

# Goldilocks and the three “Ts”: Targeting, testing, and tracking for “just right” democratic policing

Lawrence W. Sherman<sup>1,2</sup>

<sup>1</sup> Institute of Criminology, University of Cambridge, Cambridge, UK

<sup>2</sup> Cambridge Centre for Evidence-Based Policing Ltd., Cambridge, UK

## Correspondence

Lawrence Sherman, Institute of Criminology, University of Cambridge, Sidgwick Avenue, Cambridge CB3 9DA, UK.

Email:

[lawrence.sherman@crim.cam.ac.uk](mailto:lawrence.sherman@crim.cam.ac.uk)

## Abstract

**Research Summary:** Police are often criticized for doing “too much” or “too little” policing in various situations. These criticisms amount to testable hypotheses about whether “less” force, or intensity, or enforcement would have been enough, or whether “more” was needed. The rise of evidence-based policing provides a starting point for public dialogues about those hypotheses, in ways that could help to build police legitimacy. Such dialogues can be focused on the questions posed by the three “Ts”: (1) Is police action *targeted* in a way that is proportionate to the harm that it can prevent? (2) Has the action been *tested* and found effective with the kinds of targets, and their levels of harm, where it is being used? (3) Is police action *tracked* to ensure it is delivered in the way that has been tested, and in compliance with relevant legal requirements? In this lecture, I frame the issue as follows: *Can more widespread use of better research evidence on targeting, testing, and tracking police actions, shared more clearly among the public and police, help reduce the wide range of oscillation between over-policing and under-policing?*

**Policy Implications:** The use of these questions in public dialogue would be especially relevant to the three biggest threats to police legitimacy in the aftermath of George Floyd’s murder: (A) police killing people, (B) police stopping people, and (C) police under-patrolling

-----  
This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2022 The Authors. *Criminology & Public Policy* published by Wiley Periodicals LLC on behalf of American Society of Criminology.

high-crime hot spots (while over-patrolling low-crime areas). One result of applying the three-Ts questions to these threats, for example, could be the end of the vast overuse of stop and search in low-violence areas. At the same time, this approach could also lead to reductions in homicide by increasing stops in highest violence hot spots. Such changes could demonstrate how the “Goldilocks principle” for the three Ts could get policing closer to “just right” for each place and person being policed.

**KEYWORDS**

evidence-based policing, over-policing, police legitimacy, under-policing

## 1 | THE GOLDILOCKS PRINCIPLE IN SCIENCE AND POLICING

“There is far too little research evidence cited by police leaders in public communications about controversial issues, compared to other areas of government.”

Commander Alex Murray, Metropolitan Police Service, UK, January 12, 2022

How much policing is too much, too little, or “just right” in a liberal democracy? How can we tell? How can we design early warning systems for changes in the answers to these questions—not just for entire nations or cities, but for neighborhoods and communities? How can we even define “policing” in terms of “more” or “less,” when it has so many dimensions?

Answering these questions may be aided by retelling a children’s story, one that Gladwell (2000) would describe as “sticky” (easily remembered as an aide to learning). The children’s story does not answer the policing question as much as it helps formulate it. That simple formulation could help many people understand the potential role of better use of a framework for statistical evidence—the “Triple-T” (Sherman, 2013)—in the governance of democratic policing. It is for that reason that the story is already widely applied across a range of sciences, in which it is known as the “Goldilocks principle.” The use of this English story (Pyle, 1918) as a mnemonic is long overdue in criminology, even though the substantive principle is as old as criminology itself (Beccaria, 1764). Here is a summary of the story:

Once upon a time, a juvenile female burglar named Goldilocks illegally entered the residence of a family of three bears. The family was not home at the time of the crime, since they had just prepared a meal that was too hot to eat. The three bears left their home for a short walk to allow the meal to cool off, with three bowls left out on the kitchen table: one for Father bear, one for Mother bear, and one for the child bear. Goldilocks went first to the Father bear’s bowl, which was still steaming. She picked

up the spoon next to the bowl to gather evidence on the temperature of the food. “Ow! Too hot!” she cried. Then she gathered evidence on the Mother bear’s bowl, which had no steam over it. “Ugh! Too cold,” she said. Then she drew evidence from the child bear’s bowl, and said “Yum! Just right!” And then she ate up all the evidence in that bowl.

In the Triple-T framework of evidence-based policing (Sherman, 2013), the story of Goldilocks might be described as an exercise in *targeting* resource investment for optimal success. The idea of optimal targeting is widely used across a range of sciences. As Hawking and Mlodinow (2010, p. 153) said about the origins of life on earth, the Goldilocks principle is fundamental: “like Goldilocks, the development of intelligent life requires that planetary temperatures be ‘just right.’” Astrobiologist Paul Davies (2008) explores this idea empirically across planetary systems to conclude that our planet is an “enigma,” a mystery that lies at the heart of climate change and sustaining the temperature range we require to exist. In economics, a “Goldilocks economy” sustains moderate economic growth and low inflation, with neither too much growth (with too much inflation) or too little. In sports medicine, there is growing evidence of the risk of both too *little* exercise and too *much* for maximizing lifelong health (Aschwanden, 2022; Schnohr et al., 2021).

In evidence-based policing, the “Koper Curve” (Koper, 1995) is widely known as the optimal dose–response balance for the length of police patrols in a crime hot spot: police staying less than 10 min produces little lasting deterrence after they leave the location, but staying more than 15 min reduces residual deterrence. The Goldilocks “sweet spot” for hot spot patrols, as Koper found by analyzing 24,813 instances of police arrivals and departures (including drive-throughs), was between 10 and 15 min.

Similarly, in a recent Australian experiment, Barnes et al. (2020) found that one police patrol under 15 min can prevent crime in hot spots for up to 4 days before crime and harm start to rise sharply. For that location, “just right” policing was at least once every 4 days. Replications of that experiment in the United Kingdom, in contrast, found somewhat more frequent (daily) patrols of 15 min were needed to achieve substantial reductions in crime harm (Basford et al., 2021; Bland et al., 2021).

Mere patrol presence, however, may be among the most benign ways in which police can intrude on citizen lives in public places; its absence can be less benign, as in uncontrolled outbreaks of violence where patrols are insufficient.

Other activities, such as stop and search, produce far more anger and controversy, dating back to the first years of the Metropolitan Police in London (Moore, 2021). This issue develops in thousands of nonviolent events, each of which may chip away at public support for the police institution. They also inflict direct harm on the subjects of the stops, which can only be justified by evidence that the harm such stops prevent is even greater, at a point of “optimal policing” (Manski & Nagin, 2017).

Yet the greatest damage to public perception of policing being “just right” can derive from a single, extreme case of police harming people, such as the murders of George Floyd in the United States and Sarah Everard in the United Kingdom. No matter how rare (Taleb, 2007), such extreme events may be taken as evidence that the entire police institution is morally bankrupt and unsustainable, even across democratic nations. The 2014 death of Eric Garner at the hands of the New York City Police Department (NYPD), for example, apparently reduced confidence among Afro-Caribbean residents of London in the London police (Laniyonu, 2021). These events force police in each locale to talk about what police are doing “here” to prevent such extreme events, and the evidence to show whether they are getting police effectiveness “just right.”

Tankebe (2013) offers an evidence-based theory for the claim that policing must be both effective and fair to earn its monopoly. Police legitimacy in the eyes of the populace, his research suggests, requires that they must be able to protect the public from violent injury by both other citizens and the police themselves. Crucially, the use of force by the police must be seen to be not only “effective,” but also not “excessive.” Police usage of their many powers must be “not too little and not too much.” It must be, as Goldilocks put it, “just right.”

That is not to suggest that police are the only, or even primary, cause of public safety, nor that crime rates depend solely on the actions or inaction of policing. The importance of the Goldilocks principle, like the Hippocratic oath, is that police can not only “help,” at least to some degree, but they can also “hurt,” causing more harm than good. In policing, as in medicine, tracking the “just right” pathway is hard to do without systematic evidence. One lesson of the COVID pandemic is that evidence can become a much more central element of debate about policies that affect people’s lives, from vaccines to masks to school closures and lockdowns. As Nadhim Zahawi, former vaccines Minister for the United Kingdom has put it:

“I am obsessed by making sure we collect data and publish it,” Zahawi said. “Data and transparency are my allies on this journey. The way you get complex systems to improve and deliver is by being transparent and publishing.” (As quoted in Wheeler & Griffith, 2022).

The need for data during the COVID pandemic became clear in proportion to the extent of restriction on individual liberty that the U.K. government imposed. The daily death toll was evidence widely discussed by voters, as a rising threat that offered great clarity about the proportionate character of lockdown restrictions. Yet even that clarity required more evidence as soon as differences in death rates emerged across different regions of the country. This led to a system of “tiers” of harm that could be matched with tiers of loss of liberty (Sherman, 2022). This system was designed to provide more evidence that policies were proportionate to harm. That evidence did not stop the disputes, but it did make the debates more fact based than they would have otherwise been. Crucially, they may have increased the legitimacy of restrictions on liberty as proportionate and necessary.

Building on that hypothesis, this Vollmer Lecture addresses the following question:

Can more widespread use of better research evidence on targeting, testing and tracking police actions, shared more clearly among the public and police themselves, help reduce the wide range of oscillation between over-policing and under-policing?

The answer, of course, may depend on whether police themselves will accept and understand the implications of a “dialogue” (Bottoms & Tankebe, 2012) between police and the public about the specific uses of police powers and their effects. Equally important is whether universities and government can help promote public understanding of disparities in both victimization and policing by race or other identity groups (e.g., Kumar et al., 2020).

In meeting all these challenges, one clear place to start is illustrating and defining the distinction between over-policing and under-policing. With an evidence-based criterion of using the *minimum* level of police intrusions on citizens’ liberties to provide adequate public safety (Sherman, 2009), criminology would continue the quest launched by Beccaria in 1764 to punish only to the “degree” of severity needed to prevent crime. But that degree must be understood at the macro level of national and local history and not just on a case-by-case basis.

The next part of this lecture (Part 2) reviews a half-century of oscillation between too “little” and too “much” policing. Part 3 disaggregates the idea of policing as a unidimensional quantity, suggesting that police should speak to high-priority concerns about police “actions” as well as to numbers of police “officers.” The most important actions for “just right policing” seem to be these: (A) police use of deadly force, (B) stop and search, and (C) preventive patrols in high-violence locations. By targeting these three kinds of police actions to get them “just right,” police can concentrate on a dialogue about the three Ts of targeting, testing, and tracking police actions.

The evidence needed for police success in satisfying the Goldilocks principle must be clear to both majority and minority groups, with all relevant evidence displayed. Although such dialogues will no doubt be challenging, the final section offers examples of evidence-based practice that offer hope, if not proof, that criminological research can help preserve legitimacy for democratic policing.

## 2 | FROM “CHAOS,” TO “CRIME DROP,” TO “DEFUND THE POLICE”—AND BACK

The recent history of U.S. policing shows enormous oscillations in public demands on the police. As a participant observer of what those demands looked like inside a range of U.S. and overseas police agencies for five decades, I have been able to track voters’ perceptions of police doing too little or too much. These observations began at a point perceived under-policing when I began my work as a research analyst in the NYPD in 1971. In that year, for example, the NYPD began a journey from what has been described as follows:

1. **CHAOS (1960–1990):** what one police historian called what happens when the “cops lose the streets” (Reppetto, 2012, p. 130), to
2. **CRIME DROP (1990–2012):** what Zimring (2012) called the “city that became safe” became a “demonstration project of effective policing,” to
3. **DEFUND THE POLICE (2020–2021):** in which the New York City Mayor agreed, after the murder of George Floyd in Minneapolis, to cut \$1 billion from the annual \$6 billion budget of the NYPD (Rubinstein & Mays, 2020), to
4. **RETURN TO CHAOS (2020–2021):** in which shootings (i.e., woundings or murders) doubled in New York (from 748 to 1480), and murders rose 40% (312 to 436) from 2019 to (most of) 2020 (Blau, 2020).

In those five decades, I spent much of my career focused on a key dimension of over-policing: police shootings of citizens (Sherman, 1980, 1983, 2018). Later to be cited by the U.S. Supreme Court (*Tennessee v. Garner*, cited as 471 U.S. 1 [1985]), my research was shaped in part by the two peak statistics in New York in the year I arrived there. One was that NYPD officers killed 93 people, a number unequalled since then. The other was that 10 police officers were murdered in the line of duty, a number also unequalled since then.

Although annual killings of citizens by police dropped from 93 in 1971 to 25 by 1980 (Sherman, 2018, p. 426), total murders in New York had increased by 350%, from 390 in 1960 to 1826 in 1980 (Reppetto, 2012, p. 130). Even as violence dropped in police encounters with citizens, murder in general continued to increase for two more decades—not just in New York, and not just in the United States, but more broadly across many nations, including England and Wales (while at much lower rates than the United States).

Yet by the start of the 21st century, the Chaos in New York had suddenly ended in a massive Crime Drop, with police getting much credit for reversing the rising trend of murder. Although no one could precisely explain the many possible causes of the Crime Drop, some criminologists (e.g., Reppetto, 2012; Zimring, 2012) tied specific changes in geographic targeting of policing to the reduction of violence. More precisely, Reppetto (2012, pp. 173–174) identified microlocal increases in “dosage” of police activities (to use a public health concept) in evidence-based hot spots, from the 1990s through 2021. Even as critics said the crime drop would not continue, murder kept dropping in New York by 43% in the first two decades of the 21st century, from 673 in 2000 to 289 in 2018. The continued drop was all the more remarkable because it was sustained without increases in incarceration rates, and with a massive reduction in stop and search in the second decade (from some 750,000 per year to some 20,000). Reppetto (2012, p. 174) concluded that “The crux of the crime-fighting programs developed by [NYPD Commissioner] Bratton, [Lawrence] Sherman and [NYPD Commissioner] Kelly has been to maintain police focus on the primary mission.”

Praise for policing, however, evaporated well before the Crime Drop ended. By 2013, a Mayoral candidate won an election on a campaign against “over”-policing. By then, the annual murder count had dropped to 332, or 85% from its 1990 peak of 2245. That drop was a distant memory by the time the recorded stop and search count had risen to 685,724 in 2011 (New York Civil Liberties Union, 2019) when there were 515 murders, or about 1300 stops per murder. (In contrast, England and Wales in 2020–2021 had 695,009 stop and searches recorded alongside 600 murders, for about 1200 stops per murder, almost identical to the 2011 peak of stops in New York City and a far higher rate per murder than New York has seen since 2014).

The 2013 election debate about over-policing in New York was followed the next year by the killing of Michael Brown by a Ferguson, Missouri police officer in 2014. At the same time, the NYPD was slashing the numbers of stop and searches by 98% (from 695,00 to 13,459 by 2019), but not the numbers of police. The number-of-police issue did not become part of the over-policing debate in New York until the murder of George Floyd on May 25, 2020. Yet in the few months after Floyd’s murder, the Mayor and Council approved a \$1 billion cut (out of \$6 billion) and cancellation of plans to hire over 1000 police. What happened next, for unclear reasons, was a rapid resurgence of crime that whipsawed both local and national politics.

At the national level, the most striking evidence that U.S. citizens had condemned over-policing was a July 2020 Gallup poll of 36,000 U.S. citizens, of whom 74% supported the idea of banning stop-and-frisk policing altogether. Black Americans supported ending stop and frisk at 93%, with strong majorities of Hispanic (76%) and White Americans (70%) as well (Crabtree, 2020). But the dominant rhetoric was about defunding the police, not about stop-and-frisk. That dominance did not last long. By the November 2020 elections, a resurgence of violent crime in many cities had soured many voters on claims of over-policing. Although no incumbent Republicans in the U.S. House of Representatives were defeated, 13 incumbent Democrats lost seats. Luscombe (2020) reports that many Democratic House Representatives saw “defunding the police” as a vote-losing proposal that threatened their party’s majority in Congress.

By November 2021, even the proposal to abolish the Minneapolis Police Department had been defeated in a referendum. That police agency, which had retained George Floyd’s killer despite many prior complaints against the officer, was saved by opposition to the abolition by 75% of the votes in predominantly Black neighborhoods—compared to slight majorities for abolition in progressive white areas. Citywide, the plan to replace the Minneapolis Police Department with an undescribed “Department of Public Safety” was defeated with a 56% majority opposed (Navratil et al., 2021). The mayor who had refused to endorse the abolition proposal at the height of the

protests after Floyd's murder was reelected. Yet in the interim, 30% of Minneapolis police officers had quit, and homicides had doubled from 46 in 2019 to 96 in 2021.

The 2021 developments in Minneapolis echoed the national evidence from two polls the preceding summer. A Gallup poll of 1381 respondents (Jones, 2021) showed that there was a modest recovery of confidence in police among Black Americans, from 19% in July 2020 to 27% in 2021 (although still nowhere near the 37% peak of Black approval in the crime drop years before the 2014 events in Ferguson); white approval remained steady in 2020–2021 at 56% (also below the 2001–2010 decade peak, of 63%); and overall confidence in police hit an all-time low (over three decades) of 48% in 2020, inching up to 51% in 2021.

The 1201 respondents to the Ipsos (2021) poll for *USA Today* in the same month offered more details about police issues, including the finding that crime and gun violence was the most “worrying” public issue (32%), well above COVID (21%) and climate change (25%). This sample showed far greater “favorable opinion” of the police (at 72%) than the Gallup Poll, but somewhat fewer respondents (64%) trusted police to “handle crime and safety issues.” More important was the strong support respondents offered for reforms that most police officers may oppose: 90% supported de-escalation training, 81% supported independent institutions to investigate police-involved shootings, and 64% opposed wide police discretion in searching for weapons with “stop-and-frisk.” As for defunding the police, 62% supported using some of the police budget for “community policing and social services.”

Against this backdrop, the 2021 national pattern of policy changes saw majority-Democratic party cities shift from defunding police to restoring the funds cut after Floyd's murder. San Francisco, Seattle, Portland, Chicago, New York City, and other centers of protest at over-policing all saw the pendulum swing back rapidly to fear of under-policing (Economist, 2022). Some might call these events simply the chaos of democracy, especially under a presidency that ended in the insurrection at the U.S. capitol. But others might call it evidence that a more systematic framework is needed to govern policing. That framework, as described below, would be like the use of data in the United Kingdom's COVID response, and continue to be accountable to democratic oversight. Unlike the COVID response, however, police have—so far—been unable to draw on a robust body of evidence to take part in a public dialogue.

Public dialogue was especially difficult at the long tail of the crime drop from 2010 to 2020, when “excessive policing” seemed historically disproportionate to the low murder rates. A paradox of the post-2020 Return to Chaos is that it has made evidence of “what works” to create safer streets much more relevant again. This relevance opens the door to what we have already done with COVID: bringing data to the dialogue, adding evidence to emotions. In the process, we (as a society) now have space in which to consider scientific evidence about the effectiveness of various police activities. That evidence can then be weighed in relation to the pain they cause to individual citizens and the public in general.

In the process of a dialogue about weighing that evidence, we would then do well to become more precise about which police activities we should target, asking where, when, and how (if ever) they should be used. That precision sets the stage for deciding what police actions are “proportionate” to the harm they can prevent. It also sets the scene for distinguishing highest harm policing from all other dimensions of policing. By focusing an evidence-based dialogue on the most harmful things police do, we may be able to target the “power few” of police actions (Sherman, 2007) that drive a democracy's perceptions of too much, too little, or “just right” policing.

### 3 | JUST MEASURES OF PAIN: DISAGGREGATING POLICE INTENSITY

Three police decisions can be targeted as the “power few” issues causing pain to many citizens. Better management of those issues, with a more evidence-based police–public dialogue, might help to create a more stable balance of “just right” policing when police are (A) killing people, (B) stopping them without probable cause to arrest, and (C) deciding where not to conduct preventive “patrol.” Police do many other things that can help or harm, especially prosecuting people, for which evidence could become a central part of a public dialogue (e.g., Neyroud, 2018). But the 2021 Ipsos poll suggests a pattern observed over the last five decades: deaths, stops, and patrol appear to be the three targets of the most intense public anger.

My hypothesis is that if police can get those three “just right,” the police institution can continue to support democracy and the rule of law. Conversely, if majority or minority concerns about these three dimensions cannot be alleviated, police will remain trapped on a fault line within the culture wars. A fault line is not a good place to remain, because it generates great pressure to move to one side or the other. But “taking sides” is exactly what police cannot do to remain nonpartisan and support a rule of law.

These three targets are proposed as a solution to a poorly stated problem. The idea of “too much” (or too little) policing is, even as I have used it here, far too general for meaningful action. We cannot even measure what we mean by an excess or deficiency until we agree on the metrics. That may work for polemics, but not for science. On a dust jacket for a popular recent book on policing, for example, there is a claim that “the problem is the... unprecedented expansion and intensity of policing in the last forty years, a fundamental shift in the role of police in society” (Vitale, 2017). This statement arguably aggregates the concept of policing into a unidimensional phenomenon, one that a single metric could capture over four decades. This premise is unworkable, given the wide range of things that police do across the wide range of different kinds and sizes of police agencies. It also offers no room for a dialogue about changing specific dimensions or tasks of policing, because the dust jacket insists that “the problem is policing itself.”

What does the public—or the multiple publics for policing—think? They certainly do not think the police are perfect. Yet the 2021 polls of U.S. adults reported above suggest that most U.S. citizens, such as the referendum voters of Minneapolis, would oppose abolition of the police. Even those who might prefer an alternative might reasonably ask to inspect the new model that would replace the police as we know it. Such transitional models of police transformation have been suggested, such as my own plan for a more highly paid, evidence-based professional model that would be launched in parallel to an existing police agency, with the old system continuing until the last employee retired (Sherman, 2002). Yet changing the institutional design does not necessarily capture the ideas of “expansion” or “intensity” of policing—let alone the language of “dosage” that is used to describe such metrics as “minutes of police presence in crime hot spots” (e.g., College of Policing, 2018).

A key element of science is arguably the use of taxonomy (classification), by which different categories of a phenomenon are separated as mutually exclusive—or even overlapping. The three “Ts” of evidence-based policing (Sherman, 2013) begin with targeting, which depends on both taxonomy and “stratification”: rank-ordering potential targets of police activity by a constant metric, such as volume of crime at a given street corner, or a crime-harm index total (Sherman et al., 2016) from those crimes. In a business or strategic discussion, stratification across categories is described as setting “priorities.” Hence my suggestion that police can work toward legitimacy with a public



dialogue (Bottoms & Tankebe, 2012) about these three priorities—not as general outcomes, but precisely defined outputs: things police “do” that should be done less or more often, depending on the precise circumstances.

## 4 | APPLYING THE THREE TS FOR THREE KINDS OF PAIN

How can the three Ts help police create a dialogue with the public about getting policing “just right”—neither over-policing nor under-policing? One answer is to show how to use the relevant evidence to measure whether police are indeed getting it right. Once police leaders can identify the issues or areas on which they *do* have it just right, they can concentrate on those in which they *do not*. Preparing to be held accountable in a public dialogue should concentrate the mind. Even if the public is not completely persuaded, the police organization will have gathered the necessary evidence. That achievement alone could realize the Goldilocks principle.

### 4.1 | Killing people

The top priority for “just right” policing is for police to kill as few people as possible (Sherman, 2018). There are many ways to do that, even while the United States consistently witnesses some 1000 deaths a year reported by the *Washington Post* that are caused by police. Little seems to be done about these cases in the United States, although they remain very rare in Europe. Yet wherever they occur, they are often attacked as “too much” policing. There seems no doubt that whether police kill people in the line of duty (as in the case of George Floyd) or off-duty, and whether or not the officer commits a heinous crime off-duty (as in the case of Sarah Everard), police legitimacy is highly threatened by these most visible and publicized cases of police causing death.

How can using the triple-T framework of evidence-based policing help to promote dialogue between police and the public about people killed by police? If we limit this discussion to those killings committed in the line of duty, the answers to our three questions may prove instructive.

#### 4.1.1 | Is police action targeted in a way that is proportionate to the harm that it can prevent?

This is the core of the case-by-case discussion, in which critics say police should have de-escalated a situation, or at least not put themselves into harm’s way. As long as these discussions remain free of statistical evidence of risk and safety, there is little chance of reaching an agreement between critics and police leaders (Sherman, 2018). Yet if there were an adequate research base that showed the risks of harm to police officers or innocent others if police tried to avoid an armed confrontation, there could be some very productive dialogue with communities claiming the killing was disproportionate (or “lawful but awful”). The research would go beyond the particulars of a specific case to place it in a context of risks and benefits.

One central example of statistical dimensions of targeting deadly force is literally whether it is necessary (or even possible) for police to “target” their shots fired at the parts of a human body where bullets are most likely to cause death. Police Chief Lou Dekmar of LaGrange, Georgia (pop. 31,000), for example, recently challenged the doctrine of aiming at core body mass for “shooting

to stop.” He made that challenge as a former President of the International Association of Chiefs of Police, and as a police chief himself since 1995. He has trained his own 85 officers in a new approach he hopes will be less likely to kill people: “shoot to incapacitate” (Thompson, 2021). This plan has attracted academic interest but widespread condemnation from many police trainers, who hold his ideas to be dangerous and likely to increase the risk of death of police officers who aim for the pelvis or legs rather than the chest.

The point of this example is not whether the Dekmar hypothesis is correct, but whether (1) a program of research could and should be developed to test it, and (2) whether this kind of evidence could promote better dialogue about the legitimacy of police practices. “Every time we save a life,” Dekmar said, “we build trust.” If police killings could be cut even without cutting police shootings, this research question is well worth pursuing. A thought experiment could help: imagine that a public dialogue over new evidence from testing that hypothesis *could* lead to fewer fatal shootings—even by using evidence based on retrospective analysis of past shootings. Imagine if the evidence further showed that a change in training *did* lead to fewer fatal police–citizen encounters. That experiment illustrates how evidence on “targeting” police shootings might improve police legitimacy.

Other kinds of targeting evidence might include situations where a subject is mentally ill or just acting strangely. With better data systems, the police might be able to mobilize mental health professionals as first responders, and maintain police out of sight as a backup for protecting either professionals or bystanders. The question of who is the best responder for minimizing violence is an unanswered one, but the research needed is feasible, and can build on an existing evidence base (Pollack & Humphreys, 2020). Excellent analysis of newspaper compilation evidence is already available on the distribution of police killings across different kinds of situations (Zimring, 2017). New research could establish a better basis for selecting a subset of situations to which police would not be sent if safer options are available.

It is worth noting that even in the context of one or two police killings a year, the United Kingdom could also have gained benefit from a public dialogue over research evidence. For example, the 2011 London riots were sparked by the killing of Mark Duggan in a proactive police operation designed to prevent him committing a murder, through his arrest for carrying a gun illegally in a taxi cab. Whatever the arguments for or against that operation, police were crippled by legal rules prohibiting public dialogue about a case until a full and independent investigation was completed. Had such a dialogue begun the moment the killing was announced, it remains possible that the riots may never have developed. The dialogue could have included research evidence on a substantial reduction in homicides of young men in London over the previous decade, associated with proactive policing of weapons and gangs.

#### 4.1.2 | Has the police action been tested and found effective with the kinds of targets, and their levels of harm, where it is being used?

Responses to various situations by police versus other first responders have not yet been tested in any systematic experiments. If they would be, an evidence-based dialogue could be launched every time a person with psychiatric or drug-related disorders is killed (Pollack & Humphreys, 2020). These cases comprise an estimated 21%–25% of all U.S. police fatal shootings per year (Kindy et al., 2020). The quasi-experiments of policy changes in the 1970s provided evidence for major changes in shooting fleeing felons (Sherman, 2018), but the recent debates have been largely evidence free. If mental health or other responders can be randomly assigned to hundreds or

thousands of incidents, and compared to police-only responses, we would have very strong evidence on what differences there may be in total deaths or injuries under the two kinds of responses. Such evidence would be a major contribution to public dialogue in any police force managing mental health and violence cases.

#### 4.1.3 | Is police action tracked to ensure it is delivered in the way that has been tested, and in compliance with relevant legal requirements?

Systematic tracking of police use of force and deadly force by individual police officers, units, and social networks remains a flashpoint in many cases, largely because of decisions to retain officers in frontline positions after they shoot or kill people. The killer of George Floyd is one example. Another is a Pennsylvania State police officer who has killed four people in four separate incidents (Barker et al., 2022), all involving either mental illness, drugs, or both, in rural areas with low levels of crime, and at least one decedent who was reportedly unarmed. Journalists are increasingly tracking such cases, but without the advantage of using big data on large cohorts of police to predict those who are most likely to shoot or kill someone. Better research using such data can even identify social network patterns in police agencies that raise the risks of shootings (e.g., Zhao & Papachristos, 2020). Other algorithmic methods of tracking officer behavior could be developed as more accurate early warning systems about police who may kill unnecessarily (or criminally).

If the police were to track patterns of shooting across all officers, they could justifiably move officers away from frontline duties after they have demonstrated a propensity to shoot rather than choosing alternative dispositions. They would not need to terminate officers at high risk; the rising use of rapid video responses by officers in uniform who speak to callers without leaving their police agency's control rooms, as tested successfully in Kent, UK (Rothwell, 2021), would provide ample jobs for officers who are too risky to assign to face-to-face encounters. Evidence that a police force tries to minimize police killings in this way could have, for example, helped to manage the fury over George Floyd's death, with its international repercussions for democratic policing. Such a policy might even have saved George Floyd's life.

## 4.2 | Stop and search

A second priority for "just right" policing is to minimize stop-and-frisk to the greatest degree possible. Although a majority of U.S. citizens may wish to ban stops without probable cause altogether (Ipsos, 2021), that could well cause an increase in murders, especially in high-violence poverty areas. As the U.S. National Academies of Science, Engineering and Medicine's (2018, p. 310) report on Proactive Policing concluded, stop-and-frisk in areas of frequent violence with weapons can help to prevent such crimes, at least in the short term:

CONCLUSION 4-9: Evaluations of focused uses of stop, question, and frisk (SQF) (combined with other self-initiated enforcement activities by officers), targeting places with violence or serious gun crimes and focusing on high-risk repeat offenders, consistently report short-term crime-reduction effects; jurisdictional impacts, when estimated, are modest. There is an absence of evidence on the long-term impacts of focused uses of SQF on crime.

As Gladwell (2019, pp. 287–342) points out, the fact that this tactic works in the small number of very violent areas is no justification for using it all over the United States—most of which has very low rates of violence. His praise for the tactic as tested in the highest gunfire beat of Kansas City (Sherman & Rogan, 1995) is coupled with his withering criticism of its use in low-crime areas. As he wrote about the tragic case of Sandra Bland, whose suicide followed her two nights in Texas Rangers lockup for changing lanes without signaling, she was not stopped in a high-crime area; the location of the stop was 3 miles away from the nearest locations with any violent crimes. Quoting criminologist David Weisburd, Gladwell (2019, p. 341) writes: “Why are you stopping people in places where there’s no crime? ... That doesn’t make sense to me.” Gladwell even cites the temporal dimension of proactive policing in Kansas City (Sherman & Rogan, 1995), in which the experiment was limited to night-time patrols, quoting me as saying “that’s the only time of day when the crime rate was high enough to justify aggressive policing” (Gladwell, 2019, p. 341). Thus, he provides a clear answer to the first question for evidence-based dialogue about stop and search: targeting *only* high-violence areas.

#### 4.2.1 | Is police action targeted in a way that is proportionate to the harm that it can prevent?

Introducing the three Ts into this discussion could help remedy the enormous volatility in using this tactic. Rather than using the logic of increasing stops “everywhere” whenever violent weapons crime goes up—and then slashing them if crime drops—a “tiered” system of targeting only the most violent areas could use far fewer stops overall to accomplish even more violence prevention (Sherman, 2022). As Gladwell (2019, p. 310) said about the North Carolina State Highway Patrol doubling its traffic stops from 400,000 to 800,000 over 7 years, “Does that sound like focused and concentrated policing? ... The part they missed [from the Kansas City experiment] was that aggressive patrol was supposed to be confined to places where crime was concentrated.”

This lack of concern is not distinctly American. In London, as of 2021, the Metropolitan Police Service had no digital reporting that can even track where stops are occurring. The form that police fill out every time they conduct a stop had no field for entering the exact location of the stop. Yet precise locations could be used in maps showing whether stops are targeted where murders and serious violence are most likely to occur. Absent such an evidence-based framework for discussing police stops as a strategy for preventing violence (as distinct from doing drug enforcement), police in London, as elsewhere in the United Kingdom and the United States, are likely to continue to see numbers of stop and search swing up and down across wide areas, blinded to the distribution of stops within those areas (Figures 1 and 2). Such swings are much more reflective of changing political directives than of an evidence-based approach to violence prevention

The hypothesis that targeting stops in high-violence areas will reduce serious violence remains to be tested by a randomized trial in either the United States or the United Kingdom, but that may soon change. Dorset (UK) Police are about to launch two such experiments (Browning, 2022). With a planned day-by-day crossover design comparing days with high-frequency stops to days of low frequency in two high-violence hot spots, Dorset police will target stops where they could do the most good. The experiments will also address the issues concerning police targeting of stops by the race of persons stopped. The tool they will use to assess targeting by race is a Risk-Adjusted Disparity (RAD) Index (Sherman & Kumar, 2021). This tool computes racial disparity rates by using the number of White and non-White victims of violent crime as the denominator, and the number of stops of each group as the numerator (Sherman & Kumar, 2021); the disparity

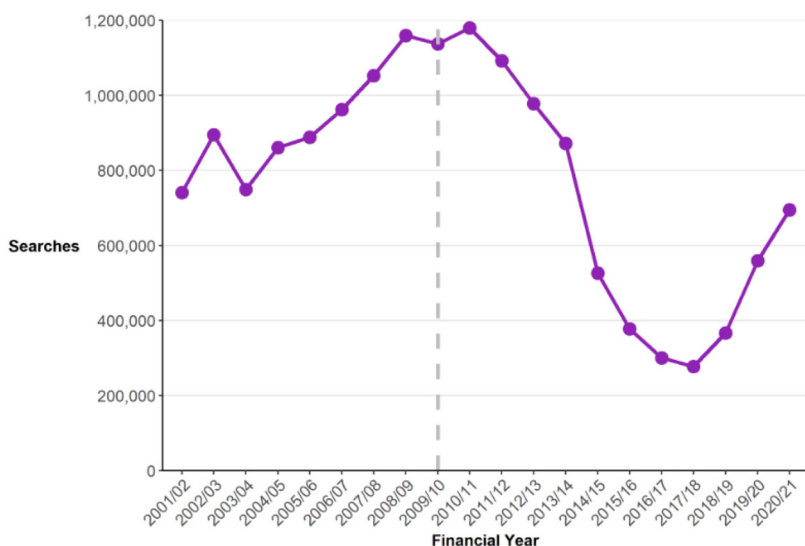
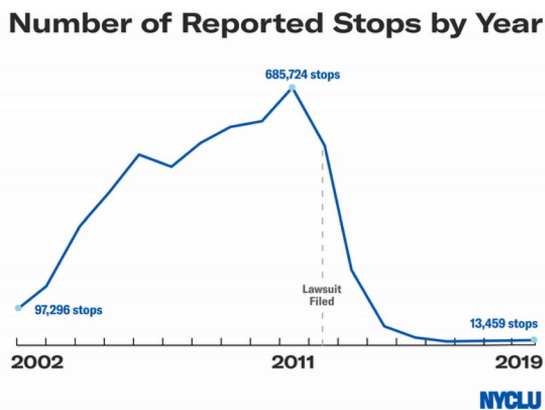


FIGURE 1 Total stop and searches by year, England and Wales, 2001--2021 [Color figure can be viewed at wileyonlinelibrary.com]

Source: <https://www.gov.uk/government/statistics/police-powers-and-procedures-stop-and-search-and-arrests-england-and-wales-year-ending-31-march-2021/police-powers-and-procedures-stop-and-search-and-arrests-england-and-wales-year-ending-31-march-2021>

FIGURE 2 Stop and search by year, New York City Police [Color figure can be viewed at wileyonlinelibrary.com]

Source: <https://www.nyclu.org/en/stop-and-frisk-data>, downloaded January 17, 2022



index is the ratio of the two, in which non-Whites could have lower or higher rates than Whites. This general method of testing for racial disparity (as distinct from resident population as the denominator) has already been adopted by a national statistics program in the UK Home Office, which produced Figure 3 using suspect race for violent crimes as the denominators (rather than victim race as recommended by Sherman & Kumar, 2021).

As Figure 3 shows, the suspect-based RAD Index of racial disparity in stops varies widely across London. In seven of 32 boroughs, Whites are more likely to be stopped per 100 White suspects of violent crime than Blacks per 100 Black suspects of violent crime. Yet in 22 boroughs, Blacks are more likely to be stopped by police per 100 Black suspects of violent crimes than are Whites per 100 White violence suspects. Ten of 32 boroughs have no racial disparity against Blacks, and the overall

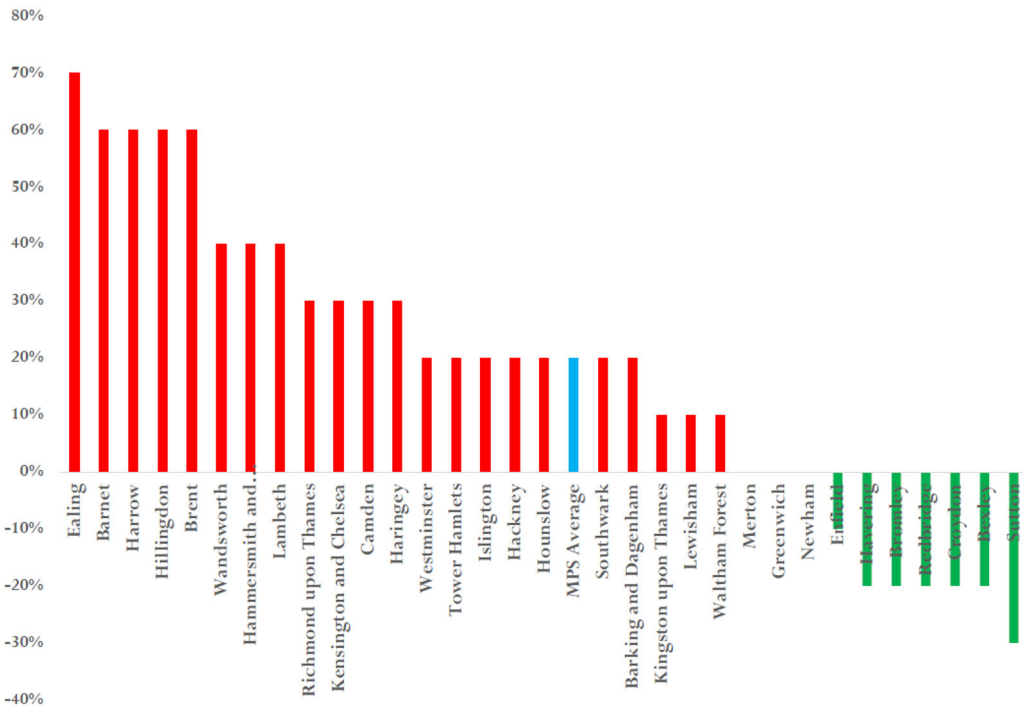


FIGURE 3 Stop and search disparity using Suspect-Based RAD. Black people compared with White people, Metropolitan Police Service (MPS) boroughs, year to March 2021 [Color figure can be viewed at wileyonlinelibrary.com]

Source: Home Office, UK. <https://www.gov.uk/government/statistics/police-powers-and-procedures-stop-and-search-and-arrests-england-and-wales-year-ending-31-march-2021/police-powers-and-procedures-stop-and-search-and-arrests-england-and-wales-year-ending-31-march-2021>

disparity is a 20% higher risk of non-Whites being stopped than Whites. These findings are a far cry from the Dorset disparity based on resident population, which shows a 2000% higher risk for Blacks being stopped compared to Whites. In the latter case, however, the disparity is complicated by the large number of nonresident tourists in the beach towns of Dorset on high-crime, highly policed summer days.

Whether using evidence analyzed with a RAD Index can help improve police–public dialogue about police stops remains to be seen. But it can also address concerns voiced by parents of murdered young people in London, who have demanded more stops of their own ethnic groups to reduce the violence in their neighborhoods (Sherman & Kumar, 2021).

#### 4.2.2 | Has the action been tested and found effective with the kinds of targets, and their levels of harm, where it is being used?

Given the level of controversy and criticism generated by stop and search, it is a striking fact that there has never been a randomized controlled trial of the practice anywhere in the world (to my knowledge). Despite the consistent quasi-experimental evidence noted by the National

Academies of Science, Engineering and Medicine (2018), the rigor of an randomized controlled trials (RCT) remains missing in this search for “just right” policing. And although Dorset may soon produce such evidence, the county is a relatively low-crime area; it is not even eligible for national antiviolence funding. What it may do is to prove the concept of randomized trials, so that other police agencies in the United Kingdom and elsewhere can follow a template and replicate the 1-year Dorset test.

At the same time, other forces can deploy a RAD Index for tracking who is targeted for stops, by race or ethnicity. The RAD Index should be based on characteristics of victims of violence, rather than suspects, for good statistical reasons. Although the race of victims can be identified (in principle) in 100% of recorded violent crimes, there will never be 100% clarity about the race of suspects. Because most violent crime is intraracial, it may appear to be a minor distinction. Yet the legitimacy of police stops derives from doing something to help “protect the people stopped,” as well as everyone else in the area. What better way can there be to demonstrate that concern than to use victim data as the denominator?

#### 4.2.3 | Is the police action tracked to ensure it is delivered in the way that has been tested, and in compliance with relevant legal requirements?

As the RAD Index demonstrates, there is now a way to monitor each police officer and policing area for potential racial bias against people stopped, controlling for the characteristics of people (specifically victims) on the streets relevant to the tracking. This tracking could be just the start.

Body-worn cameras are also being used to sample the *qualitative* dimensions of each stop, following the template developed by Nawaz and Tankebe (2018). In one U.K. force, dimensions of procedural justice are coded by supervisors to review with officers recorded during citizen encounters. Although NYPD officers may conduct more stops of minorities when wearing body-cameras, and with more questionable legal bases (Braga et al., 2021), that is exactly the kind of evidence police leaders can use to manage police compliance for different (and more lawful) results. A weekly meeting to review Body-worn camera (BWC) records of stops with local officers, one-by-one, is a viable method for feeding back the evidence to the officers. Such tracking could be tested in a randomized trial in itself, with some officers receiving feedback and others not.

No matter how the evidence develops, U.S. and U.K. police have several decades of experience filing reports on all stops. There is an expectation that such record-keeping must occur. Building on that expectation should help police leaders who try to generate better records, and better procedural justice when police conduct stops—especially in the high-violence areas where stops can sometimes do more good than harm.

### 4.3 | Preventive patrol in hot spots

As the 2021 USAToday/IPSOS (2021) poll showed (see Topline & Methodology, p. 8), 77% of U.S. citizens want to deploy “more officers to street patrols.” The evidence that such patrols are effective has been reviewed by the U.S. National Academies of Science, Engineering and Medicine (2018, p. 307) in its Proactive Policing report, which concluded as follows:

CONCLUSION 4-1: The available research evidence strongly suggests that hot spots policing strategies produce short-term crime-reduction effects without simply dis-

placing crime into areas immediately surrounding targeted locations. Hot spots policing studies that do measure possible displacement effects tend to find that these programs generate a diffusion of crime-control benefits into immediately adjacent areas. There is an absence of evidence on the long-term impacts of hot spots policing strategies on crime and on possible jurisdictional outcomes.

What the National Academies of Science (NAS) report did not say is the converse of its conclusion, which may have occurred in many cities that rapidly lost police officers who quit during and after the events of 2020: reducing or *suspending patrols in high-crime hot spots may lead to increases in violent crime*. This finding is central to the question of police effectiveness. Evidence from police strikes (Sherman, 1992, pp. 192–193) shows massive increases in crime when levels of general police patrol are slashed, as does the recent experience of a police-free zone in Seattle (Piza & Connealy, 2022). And because violent crime is heavily concentrated in a small number of locations, similar effects may be generated in those locations with highly localized reductions (or failures to deliver) patrols (see, e.g., Barnes et al., 2020).

For Tankebe's (2013) requirement of police to be effective against crime in order to earn legitimacy, the decisions about where to conduct patrols may comprise the most important dimension of the Goldilocks principle in policing. Placing this issue on the same level as police killings and stops is no exaggeration of the importance of hot spots patrols. Given the claim that the Crime Drop was driven, at least in part, by *more* patrolling of hot spots (Repetto, 2012, p. 174), the 2020 Renewal of Chaos may have been driven by *less* patrolling in those same places—and perhaps by fewer targeted searches as well. Whether there was a patrol reduction, and if so whether it was due to COVID or public backlash over the Floyd murder are all unanswered empirical questions. But if such evidence was routinely used in public dialogue with police about neighborhood crime trends, such evidence might become readily available. That would make it all the more important for police to deploy the Triple-T in public dialogues about where and when they patrol.

#### 4.3.1 | Is police action targeted in a way that is proportionate to the harm that it can prevent?

At present, few police forces in the United Kingdom or the United States appear to be targeting patrols to the places and times at which most crime harm (Sherman et al., 2016) occurs in public spaces (Weinborn et al., 2017; Koper et al., 2020). Yet incentives to do so may increase. A national initiative in England and Wales was announced in 2021, for example, by which 18 police forces with highest violence rates could receive extra funding to develop hot spots patrols. The evidence needed to justify receipt of these funds was not initially rigorous, but could become more precise. But as of 2022, the kind of mapping of hot spots (Buerger et al., 1995) and harm spots (Weinborn et al., 2017) that would make the best use of police patrols is far from business as usual in most first-world police agencies.

If the police were to start tracking the concentrations of high *harm* from violence, as distinct from places with high frequencies of low-harm crime (e.g., shoplifting), they would be in a much better position to cope with spikes in homicide or other “signal” crimes. Police chiefs could appear at press conferences and display maps or graphs to pinpoint where the highest harm is occurring and where police resources are being directed. They could even display a tightening relationship between where (and when) the violence occurs and the comparable distributions of GPS-tracked



police time. This kind of focus, first demonstrated in the NYPD in 1994 as part of “COMPSTAT,” was long held to be secret information. In the second term of NYPD Commissioner Ray Kelly, he barred journalists from attending the discussions. Yet it seems unlikely that violent offenders were paying much attention to the technical details affecting the chances of their being arrested in the act of crime. It is more likely that there were concerns in many communities about under-policing or over-policing.

Seen from the perspective of the Bottoms and Tankebe (2012) call for a dialogue between police and public, the level of detail used in both research and (some) examples of practice is a prime example. The allocation of police patrol is definitely a zero-sum game, at least from a voter’s perspective. In such circumstances, greater transparency has much to offer.

#### 4.3.2 | Has this police action been tested and found effective with the kinds of targets, and their levels of harm, where it is being used?

The tragedy of the “map gap” in allocating patrols more precisely is that the testing is extensive. Not only has the National Academies of Science concluded that focusing police in hot spots “works” to reduce violence (in the short run), so has the Crime and Policing Minister of the United Kingdom. Neither source has yet produced a clear change in targeting patrol times, but there are at least efforts to move in that direction. Police resistance to hot spots policing is often based on the “displacement” myth, against which the evidence seems extremely strong in most tests (Bowers et al., 2011). Hence, like the science of vaccines, the evidence for hot spots policing may require special efforts to combat the science deniers. Public dialogue about local data could be a successful approach to that challenge.

#### 4.3.3 | Is this police action tracked to ensure it is delivered in the way that has been tested, and in compliance with relevant legal requirements?

The tracking of police patrols remains a technical and managerial challenge, even in the agencies that are trying to introduce such tracking. Promises made by police radio and smart phone providers have been broken repeatedly. The claims that their tech can track where police are at all times fail on such surprises as the phone GPS transponder not working as soon as an officer sits down inside a police car. Communications equipment will probably not improve until the policing market demands more reliable tracking of where police are, if only for police safety, let alone public safety. Yet the problem seems likely to be solved, at least technically.

The larger challenge will be to transform a culture in which police have rarely been directed to specific places for specific periods of time, nor have supervisors been asked to track the performance of those tasks. A case study of two supervisors in tracking patrols within large train stations compared two approaches, with an encouraging result (de Brito & Ariel, 2017). For a supervisor who simply emailed patrol officers with instructions and failed to speak to officers face-to-face, low levels of compliance were recorded. For a supervisor who discussed, in person and face-to-face, the foot patrol assignment to high-crime micro areas within the station, a high level of compliance was achieved. A long-term effort to shape patrol patterns around hot spots has also succeeded with the 110 police officers of Manhattan, Kansas (Koper et al., 2021). These findings suggest that at least in some police cultures, personal leadership can reshape police patrolling patterns

In terms of community dialogue, a local residents' meeting could be transformed by patrol tracking metrics with violent crime harm maps, even three-dimensional ones. Although residents in low-crime areas might complain that their patrols have declined, police can be steadfast about showing that patrols have been going to people who need them most—if they have. Although the politics of policing often drives patrols away from low-voter-turnout-but-high-violence areas, the overall murder rate in a city affects everyone.

Many other examples of targeting, testing and tracking could be cited to support a public dialogue on “just right” policing. The examples of the “Three T” facts described above for police killings, stoppings and failed prevention are just a starting point for building police legitimacy. Yet tracking the number of police killed in the line of duty will always be important for getting policing right. The fact of a 59% increase (Wilson, 2022) in officers killed feloniously in the US (from 46 in 2019 to 73 in 2021) must also be part of the dialogue.

## 5 | CONCLUSION

In the quarter-century since I first proposed a framework for evidence-based policing (Sherman, 1998), the digital revolution has transformed the potential for that idea. Calculations that were almost unimaginable in 1998 can be performed today with the touch of a button. Yet as Scotland Yard Commander Alex Murray recently told a group of police leaders studying evidence-based policing at Cambridge, policing is substantially behind the rest of government in using digital tools to make decisions. Despite massive investments in technologies, their use in producing “just right” evidence remains undeveloped in policing. The consequence is the Murray quotation used as the opening epigraph of this lecture, which calls on police to use more research evidence in public communications.

The need for more evidence-based police communications has never been greater, especially in service of the Goldilocks principle. It may be indispensable to preserving democratic policing, especially when democracy itself is under threat. Even the police defense of the U.S. capitol on January 6, 2021, for example, was immediately challenged by some observers as intentional under-policing. That claim was made despite 150 police officers being injured and several dying in defense of the electoral process. But that event was not unprecedented in liberal democracies. Similar events killed many police officers a century earlier in the Weimar Republic's pre-Hitler Germany, as meticulously documented by Liang (1970). But in 2020-2021, the US saw the largest jump.

Democratic governance and a nonpartisan rule of law require that the state maintain a monopoly on the legitimate use of force. Preserving that monopoly is probably the primary task of the police institution. But as Bottoms and Tankebe (2012) suggest, the legitimacy of that monopoly must be earned every day—not just through action, but also through dialogue about police actions to establish policing by consent. That dialogue needs evidence.

As the founder of the nonprofit Societies of Evidence-Based Policing ([sebp.police.uk](http://sebp.police.uk)), Alex Murray leads a worldwide professional social movement of some 5000 police officers. The SEBP members read research, attend conferences, and increasingly report their own research. None of this activity was in view at the time of the first statement on evidence-based policing (Sherman, 1998), when the model for building knowledge was to let the academics do it. The 21st century model crucially adds the recruitment and training of “pracademic” criminologists, like the leader of the American Society of Evidence-Based Policing, former Sacramento Police Sergeant Renee Mitchell, Ph.D., and the leader of the Australia-New Zealand Society of EBP, Superintendent

David Cowan; both have led experiments in their own police agencies. With such grass roots pracademics as advocates providing police with role models for building and using evidence, we see good reason to hope for progress at getting policing right. Just right.

## ACKNOWLEDGMENTS

The author wishes to thank Heather Strang, Paul Rock, and Justice Tankebe for comments on an earlier draft of this lecture, and the editors of *Criminology and Public Policy*, Cynthia Lum and Christopher Koper, for their comments, patience, and support.

## CONFLICT OF INTEREST

The author confirms that they have no conflict of interest to declare.

## REFERENCES

- Aschwanden, C. (2022, January 4). Is it possible to exercise too much? And if so, how can I tell if I'm overdoing it? *New York Times*. <https://www.nytimes.com/2022/01/04/well/move/too-much-exercise.html?smid=em-share>
- Barker, K., Eder, S., & Kirkpatrick, D. (2022, January 8). After 4 Killings, 'Officer of the Year' Is Still on the Job. A Pennsylvania state trooper was returned to duty following three investigations by his own agency. A fourth inquiry is underway. *New York Times*. <https://www.nytimes.com/2021/12/30/us/pennsylvania-trooper-jay-splain-investigation.html>
- Barnes, G. C., Williams, S., Sherman, L. W., Parmar, J., House, P., & Brown, S. A. (2020). Sweet spots of residual deterrence: A randomized crossover experiment in minimalist police patrol. *SocArXiv Papers*. <https://doi.org/10.31235/osf.io/kwf98>
- Basford, L., Sims, C., Agar, I., Harinam, V., & Strang, H. (2021). Effects of one-a-day foot patrols on hot spots of serious violence and crime harm: A randomised crossover trial. *Cambridge Journal of Evidence-Based Policing*, 5, 119–133.
- Bland, M., Leggetter, M., Cestaro, D., & Sebire, J. (2021). Fifteen minutes per day keeps the violence away: A crossover randomised controlled trial on the impact of foot patrols on serious violence in large hot spot areas. *Cambridge Journal of Evidence-Based Policing*, 5(3), 93–118.
- Blau, R. (2020, December 21). Why are shootings up in New York City? We asked four people on the front lines. *The City*. <https://www.thecity.nyc/2020/12/21/22189682/why-are-shootings-up-in-new-york-city-in-2020-nypd>
- Beccaria, C. (1764). *Of crimes and punishments*. Milan.
- Bottoms, A., & Tankebe, J. (2012). Beyond procedural justice: A dialogic approach to legitimacy in criminal justice. *The Journal of Criminal Law and Criminology*, 102(1), 119–170.
- Bowers, K. J., Johnson, S. D., Guerette, R. T., Summers, L., & Poynton, S. (2011). Spatial displacement and diffusion of benefits among geographically focused policing initiatives: A meta-analytical review. *Journal of Experimental Criminology*, 7(4), 347–374.
- Braga, A. A., MacDonald, J. M., & McCabe, J. (2021). Body-worn cameras, lawful police stops, and NYPD officer compliance: A cluster randomized controlled trial. *Criminology*. <https://doi.org/10.1111/1745-9125.12293>
- Browning, P. (2022). Designing an experiment for reducing both street violence and racial disparity in policing: Protocol for a randomised trial of a stop-search strategy (M.St. thesis). Institute of Criminology, University of Cambridge.
- Buerger, M. E., Cohn, E. G., & Petrosino, A. J. (1995). Defining the “hot spots of crime”: Operationalizing theoretical concepts for field research. In J. E. Eck & D. Weisburd (Eds.), *Crime and place* (pp. 237–257). Lynne Rienner Publishers.
- College of Policing. (2018). The effects of patrol dosage on crimes and on calls for service in night time economy hotspots. <https://whatworks.college.police.uk/Research/Research-Map/Pages/ResearchProject.aspx?projectid=449>
- Crabtree, S. (2020, July 22). Most Americans say policing needs 'Major Changes.' *Gallup News*. <https://news.gallup.com/poll/315962/americans-say-policing-needs-major-changes.aspx>
- Davies, P. (2008). *The goldilocks enigma*. Houghton Mifflin Company.

- de Brito, C., & Ariel, B. (2017). Does tracking and feedback boost patrol time in hot spots? Two tests. *Cambridge Journal of Evidence-Based Policing*, 1(4), 244–262.
- Economist. (2022, January 15). Refunding the police: As violent crime leaps, liberal cities rethink cutting police budgets. The tricky politics of criminal-justice reform at a time of rising fears. *The Economist*.
- Gladwell, M. (2000). *The tipping point*. Little, Brown and Company.
- Gladwell, M. (2019). *Talking to strangers*. Allen Lane.
- Hawking, S., & Mlodinow, L. (2010). *The grand design*. Bantam Books.
- Ipsos. (2021, July 8). As public safety tops the agenda, Americans want both order and justice. *USA Today*. <https://www.ipsos.com/en-us/news-polls/usa-today-crime-and-safety-2021>
- Jones, J. M. (2021, July 14). In U.S., Black confidence in police recovers from 2020 low. *Gallup*. <https://news.gallup.com/poll/352304/black-confidence-police-recovers-2020-low.aspx>
- Kindy, K., Tate, J., Jenkins, J., & Mellnick, T. (2020, October 17). Fatal police shootings of mentally ill people are 39 percent more likely to take place in small and midsized areas. *The Washington Post*. [https://www.washingtonpost.com/national/police-mentally-ill-deaths/2020/10/17/8dd5bcf6-0245-11eb-b7ed-141dd88560ea\\_story.html](https://www.washingtonpost.com/national/police-mentally-ill-deaths/2020/10/17/8dd5bcf6-0245-11eb-b7ed-141dd88560ea_story.html)
- Koper, C. S. (1995). Just enough police presence: Reducing crime and disorderly behavior by optimizing patrol time in crime hot spots. *Justice Quarterly*, 12(4), 649–672.
- Koper, C. S., Lum, C., Wu, X., & Fritz, N. (2020). Proactive policing in the United States: A national survey. *Policing: An International Journal*, 43(5), 861–876.
- Koper, C. S., Lum, C., Wu, X., & Hegarty, T. (2021). The long-term and system-level impacts of institutionalizing hot spot policing in a small city. *Policing: A Journal of Policy and Practice*, 15(2), 1110–1128.
- Kumar, S., Sherman, L. W., & Strang, H. (2020). Racial disparities in homicide victimisation rates: How to improve transparency by the Office of National Statistics in England and Wales. *Cambridge Journal of Evidence-Based Policing*, 4(3), 178–186.
- Laniyonu, A. (2021). Phantom pains: The effect of police killings of Black Americans on Black British attitudes. *British Journal of Political Science*. <https://doi.org/10.1017/S0007123421000557>
- Liang, H. H. (1970). *The Berlin police force in the Weimar Republic*. University of California Press.
- Luscombe, R. (2020, November 8). James Clyburn: ‘defund the police’ slogan may have hurt Democrats at polls; Black South Carolina congressman likens mantra at BLM protests to radical messaging in civil rights efforts that eroded movement. *Guardian*. <https://www.theguardian.com/us-news/2020/nov/08/james-clyburn-defund-police-slogan-democrats-polls>
- Manski, C. F., & Nagin, D. S. (2017). Assessing benefits, costs, and disparate racial impacts of confrontational proactive policing. *Proceedings of the National Academy of Sciences*, 114(35), 9308–931
- Moore, T. (2021). *Rowan and Mayne: A biography of the first London Police Commissioners*. Mango Books.
- National Academies of Sciences, Engineering, and Medicine. (2018). *Proactive policing: Effects on crime and communities*. The National Academies Press.
- Navratil, L., Bierschbach, B., & Faircloth, R. (2021, November 2). Minneapolis voters reject plan to replace Police Department. *Minneapolis StarTribune*. <https://www.startribune.com/minneapolis-voters-reject-plan-to-replace-police-department/600112156/>
- Nawaz, A., & Tankebe, J. (2018). Tracking procedural justice in stop and search encounters: Coding evidence from body-worn video cameras. *Cambridge Journal of Evidence-Based Policing*, 2(3-4), 139–163.
- New York Civil Liberties Union. (2019). Annual report. <https://www.nyclu.org/en/stop-and-frisk-data>
- Neyroud, P. (2018). *Out of court disposals managed by the police: A review of the evidence*. Commissioned by the National Police Chief’s Council of England and Wales. University of Cambridge. <https://www.npcc.police.uk/Publication/NPCC%20Out%20of%20Court%20Disposals%20Evidence%20assessment%20FINAL%20June%202018.pdf>
- Piza, E. L., & Connealy, N. T. (2022). The effect of the Seattle Police-Free CHOP zone on crime: A microsynthetic control evaluation. *Criminology & Public Policy*. <https://doi.org/10.1111/1745-9133.12570>
- Pollack, H., & Humphreys, K. (2020). Reducing violent incidents between police officers and people with psychiatric or substance use disorders. *Annals of the American Academy of Political and Social Science*, 687(1), 166–184.
- Pyle, K. (1918). Goldilocks and the three bears. Reprinted at Project Gutenberg. [https://www.gutenberg.org/files/49001/49001-h/49001-h.htm#Page\\_207](https://www.gutenberg.org/files/49001/49001-h/49001-h.htm#Page_207)
- Repetto, T. (2012). *American police: A history. Vol. II 1945–2012*. Enigma Books.

- Rothwell, S. (2021). Tele-policing: Improving response to calls for service. Video Presentation at Evidence-Based Policing 2021 Conference, University of Cambridge. <https://www.youtube.com/watch?v=2J3bP1SpzU>
- Rubinstein, D., & Mays, J. (2020, June 30). Nearly \$1 billion is shifted from police in budget that pleases no one. *New York Times*. <https://www.nytimes.com/2020/06/30/nyregion/nypd-budget.html>
- Schnohr, P., Okeefe, J. H., Lavie, C. J., Holtermann, A., Lange, P., Jensen, G. B., & Marott, J. L. (2021). U-Shaped association between duration of sports activities and mortality: Copenhagen City heart study. *Mayo Clinic Proceedings*, 96(12), 3012–3020.
- Sherman, L. W. (1980). Execution without trial: Police homicide and the Constitution. *Vanderbilt Law Review*, 33, 71–100.
- Sherman, L. W. (1983). Reducing police gun use: Critical events, administrative policy and organizational change. In M. Punch (Ed.), *Control in the police organization* (pp. 98–125). MIT Press.
- Sherman, L. W. (1992). Attacking crime: Policing and crime control. *Crime and Justice: A Review of Research*, 15, 159–230.
- Sherman, L. W. (1998). *Evidence-based policing*. Police Foundation. <https://www.policefoundation.org/wp-content/uploads/2015/06/Sherman-1998-Evidence-Based-Policing.pdf>
- Sherman, L. W. (2002). Fair and effective policing. In J. Q. Wilson & J. Petersilia (Eds.), *Crime: Public policies for crime control* (pp. 383–434). San Francisco: ICS Press.
- Sherman, L. W. (2007). The power few: Experimental criminology and the reduction of harm. *Journal of Experimental Criminology*, 3(4), 299–321.
- Sherman, L. W. (2009). Evidence and liberty: The promise of experimental criminology. *Criminology & Criminal Justice*, 9(1), 5–28.
- Sherman, L. W. (2013). The rise of evidence-based policing: Targeting, testing, and tracking. *Crime and Justice*, 42(1), 377–451.
- Sherman, L. W. (2018). Reducing fatal police shootings as systems crashes: Research, theory and practice. *Annual Review of Criminology*, 1, 421–449.
- Sherman, L. W. (2022). Three tiers for evidence-based policing: Targeting “minimalist” policing with a risk-adjusted disparity index. In D. Weisburd, T. Jonathan, B. Hasasi, & G. Perry (Eds.), *Evidence-based policing* (forthcoming). Cambridge University Press.
- Sherman, L. W., Neyroud, P. W., & Neyroud, E. (2016). The Cambridge crime harm index: Measuring total harm from crime based on sentencing guidelines. *Policing: A Journal of Policy and Practice*, 10(3), 171–183.
- Sherman, L. W., & Kumar, S. (2021). Equal protection by race with stop and frisk: A Risk-Adjusted Disparity (RAD) Index for balanced policing. *Cambridge Journal of Evidence-Based Policing*, 5(1), 1–19.
- Sherman, L. W., & Rogan, D. P. (1995). Effects of gun seizures on gun violence: “Hot spots” patrol in Kansas City. *Justice Quarterly*, 12(4), 673–693.
- Taleb, N. N. (2007). *The black swan: The impact of the highly improbable*. Random House.
- Tankebe, J. (2013). Viewing things differently: The dimensions of public perceptions of police legitimacy. *Criminology*, 51(1), 103–135
- Thompson, J. (2021, October 24). A ‘shoot to incapacitate’ policy puts Georgia police chief and town in the spotlight. *The Washington Post*. [https://www.washingtonpost.com/national/a-shoot-to-incapacitate-policy-puts-georgia-police-chief-and-town-in-the-spotlight/2021/10/24/d64b86f4-3378-11ec-9241-aad8e48f01ff\\_story.html](https://www.washingtonpost.com/national/a-shoot-to-incapacitate-policy-puts-georgia-police-chief-and-town-in-the-spotlight/2021/10/24/d64b86f4-3378-11ec-9241-aad8e48f01ff_story.html)
- Vitale, A. S. (2017). *The end of policing*. Verso Books.
- Weinborn, C., Ariel, B., Sherman, L. W., & O’Dwyer, E. (2017). Hotspots vs. harmspots: Shifting the focus from counts to harm in the criminology of place. *Applied Geography*, 86, 226–244.
- Wilson, M. (2022). How a Calm Call for Help Led 2 N.Y.P.D. Officers Into a Hail of Gunfire. *New York Times*, Jan. 22. <https://www.nytimes.com/2022/01/22/nyregion/nypd-officers-shot-harlem.html>
- Wheeler, C., & Griffiths, S. (2022, January 9). Universities have no excuse for not teaching face-to-face, says Nadhim Zahawi. *The Times*. <https://www.thetimes.co.uk/article/universities-no-excuse-teaching-face-to-face-masks-nadhim-zahawi-l2zgntz5b>
- Zhao, L., & Papachristos, A. V. (2020). Network position and police who shoot. *Annals of the American Academy of Political and Social Science*, 687(1), 89–112.
- Zimring, F. E. (2012). *The city that became safe: New York’s lessons for urban crime and its control*. Oxford University Press.
- Zimring, F. E. (2017). *When police kill*. Harvard University Press.

**How to cite this article:** Sherman, L. W. (2022). Goldilocks and the three “Ts”: Targeting, testing, and tracking for “just right” democratic policing. *Criminol Public Policy*, 21, 175–196. <https://doi.org/10.1111/1745-9133.12578>