%)	13 (23.2%)	11 (19.6%)	19 (33.9%)	1 (1.8%)
Question 8	1 (0.8%)	1 (0.8%)	0 (0%)	0 (0%)	4 (3.4%)	6 (5.1%)	8 (6.8%)	25 (21.2%)	34 (28.8%)	39 (33.1%)
Black	1 (1.5%)	0 (0%)	0 (0%)	0 (0%)	4 (6.1%)	4 (6.1%)	4 (6.1%)	13 (19.7%)	19 (28.8%)	21 (31.8%)
White	0 (0%)	1 (1.9%)	0 (0%)	0 (0%)	0 (0%)	2 (3.8%)	4 (7.7%)	12 (23.1%)	15 (28.8%)	18 (34.6%)
Question 9	4 (3.4%)	6 (5.1%)	4 (3.4%)	10 (8.5%)	10 (8.5%)	16 (13.6%)	13 (11.0%)	27 (22.9%)	14 (11.9%)	14 (11.9%)
Black	3 (6.0%)	4 (8.0%)	1 (2.0%)	4 (8.0%)	7 (14.0%)	8 (16.0%)	7 (14.0%)	6 (12.0%)	5 (10.0%)	5 (10.0%)
White	1 (1.5%)	2 (2.9%)	3 (4.4%)	6 (8.8%)	3 (4.4%)	8 (11.8%)	6 (8.8%)	21 (30.9%)	9 (13.2%)	9 (13.2%)
Question 10	6 (5.1%)	6 (5.1%)	14 (11.9%)	13 (11.0%)	10 (8.5%)	19 (16.1%)	14 (11.9%)	18 (15.3%)	14 (11.9%)	4 (3.4%)
Black	4 (6.6%)	4 (6.6%)	6 (9.8%)	8 (13.1%)	8 (13.1%)	10 (16.4%)	4 (6.6%)	5 (8.2%)	10 (16.4%)	2 (3.3%)
White	2 (3.5%)	2 (3.5%)	8 (14.0%)	5 (8.8%)	2 (3.5%)	9 (15.8%)	10 (17.5%)	13 (22.8%)	4 (7.0%)	2 (3.5%)
Question 11	15 (12.7%)	9 (7.6%)	22 (18.6%)	6 (5.1%)	13 (11.0%)	18 (15.3%)	14 (11.9%)	7 (5.9%)	8 (6.8%)	6 (5.1%)
Black	11 (17.7%)	5 (8.1%)	13 (21.0%)	3 (4.8%)	6 (9.7%)	10 (16.1%)	6 (9.7%)	2 (3.2%)	3 (4.8%)	3 (4.8%)
White	4 (7.1%)	4 (7.1%)	9 (16.1%)	3 (5.4%)	7 (12.5%)	8 (14.3%)	8 (14.3%)	5 (8.9%)	5 (8.9%)	3 (5.4%)
Question 12	0 (0%)	1 (0.8%)	3 (2.5%)	1 (0.8%)	3 (2.5%)	8 (6.8%)	7 (5.9%)	21 (17.8%)	27 (22.9%)	47 (39.8%)
Black	0 (0%)	0 (0%)	1 (1.8%)	0 (0%)	2 (3.6%)	5 (8.9%)	4 (7.1%)	17 (30.4%)	13 (23.2%)	14 (25.0%)
White	0 (0%)	1 (1.6%)	2 (3.2%)	1 (1.6%)	1 (1.6%)	3 (4.8%)	3 (4.8%)	4 (6.5%)	14 (22.6%)	33 (53.2%)

	Score									
	− 5	− 4	− 3	− 2	− 1	+ 1	+ 2	+ 3	+ 4	+ 5
Question 13	4 (3.4%)	4 (3.4%)	4 (3.4%)	1 (0.8%)	15 (12.7%)	19 (16.1%)	10 (8.5%)	22 (18.6%)	24 (20.3%)	15 (12.7%)
Black	0 (0%)	4 (7.4%)	2 (3.7%)	1 (1.9%)	11 (20.4%)	12 (22.2%)	4 (7.4%)	7 (13.0%)	8 (14.8%)	5 (9.3%)
White	4 (6.3%)	0 (0%)	2 (3.1%)	0 (0%)	4 (6.3%)	7 (10.9%)	6 (9.4%)	15 (23.4%)	16 (25.0%)	10 (15.6%)
Question 14	11 (9.3%)	7 (5.9%)	20 (16.9%)	14 (11.9%)	11 (9.3%)	19 (16.1%)	16 (13.6%)	11 (9.3%)	4 (3.4%)	5 (4.2%)
Black	4 (6.8%)	4 (6.8%)	8 (13.6%)	6 (10.2%)	7 (11.9%)	10 (16.9%)	12 (20.3%)	7 (11.9%)	0 (0%)	1 (1.7%)
White	7 (11.9%)	3 (5.1%)	12 (20.3%)	8 (13.6%)	4 (6.8%)	9 (15.3%)	4 (6.8%)	4 (6.8%)	4 (6.8%)	4 (6.8%)
Question 15	4 (3.4%)	4 (3.4%)	7 (5.9%)	5 (4.2%)	11 (9.3%)	19 (16.1%)	12 (10.2%)	25 (21.2%)	14 (11.9%)	17 (14.4%)
Black	2 (3.4%)	3 (5.2%)	3 (5.2%)	2 (3.4%)	5 (8.6%)	7 (12.1%)	5 (8.6%)	12 (20.7%)	9 (15.5%)	10 (17.2%)
White	2 (3.3%)	1 (1.7%)	4 (6.7%)	3 (5.0%)	6 (10.0%)	12 (20.0%)	7 (11.7%)	13 (21.7%)	5 (8.3%)	7 (11.7%)
All	107 (6.9%)	64 (4.1%)	119 (7.6%)	88 (5.6%)	135 (8.6%)	205 (13.1%)	184 (11.8%)	284 (18.2%)	269 (17.2%)	107 (6.9%)
Black	54 (6.8%)	42 (4.3%)	54 (6.8%)	44 (5.5%)	80 (10.1%)	113 (14.2%)	85 (10.7%)	136 (17.2%)	131 (16.5%)	54 (6.8%)
White	53 (6.9%)	22 (2.9%)	65 (8.5%)	44 (5.7%)	55 (7.2%)	92 (12.0%)	99 (12.9%)	148 (19.2%)	138 (17.9%)	53 (6.9%)

Authors' contributions Major contribution by Andy Brittain; substantial contribution by Christopher Sims; statistical work by Vincent Harinam; editing and review by Heather Strang. All authors read and approved the final manuscript.

Availability of data and materials Not applicable.

Declarations

Ethics approval and consent to participate Not applicable.

Competing interests Thr authors declare that they have no competing interests.

References

BBC News (2023). *Police strip-searched children as young as eight.* BBC News, 27/03/2023.

Bottoms, A. E., & Tankebe, J. (2017). Police legitimacy and the authority of the state. In A. du Bois-Pedain, M. Ulvang, & P. Asp (Eds.), *Criminal law and the authority of the state.* Bloomsbury Publishing.

Casey, B. (2023). *An independent review into the standards of behaviour and internal culture of the Metropolitan Police Service.* Online: Metropolitan Police Service.

Cockcroft, T. (2007). Police culture(s): Some definitional, methodological and analytical considerations. *Police Occupational Culture: New Debates and Directions.* (pp. 85–102)

Flacks, S. (2018). The stop and search of minors: A 'vital police tool'? *Criminology & Criminal Justice, 18*, 364–384.

Hargreaves, J. (2018). Police stop and search within British Muslim communities: Evidence from the Crime Survey 2006–11. *British Journal of Criminology, 58*, 1281–1302.

Her Majesty's Inspectorate of Constabulary and Fire and Rescue Services (HMICFRS). (2017). *PEEL: Police legitimacy.*

Hussein, D.M.E.-D.M. (2018). A survey on sentiment analysis challenges. *Journal of King Saud University-Engineering Sciences, 30*, 330–338.

Independent Office for Police Conduct (IOPC). (2022). *Operation hotton learning report - January 2022.*

McCambridge, J., Witton, J., & Elbourne, D. R. (2014). Systematic review of the Hawthorne effect: New concepts are needed to study research participation effects. *Journal of Clinical Epidemiology, 67*, 267–277.

Mikroyannidis, A., & Theodoulidis. (2006). *Heraclitus ii: a framework for ontology management and evolution in proceedings of the 2006 IEEE/WIC/ACM International Conference on Web Intelligence, Nishida, Toyoaki (Chair).* (pp. 514–521). IEEE Computer Society.

Miller, J., Bland, N., Quinton, P., & Willis, C. F. (2000). *The impact of stops and searches on crime and the community (No. 127).* Home Office, Policing and Reducing Crime Unit, Research, Development and Statistics Directorate.

Murray, K., McVie, S., Farren, D., Herliitz, L., Hough, M., & Norris, P. (2021). Procedural justice, compliance with the law and police stop-and-search: A study of young people in England and Scotland. *Policing & Society, 31*, 263–282.

Office of National Statistics. (2021). *Census.*

Quinton, P. (2011). The formation of suspicions: Police stop and search practices in England and Wales. *Policing and Society, 21*, 357–368.

Sunshine, J., & Tyler, T. R. (2003). The Role of Procedural Justice and Legitimacy in Shaping Public Support for Policing. *Law & Society Review, 37*, 513–548.

Terrill, W., Paoline, E. A., III., & Manning, P. K. (2003). Police culture and coercion. *Criminology, 41*(4), 1003–1034.

Andy Brittain Chief Superintendent Metropolitan Police Service, UK; Commander responsible for front-line policing in the Met's South BCU.

Christopher Sims Affiliated lecturer, Institute of Criminology, Cambridge University, UK and formerly Chief Constable, West Midlands Police, UK.

Vincent Harinam Research Associate, Institute of Criminology, University of Cambridge, UK.

Heather Strang Director, Lee Centre of Experimental Criminology, Institute of Criminology, University of Cambridge, UK.